Cunningham's

TEXTBOOK of VETERINARY PHYSIOLOGY

SIXTH EDITION



BRADLEY G. KLEIN

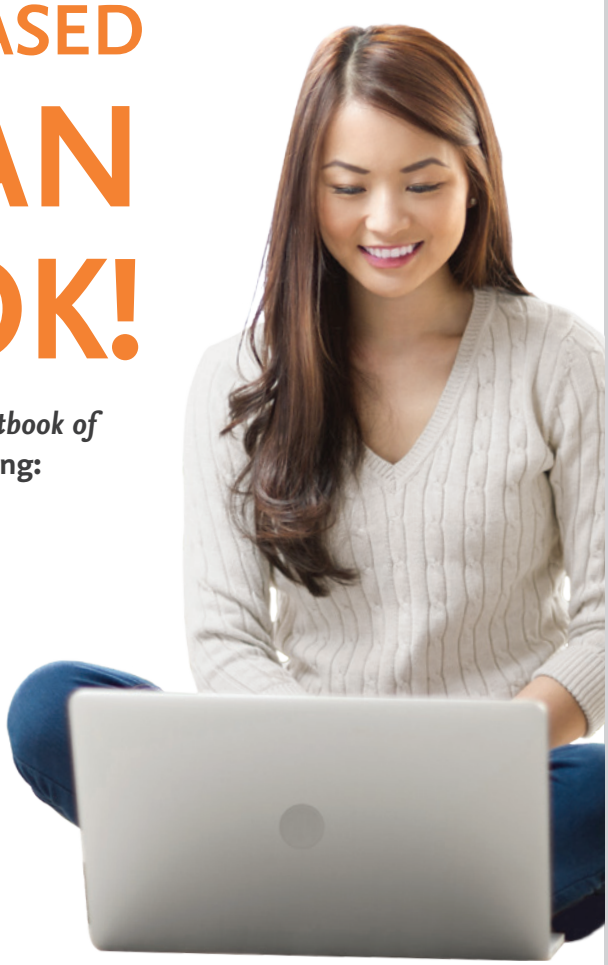
Evolve®

Student Resources on Evolve
Access Code Inside

ELSEVIER

Evolve®

YOU'VE JUST PURCHASED MORE THAN A TEXTBOOK!



Evolve Student Resources for *Klein: Cunningham's Textbook of Veterinary Physiology, Sixth Edition*, include the following:

- Animations
- Clinical Correlations
- Glossary
- Practice Tests

Activate the complete learning experience that comes with each NEW textbook purchase by registering with your scratch-off access code at

http://evolve.elsevier.com/Klein/physiology/

If you rented or purchased a used book and the scratch-off code at right has already been revealed, the code may have been used and cannot be re-used for registration. To purchase a new code to access these valuable study resources, simply follow the link above.

Place Sticker Here

REGISTER TODAY!

ELSEVIER

You can now purchase Elsevier products on Evolve!
Go to evolve.elsevier.com/shop to search and browse for products.

2018v1.0

Cunningham's Textbook of

VETERINARY PHYSIOLOGY

Sixth Edition

Bradley G. Klein, PhD

Associate Professor of Neuroscience
Department of Biomedical Sciences and Pathobiology
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

ELSEVIER

ELSEVIER

3251 Riverport Lane
St. Louis, Missouri 63043

CUNNINGHAM'S TEXTBOOK OF VETERINARY PHYSIOLOGY, SIXTH EDITION

ISBN: 978-0-323-55227-1

Copyright © 2020 by Elsevier, Inc. All rights reserved.

Previous edition copyrighted 2013, 2007, 2002, 1997, 1992.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds or experiments described herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. To the fullest extent of the law, no responsibility is assumed by Elsevier, authors, editors or contributors for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN 978-0-323-55227-1

Senior Content Strategist: Jennifer Flynn-Briggs
Senior Content Development Manager: Ellen Wurm-Cutter
Content Development Specialist: Kathleen Nahm
Publishing Services Manager: Deepthi Unni
Senior Project Manager: Kamatchi Madhavan
Project Manager: Haritha Dharmarajan
Designer: Bridget Hoette

Printed in China

Last digit is the print number 9 8 7 6 5 4 3 2 1

Working together to grow libraries in developing countries
www.elsevier.com • www.bookaid.org

Vet-ebooks.com

In loving and lasting memory of Ira Matthew Klein (1985-2013).

Vet-ebooks.com

Contributors

S. Ansar Ahmed, DVM, PhD
Professor and Associate Dean for Research
Department of Biomedical Sciences and Pathobiology
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Steven P. Brinsko, DVM, MS, PhD, DACT
Professor and Associate Department Head of Academics
Department of Large Animal Clinical Sciences
College of Veterinary Medicine and Biomedical Sciences
Texas A&M University
College Station, Texas

Autumn P. Davidson, DVM, MS, DACVIM (Internal Medicine)
Clinical Professor
Veterinary Medical Teaching Hospital
Department of Medicine and Epidemiology
School of Veterinary Medicine
University of California-Davis
Davis, California

Susan L. Ewart, DVM, PhD, DACVIM
Professor and Special Assistant to the Dean
Department of Large Animal Clinical Sciences
College of Veterinary Medicine
Michigan State University
East Lansing, Michigan

Deborah S. Greco, DVM, PhD, DACVIM-SAIM
Senior Research Scientist
Nestle Purina Petcare
St. Louis, Missouri

Steven R. Heidemann, PhD
Professor Emeritus
Department of Physiology
Michigan State University
East Lansing, Michigan

Thomas H. Herdt, DVM, MS, DACVIM, DACVN
Professor and Chief of Nutrition
Department of Large Animal Clinical Sciences and Veterinary Diagnostic Laboratory
College of Veterinary Medicine
Michigan State University
East Lansing, Michigan

Bradley G. Klein, PhD
Associate Professor of Neuroscience
Department of Biomedical Sciences and Pathobiology
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Xin M. Luo, MS, PhD
Associate Professor of Immunology
Department of Biomedical Sciences and Pathobiology
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Brian K. Petroff, DVM, PhD
Associate Professor and Section Chief, Endocrinology
Department of Pathobiology and Diagnostic Investigation
Diagnostic Center for Population and Animal Health
College of Veterinary Medicine
Michigan State University
East Lansing, Michigan

Juan E. Romano, DVM, MS, PhD, DACT
Professor
Department of Large Animal Clinical Sciences
College of Veterinary Medicine and Biomedical Sciences
Texas A&M University
College Station, Texas

John H. Rossmeisl Jr., DVM, MS, DACVIM
Professor
Neurology and Neurosurgery
Department of Small Animal Clinical Sciences
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Gerhardt G. Schurig, DVM, MS, PhD
Professor and International Programs Strategist
Department of Biomedical Sciences and Pathobiology
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Vet-ebooks.com

George H. Stabenfeldt, DVM, PhD (deceased)
Professor
Department of Reproduction
School of Veterinary Medicine
University of California-Davis
Davis, California
(Contributor of past editions)

Robert B. Stephenson, PhD
Associate Professor
Department of Physiology
Michigan State University
East Lansing, Michigan

Jill W. Verlander, DVM
Scientist
Department of Medicine
Division of Nephrology, Hypertension, and Renal Transplantation
University of Florida College of Medicine
Gainesville, Florida

Sharon G. Witonsky, DVM, PhD, DACVIM
Associate Professor
Equine Field Service
Department of Large Animal Clinical Sciences
Virginia-Maryland College of Veterinary Medicine
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Vet-ebooks.com

Preface

Physiology is the study of the normal functions of the body—the study of the body's molecules, cells, and organ systems, and the interrelationships among them. Because the study of medicine is the study of abnormal functions of the body, it is essential to understand normal physiology if one is to understand the mechanisms of disease. For this reason, physiology and other important sciences basic to medicine are introduced early in the veterinary curriculum.

Physiology is a vast subject, and the normal time constraints on veterinary students are usually prohibitive for learning all that is known about it. Therefore an effort was made to limit the concepts presented in this book to those germane to the practice of veterinary medicine. Because the scope of physiology encompasses many scientific disciplines and levels of analysis, the contributing authors represent not only the field of physiology but others, among them neuroscience, cell biology, molecular biology, immunology, and various clinical specialties. Most of these authors are also veterinarians or have consulted with veterinary clinicians regarding content. Sections on the immune system and cancer underscore the intimate relationship between the understanding of cell and molecular biology, physiological function, and veterinary medicine.

This book is designed for first-year veterinary students. The goal is to introduce the student to the principles and concepts of physiology pertinent to the practice of veterinary medicine. Other goals are to introduce the reader to physiopathology and clinical problem-solving techniques and to help the reader understand the relationship between physiology and the practice of veterinary medicine.

This book is designed to be as student friendly as possible. New concepts in the text are introduced by a declarative statement designed to summarize the essential point. This format also helps the reader survey the chapter or review for an examination. These declarative statements are also listed at the beginning of the chapter as an outline of Key Points.

Chapters include Clinical Correlations at the end. Concordant with the increasing trend toward early integration of basic science and clinical material in veterinary curricula, these Clinical Correlations are designed to show the reader how knowledge of physiology is applied to the diagnosis and treatment of veterinary patients. They also provide the student with an additional way to think through the principles and concepts presented, and they can serve as a basis for classroom case discussions.

Several Practice Questions are included in each chapter as another method for students to review the book's content. The Bibliography for each chapter is designed to lead the reader to more advanced textbooks, as the aforementioned time constraints can often prevent veterinary students from reading original literature. However, for those who may find the time, some original literature references are also included in several chapters.

Accompanying resources for the text can be found on Elsevier's Evolve website. These include an online repository of the book's Practice Questions and Clinical Correlations, as well as some relevant animations from Elsevier's existing collection. Students will appreciate the ability to access the Practice Questions as Practice Tests. Instructors will appreciate the items in the illustration bank, which can be downloaded for lecture use. There is also a Glossary covering many of the italicized words from the printed text.

Immense gratitude is due to the contributing authors who have worked hard to ensure that the information in this latest edition is accurate and up-to-date. This has resulted in new illustrations and Clinical Correlations, as well as addition of material on new discoveries (e.g., a brain lymphatic system; neutrophil extracellular traps). The expertise of two additional section authors is welcomed, Dr. Susan L. Ewart for Respiratory Function and Homeostasis, and Dr. Brian K. Petroff for Endocrinology. Dr. Xin M. Luo contributed some material to the Immunology section. Notable recognition is in order for Dr. N. Edward Robinson, who is retiring from this project after contributing to all of its previous editions, and who authored the initial material on respiratory physiology and on homeostasis. Thanks also to Dr. Ayman Sayegh for contributing to an earlier edition of the introductory chapter on regulation of gastrointestinal function. Suggestions of ways to improve this text in subsequent editions are always welcome.

Particular gratitude is due to the book's medical illustrator, Jeanne Robertson, who drew the new illustrations for this edition and revised much of the preexisting artwork. Thanks are also in order for the folks at Elsevier who were instrumental in producing the sixth edition, among them Penny Rudolph, Jennifer Flynn-Briggs, Ellen Wurm-Cutter, Anna Miller, Madhavan Kamatchi, and Haritha Dharmarajan. Special recognition is due for Kathleen Nahm, the Elsevier content development specialist who shepherded almost all aspects of this project through various trials and tribulations and who did an admirable job managing my not infrequent tardiness yet otherwise fastidious nature. Drs. Jonathan Abbott, Virginia Buechner-Maxwell, and Shireen Hafez unselfishly provided their valuable opinions on various aspects of the book that resulted in its improvement. Dr. Bonnie Smith was particularly helpful in providing advice on the accuracy of the cover illustration. Again, this book would not exist without the invaluable expertise of the section authors who worked so hard to make this the best veterinary physiology text possible.

The most immense debt is due to Dr. Jim Cunningham, whose vision, guidance, and expertise made the *Textbook of Veterinary Physiology* a reality and a success, and has made "Cunningham's" a household name among veterinary students and faculty of several generations. Placing his name in the book's title is in partial recognition of that debt. He is also recognized here for authoring the initial material on neurophysiology upon which I have built over the years. The instructional style he instituted continues in this edition and will continue in future editions of the text.

I would also like to thank my family whose love and support sustains my existence, and Ira, we miss you so very much. You are always in our hearts and our memories.

Finally, this book is a tribute to the veterinary students throughout the world. It is these students who give pleasure, meaning, and value to our teaching.

Brad Klein

Vet-ebooks.com

Contents

Section I: The Cell

1 The Molecular and Cellular Bases of Physiological Regulation, 1
Steven R. Heidemann

Clinical Correlations:
Peripheral Edema, 27
Horse With Hyperkalemic Periodic Paralysis, 27

2 Cancer: A Disease of Cellular Proliferation, Life Span, and Death, 29
Steven R. Heidemann

Clinical Correlations:
Dog That Collapsed While Running, 48
8-Year-Old Male Neutered Boxer with Seizure, 49

Section II: Neurophysiology

3 Introduction to the Nervous System, 51
Bradley G. Klein

Clinical Correlations:
Neurological Disease in a Horse, 55
Neurological Signs in 10-Year-Old Male Neutered Pembroke Welsh Corgi, 55

4 The Neuron, 57
Bradley G. Klein

Clinical Correlations:
Hypoglycemia, 63
Salt Toxicity in Pot Belly Pig, 63

5 The Synapse, 66
Bradley G. Klein

Clinical Correlations:
Myasthenia Gravis, 71
Tetanus, 71

6 The Physiology of Muscle, 73
Bradley G. Klein

Clinical Correlations:
Down Cow After Calving, 80
Malignant Hyperthermia, 80

7 The Concept of a Reflex, 82
Bradley G. Klein

Clinical Correlations:
Trauma in a Foal, 85

8 Skeletal Muscle Receptor Organs, 86
Bradley G. Klein

Clinical Correlations:
Femoral Nerve Mononeuropathy, 91
Obturator Nerve Paralysis in a Cow Post-Calving, 91

9 The Concept of Lower and Upper Motor Neurons and Their Malfunction, 93
Bradley G. Klein

Clinical Correlations:
Lower Motor Neuron Disease, 95
Upper Motor Neuron Disease, 95

10 The Central Control of Movement, 97
Bradley G. Klein

Clinical Correlations:
Focal Lesion of the Motor Cortex, 107
Cow With Hypomagnesemia, 107

11 The Vestibular System, 109
Bradley G. Klein

Clinical Correlations:
Vestibular Syndrome in a Dog, 116
Horse With Vestibular Disease, 116

12 The Cerebellum, 118
Bradley G. Klein

Clinical Correlations:
Cerebellar Hypoplasia, 124
Newborn Calf Unable to Rise, 124

13 The Autonomic Nervous System, 126
Bradley G. Klein

Clinical Correlations:
Horner's Syndrome, 135
Colic From Administration of Atropine to Treat Corneal Ulcer, 136

14 The Visual System, 138
Bradley G. Klein

Clinical Correlations:
Homonymous Hemianopia, 147
Uveitis in a Horse, 147

Vet-ebooks.com

15 Cerebrospinal Fluid and the Blood-Brain Barrier, 149
Bradley G. Klein
Clinical Correlations:
Increased Intracranial Pressure, 155
Seizures in a Foal, 156

16 The Electroencephalogram and Sensory-Evoked Potentials, 158
Bradley G. Klein
Edited by John H. Rossmeisl, Jr.
Clinical Correlations:
Brain Tumor, 164
Outbreak of Neurological Disease in Goats, 164

17 Hearing, 166
Bradley G. Klein
Clinical Correlations:
Congenital Deafness, 171
Deafness in American Paint Horse Associated With Genetic Defect, 171

Section III: Cardiovascular Physiology

18 Overview of Cardiovascular Function, 173
Robert B. Stephenson
Clinical Correlations:
Lethargic Kid Goat, 184
Colic and Endotoxic Shock in Horse Secondary to Strongylus Parasitism, 184

19 Electrical Activity of the Heart, 186
Robert B. Stephenson
Clinical Correlations:
Third-Degree Atrioventricular Block, 201

20 The Electrocardiogram, 203
Robert B. Stephenson
Clinical Correlations:
Dilative Cardiomyopathy With Paroxysmal Atrial Tachycardia, 213
Race Horse Fades During Workout, 213

21 The Heart as a Pump, 216
Robert B. Stephenson
Clinical Correlations:
Pulmonic Stenosis, 226
Older Horse With Exercise Intolerance, 227

22 The Systemic and Pulmonary Circulations, 230
Robert B. Stephenson
Clinical Correlations:
Canine Heartworm Disease With Pulmonary Embolism, 239
Dummy Foal: Hypoxemic Ischemic Encephalopathy, 239

23 Capillaries and Fluid Exchange, 241
Robert B. Stephenson
Clinical Correlations:
Acute Protein-Losing Enteropathy in a Horse, 250
Clinical Examination, 250
Comment, 250
Treatment, 250
Vasculitis in a Mare, 251

24 Local Control of Blood Flow, 253
Robert B. Stephenson
Clinical Correlations:
Patent Ductus Arteriosus, 258
Endotoxemia in a Foal, 259

25 Neural and Hormonal Control of Blood Pressure and Blood Volume, 262
Robert B. Stephenson
Clinical Correlations:
Intraoperative Hemorrhage, 269
Three-Day-Old Heifer Calf With Diarrhea, 269

26 Integrated Cardiovascular Responses, 272
Robert B. Stephenson
Clinical Correlations:
Exercise Intolerance Secondary to Congestive Heart Failure, 279
Cow With "Hardware Disease", 280

Section IV: Physiology of the Gastrointestinal Tract

27 Regulation of Gastrointestinal Function, 284
Thomas H. Herdt

28 Motility Patterns of the Gastrointestinal Tract, 293
Thomas H. Herdt
Clinical Correlations:
Equine Rabies, 304
Horse With Colic Due to Large Colon Displacement, 305

29 Secretions of the Gastrointestinal Tract, 307
Thomas H. Herdt
Clinical Correlations:
Horse in Pain With Weight Loss, 314
Pancreatitis in a Dog, 314

30 Digestion and Absorption: The Nonfermentative Processes, 316
Thomas H. Herdt
Clinical Correlations:
Diarrhea With Dehydration and Acidosis in a Calf, 337
Juvenile Pancreatic Atrophy in a Dog, 337

Vet-ebooks.com

31 Digestion: The Fermentative Process, 339
Thomas H. Herdt
Clinical Correlations:
Grain Engorgement Toxemia, 359
Impaction Colic, 359

32 Postabsorptive Nutrient Utilization, 361
Thomas H. Herdt
Clinical Correlations:
Hepatic Lipidosis in a Cat, 375
Hyperlipemia in a Horse, 376

Section V: Endocrinology

33 The Endocrine System, 378
Brian K. Petroff and Deborah S. Greco
Clinical Correlations:
Equine Pituitary Pars Intermedia Dysfunction (PPID), 392
Agalactic Mare, 393

34 Endocrine Glands and Their Function, 394
Brian K. Petroff and Deborah S. Greco
Clinical Correlations:
Diabetes Mellitus, 426
Pancreatic Tumor in a German Shepherd, 426

Section VI: Reproduction and Lactation

35 Reproductive System Development and Sexual Differentiation, 428
Autumn P. Davidson and George H. Stabenfeldt†
Clinical Correlations:
Androgen Insensitivity, 434
Ovarian Remnant Syndrome, 435

36 The Ovarian Cycle, 436
Autumn P. Davidson and George H. Stabenfeldt†
Clinical Correlations:
Pseudopregnancy in the Bitch Presented for Infertility, 440
Persistent Luteal Phase in the Mare, 440

37 The Estrous/Menstrual Cycle, 442
Autumn P. Davidson and George H. Stabenfeldt†
Clinical Correlations:
Sexual Attractiveness in the Spayed Bitch, 448
Trying to Get a Mare Pregnant, 448

38 Pregnancy and Parturition, 450
Autumn P. Davidson and George H. Stabenfeldt†
Clinical Correlations:
Prolonged Gestation in the Cow, 456
Prolonged Gestation in the Bitch, 456

39 The Mammary Gland and Lactation, 458
Autumn P. Davidson and George H. Stabenfeldt†
Clinical Correlations:
Pregnant Mare That Does Not Have Sufficient Milk or Udder Development, 469
Neonatal Isoerythrolysis, 469

40 Reproductive Physiology of the Male, 471
Juan E. Romano and Steven P. Brinsko
Clinical Correlations:
Infertility in a Stallion, 478
Infertility in a Bull, 479

Section VII: Renal Physiology

41 Glomerular Filtration, 480
Jill W. Verlander
Clinical Correlations:
Chronic Renal Failure, 487
Glomerulonephritis, 487

42 Solute Reabsorption, 489
Jill W. Verlander
Clinical Correlations:
Glucosuria, 499
Hypoadrenocorticism, 499

43 Water Balance, 501
Jill W. Verlander
Clinical Correlations:
Diabetes Insipidus, 507
Chronic Renal Insufficiency, 507

44 Acid-Base Balance, 509
Jill W. Verlander
Clinical Correlations:
Respiratory Acidosis With Renal Compensation, 516
Metabolic Alkalosis With Paradoxical Aciduria, 516

Section VIII: Respiratory Function

45 Overview of Respiratory Function: Ventilation of the Lungs, 518
Susan L. Ewart
Clinical Correlations:
Lung Fibrosis in a Dog, 530
Chronic Airway Disease in a Horse, 530

46 Blood Flow Through the Lungs, 532
Susan L. Ewart
Clinical Correlations:
Bovine High-Altitude ("Brisket") Disease in a Heifer, 538
Stallion With Bilateral Epistaxis Postracing, 538

†Deceased

Vet-ebooks.com

47 Gas Exchange, 540
Susan L. Ewart
Clinical Correlations:
Hypoventilation in an English Bulldog, 548
Hypoxemia in an Anesthetized Clydesdale Horse, 548

48 Gas Transport in the Blood, 550
Susan L. Ewart
Clinical Correlations:
Flea Infestation in a Cat, 558
Atrial Fibrillation in a Horse, 558

49 Control of Ventilation, 560
Susan L. Ewart
Clinical Correlations:
Hypoxemia With Hyperventilation in a Samoyed Puppy, 566
Hypoventilation in an Anesthetized Saint Bernard, 566

50 Nonrespiratory Functions of the Lungs, 568
Susan L. Ewart
Clinical Correlations:
Pleuritis in a Thoroughbred Horse, 573
Mitral Insufficiency in a Dog, 574

Section IX: Homeostasis

51 Fetal and Neonatal Oxygen Transport, 576
Susan L. Ewart
Clinical Correlations:
Patent Ductus Arteriosus in a Pomeranian, 581
Pregnant Mare With Placentitis, 581
Dyspneic Foal, 582

52 Acid-Base Homeostasis, 584
Susan L. Ewart
Clinical Correlations:
Upper Airway Obstruction in a Boston Terrier, 592
Torsion of the Abomasum in a Cow, 593
Neonatal Diarrhea in a Foal, 594

53 Thermoregulation, 596
Susan L. Ewart
Clinical Correlations:
Influenza in Pigs, 606
Heat Stroke in a Boxer Dog, 606

Section X: The Immune System

54 Antigens and Innate Immunity, 608
S. Ansar Ahmed and Gerhardt G. Schurig
Clinical Correlations:
Swollen Lymph Nodes in a Colt, 615
Heifer Calf With Recurrent Infection, 615

55 The Specific Immune Response: Acquired Immunity, 617
S. Ansar Ahmed and Gerhardt G. Schurig
Contributions by Xin M. Luo
Clinical Correlations:
Unthrifty Foal, 624
Swelling Under Jaw in 20-Year-Old Rocky Mountain Horse Gelding, 625

Appendix A, Answers to Practice Questions, 627

Clinical Correlations
Edited by Sharon G. Witonsky

Index, 628

Vet-ebooks.com

1

The Molecular and Cellular Bases of Physiological Regulation

STEVEN R. HEIDEMANN

KEY POINTS

1. All physiological change is mediated by proteins.
2. Protein function depends on protein shape and shape changes.
3. A series of enzymatic reactions converts tyrosine into the signaling molecules dopamine, norepinephrine, and epinephrine.
4. Muscle contraction and its initiation and cessation depend on the binding specificity and allosteric properties of proteins.
5. Biological membranes are a mosaic of proteins embedded in a phospholipid bilayer.

Transport

1. Only small, uncharged molecules and oily molecules can penetrate biomembranes without the aid of proteins.
2. Molecules move spontaneously from regions of high free energy to regions of lower free energy.
3. Important transport equations summarize the contributions of the various driving forces.
4. Starling's hypothesis relates fluid flow across the capillaries to hydrostatic pressure and osmotic pressure.
5. Membrane proteins that serve the triple functions of selective transport, catalysis, and coupling can pump ions and molecules to regions of higher free energy.
6. Many membrane proteins selectively facilitate the transport o f ions/molecules from high to low electrochemical potential.
7. Passive transport of K^+ across the plasma membrane creates an electrical potential.
8. Spatial organization of active and passive transport proteins enables material to pass completely through the cell.
9. Membrane fusion allows for a combination of compartmentalization and transport of material.

Information Transmission and Transduction

1. Cell signaling often occurs by a lengthy chain of sequential molecular interactions.
2. Signaling pathways begin with the binding of an extracellular molecule to a receptor.
3. Specific physiological information is inherent in the receptor/ligand complex, not in the hormone/neurotransmitter molecule.
4. G-protein–coupled receptors are the largest family (a *superfamily*) of receptors and help regulate almost all physiological processes.
5. Most G-protein–linked information is sent to the cytoplasm by *second messengers*.
6. Ca^{2+} transport across plasma and intracellular membranes is an important second messenger.
7. Cyclic adenosine monophosphate (AMP) is produced by activation of a membrane-bound enzyme in response to hormone/neurotransmitter binding to receptors.
8. The receptor-mediated hydrolysis of a rare phospholipid of the plasma membrane produces two different second messengers with different actions.
9. Steroid hormones and other lipid signals interact with nuclear receptors, which are transcription factors within the cell.

Physiology is the study of the regulation of change within organisms, in this case higher animals. Our understanding of physiology has changed dramatically in the past 35 years as a result of insight into the molecular basis of biological regulation. This chapter summarizes (and simplifies considerably) our current understanding of the molecular and cellular basis of that regulation. Most of the principles in this chapter apply to all animal cells. The approach taken is one of functional molecular anatomy. That is, the molecular structure of the cell is examined with particular emphasis on the physiological function, in the intact animal, of the molecules and supramolecular structures responsible for the function. Only those aspects of cell function that illuminate the medical physiology of the higher animals are discussed; the reader is referred to the Bibliography for more complete coverage of the cell. Some review of basic concepts and vocabulary is presented. However, the discussion assumes that the reader is familiar with the cell and its constituent molecules as presented in courses in general biology and undergraduate courses in biochemistry.

Vet-ebooks.com

All Physiological Change Is Mediated by Proteins

All physiological change is mediated by a single class of polymeric macromolecules (large molecules), the proteins. Protein function can be subdivided into a number of categories: catalysis, reaction coupling, transport, structure, and signaling.

Catalysis is the ability to increase greatly the rate of a chemical reaction without altering the equilibrium of the reaction. The majority of biochemical reactions occur at a physiologically useful rate only because of protein catalysts, called *enzymes.* Examples of enzymatic catalysis in the synthesis of a class of physiological regulator molecules, catecholamines, are given later in this chapter.

In *reaction coupling,* two reactions are joined together with the transfer of energy. Energy from a spontaneous reaction (similar to water flowing downhill) is funneled to a nonspontaneous reaction (e.g., sawing wood) so that the sum of the two reactions is spontaneous. That is, the energy liberated by the "downhill" reaction is used to drive the "uphill" reaction. This is the basic function of a motor; the "downhill" burning of gasoline is coupled with the "uphill" movement of the car. The ability of proteins to couple spontaneous and nonspontaneous reactions allows cells to be chemical motors, using chemical energy to do various jobs of work. One such job of work, the contraction of striated muscle, is discussed later with particular emphasis on the proteins involved.

Proteins provide a pathway for the *membrane transport* of most molecules and all ions into and out of the cell. Transport and transport proteins are discussed more fully after a discussion of the lipid bilayer membrane, the major obstacle to transport.

Proteins that form filaments and glue cells to each other and to their environment are responsible for the *structure* and organization of cells and multicellular assemblies (i.e., the tissues and organs of animals). The internal structure of the muscle cell, as well as its ability to do work, is a result of the properties of the muscle proteins discussed later.

At its most basic level, *signaling* requires only a controlled change or difference. Human signaling occurs by way of open and closed electrical circuits (telegraphy), puffs of smoke in the air, and complex black marks on a contrasting background (numbers and letters). As discussed next, a fundamental property of proteins is the ability to change shape. The cell can use changes of protein shape directly to send signals, and the function of some proteins is purely informational. That is, all that some proteins do by changing shape is to transmit and transduce information. *Information* can be defined as "any difference that makes a difference," or more simply, any difference that regulates something. Catalysis, coupling, transport, structural, and signaling functions can be combined on individual protein molecules. As will become apparent, such multifunctional proteins carry out many important physiological functions. Also important is that a change in one or more of these protein functions can be used to carry information, to serve as a signal within the cell. Thus in addition to proteins specialized exclusively to carry information, changes in enzymatic activity or ion transport can also make a difference, transmitting information and triggering an appropriate response.

Protein Function Depends on Protein Shape and Shape Changes

Protein function is founded on two molecular characteristics: (1) proteins can bind to other molecules very specifically; and (2) proteins change shape, which in turn alters their binding properties and their function. The binding specificity of proteins is the result of their complex three-dimensional structure. Grooves or indentations on the surface of protein molecules, called *binding sites,* permit specific interactions with a molecule of a complementary shape, called the *ligand.* This complementary-shape mechanism underlying binding is similar to the shape interaction between a lock and key.

Several aspects of the lock-and-key analogy are worth noting. As with a lock, only a small part of the protein is engaged in binding. The binding is very specific; small changes in the shape of the binding site (keyhole) or the shape of the ligand (key) can cause major changes in protein (lock) behavior. Similar to the lock and key, the complementary-shape interaction serves a recognition function; only those molecules with the right shape affect protein function. This recognition function plays a primary role in information transfer. The protein recognizes a particular signal by binding to it, thus changing the protein's shape and its function. Unlike the majority of locks, however, proteins frequently have multiple binding sites for multiple ligands.

Thus the three-dimensional shape of a protein, its *conformation,* determines protein function. A major force that stabilizes protein conformation is the *hydrophobic interaction.* Oily, *hydrophobic* (water-fearing) amino acids tend to congregate in the middle of a protein away from water, whereas *hydrophilic* (water-loving) amino acids tend to be found on the protein's outer surface interacting with the abundant cellular water. The hydrophobic interaction is also important in stabilizing the interaction of proteins with the lipids of biological membranes, as will be discussed shortly. Protein shape is also stabilized by *hydrogen bonding* between polar amino acid pairs in the polypeptide (protein) chain.

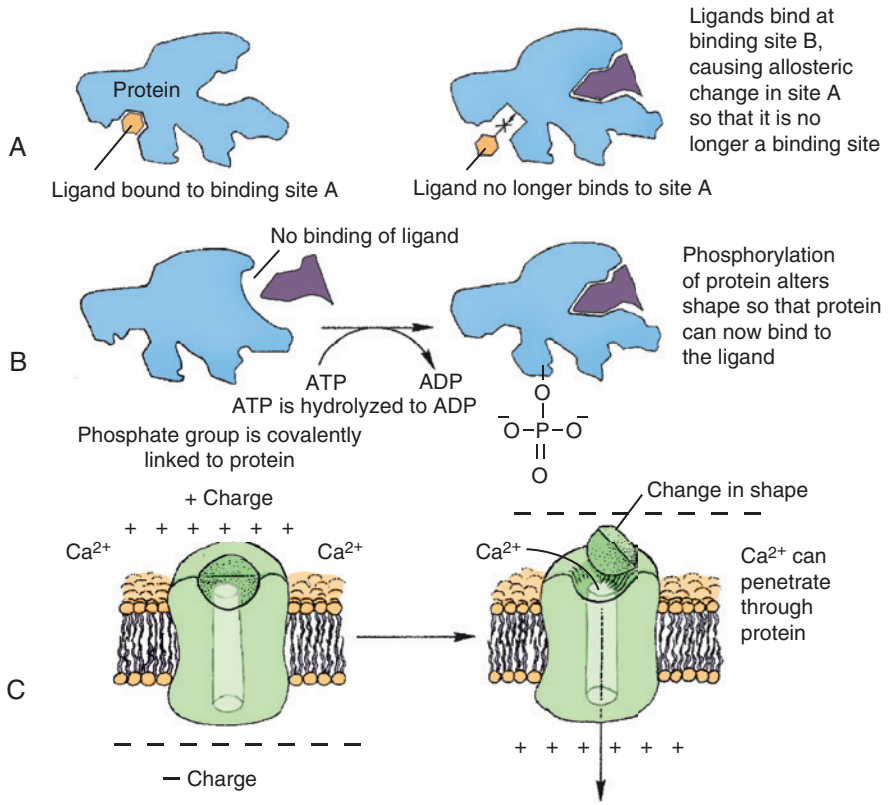
The same weak forces responsible for protein conformation are also used to hold the ligand in the protein-binding site. The position of the ligand in the binding site is stabilized by hydrogen bonds between the polar groups of the ligand and polar amino acid side groups lining the binding site, just as hydrogen bonds within the polypeptide chain stabilize the shape of the polypeptide. Precisely because the same forces are responsible for the shape of the protein and for its binding properties, shape influences binding, and in turn, binding can influence protein shape. The ability of proteins to change shape is called *allostery* (Greek for "other shape").

Allosteric changes in protein conformation arise in four general ways. One way, just mentioned, is that most proteins change shape depending on which ligands are bound at particular binding sites (Fig. 1.1A). The sequence—specific ligand binding → protein shape change → change in protein-binding properties and protein function → this change regulates something—is a common molecular mechanism underlying physiological control. This method involves no alteration in the covalent structure of the protein.

A second method of producing conformational change, however, occurs as a result of the covalent modification of one or more of the amino acid side groups of the protein (see Fig. 1.1B). By far the most common such change is the covalent addition of a phosphate group to the hydroxyl (–OH) group on the side chain of serine, threonine, or tyrosine residues in the protein. This modification is called *phosphorylation.* Because the phosphate group is highly charged, phosphorylation of a protein alters hydrogen bonding and other electrostatic interactions within the protein chain, altering its conformation and functional properties.

In a third method, some physiologically important proteins change shape in response to the electrical field surrounding the protein (see Fig. 1.1C). These respond to a voltage change

Vet-ebooks.com



• **Fig. 1.1** Three common mechanisms of allosteric shape change in proteins. (A) Ligand binding. Ligand binding to an allosteric site (site B) on a protein changes the protein's conformation such that binding site A is altered; ligand no longer binds at site A because of the binding event at site B. (B) Phosphorylation. Addition of a phosphate group to a serine, threonine, or tyrosine residue of a protein alters the protein's conformation, changing its binding characteristics. In this hypothetical example, phosphorylation activates an otherwise inactive protein. Some proteins inactivate by this mechanism. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate. (C) Voltage-dependent proteins. The conformation of some proteins, particularly ion channels, is altered by the electrical field surrounding the protein. Shown here is the opening (activation) of a voltage-dependent, gated Ca^{2+} channel when the membrane depolarizes.

by altering the position of charged amino acids, thus altering protein shape.

The fourth method of protein shape change is the least well understood (not shown). Some proteins change shape in a controlled manner in response to mechanical forces. Although this is not surprising, because all solids and solid-like substances change shape at least slightly in response to force, we know relatively little about mechanosensitive proteins. The best current example is a protein involved in the very early events of hearing that changes its transport of ions in response to the mechanical stimulation by sound (small changes of air pressure in waves).

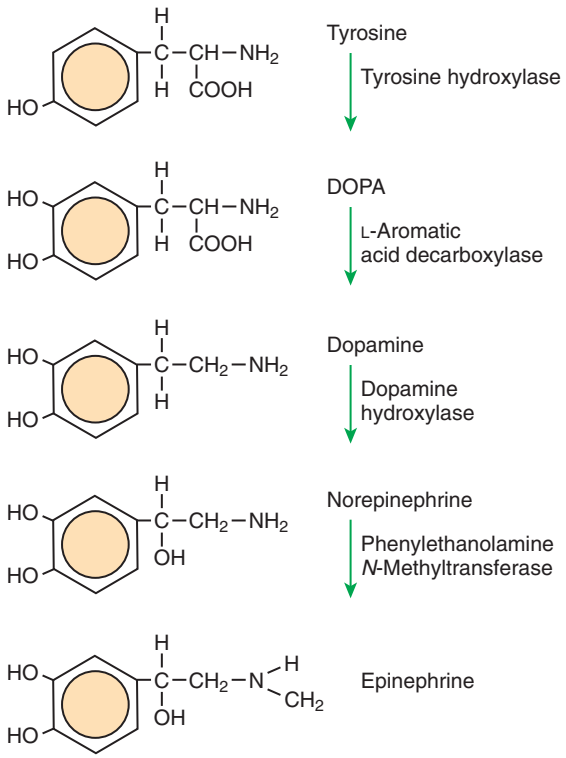
The significance of binding specificity and allostery can be better appreciated with two examples of their roles in physiological function. The first example is the role of enzymes in synthesizing three small, structurally similar, nonprotein signaling molecules. This example shows how binding specificity is important in catalytic function and how allostery underlies the regulation of the synthesis. The second example is more complex: the role of proteins in the contraction of muscle. The contraction of muscle shows how proteins can exploit the basic properties of specific binding and allosteric shape change to do more than one job of work at the same time; muscle proteins serve a structural role, serve a catalytic function, and couple the "downhill" hydrolysis of *adenosine triphosphate* (ATP) to do mechanical work, the "uphill" lifting of weight.

A Series of Enzymatic Reactions Converts Tyrosine Into the Signaling Molecules Dopamine, Norepinephrine, and Epinephrine

Fig. 1.2 is a diagram of the series of reactions by which the amino acid tyrosine is converted into three different signaling molecules: (1) *dopamine,* a brain neurotransmitter; (2) *norepinephrine,* a neurotransmitter of the brain and peripheral autonomic nervous system; and (3) *epinephrine,* an autonomic neurotransmitter and hormone. Dopamine, norepinephrine, and epinephrine share a similar structure. All contain a phenyl (benzene) ring with two hydroxyl groups (i.e., catechol) and an amine group (thus the term *catecholamines*). They are among the large number of molecules that function as neurotransmitters. That is, the electrically coded information sent along nerve cells causes the release of a chemical, the neurotransmitter, at the terminal of the neuron, which is next to a target cell, such as another nerve, a muscle, or an endocrine cell. The electrically encoded information of the nerve is transmitted to the target cell by the binding of the neurotransmitter to proteins on the surface of the target cell. Proper neurotransmitter synthesis is crucial to nervous function and physiological regulation.

In the first step of catecholamine biosynthesis, tyrosine binds to the enzyme tyrosine hydroxylase, which catalyzes the addition of

Vet-ebooks.com



• **Fig. 1.2** Epinephrine biosynthetic pathway. The amino acid tyrosine is metabolized to the neurotransmitters dopamine, norepinephrine, and epinephrine. The diagram shows the names and structural formulas for each compound in the path and the names of the enzymes that catalyze each reaction. *DOPA,* Dihydroxyphenylalanine.

another hydroxyl group to the phenyl group to form dihydroxyphenylalanine, almost always called *DOPA.* This hydroxyl group alters the enzyme-ligand interaction; the key no longer fits the keyhole. DOPA is released from the tyrosine hydroxylase and is then bound by another enzyme, L-aromatic amino acid decarboxylase. As the name implies, this enzyme catalyzes the removal of the carboxyl group, converting DOPA to dopamine. Dopamine is converted into norepinephrine by the activity of dopamine hydroxylase, which adds yet another hydroxyl group, this time to the two-carbon tail of dopamine. Finally, addition of a methyl group to the amino nitrogen by phenylethanolamine *N*-methyltransferase gives rise to epinephrine (also called *adrenalin*). Note the binding specificity of the enzymes; whereas the catecholamine structures are all similar to one another, different enzymes bind each one (e.g., epinephrine does not bind to dopamine hydroxylase).

The allosteric properties of one enzyme in this pathway provide an example of physiological regulation. Certain hormones and neurotransmitters cause the *phosphorylation* of tyrosine hydroxylase, the first enzyme in the pathway, increasing its activity. That is, phosphorylation of the enzyme increases the rate at which it catalyzes the conversion of tyrosine to DOPA. Because this step is the slowest in the pathway, an increase in the activity of this protein increases the net rate of synthesis of all the catecholamines. Regulated decreases in the rate of catecholamine synthesis are achieved by a different allosteric mechanism: binding of end products to the enzyme. Dopamine, norepinephrine, and epinephrine can all bind to tyrosine hydroxylase at a site different than the site for tyrosine. These binding events inhibit the enzymatic activity. The inhibition of the pathway by its own end products makes this a classic case of allosteric control called *end-product inhibition.* Many substances regulate their own synthesis by inhibiting an initial enzyme in the pathway. If the cell has enough end products, these products inhibit further synthesis by allosteric changes in the enzyme. This is an example of the following sequence: specific binding → protein shape change → change in protein-binding properties and protein function → this change regulates something.

Muscle Contraction and Its Initiation and Cessation Depend on the Binding Specificity and Allosteric Properties of Proteins

There are three types of muscle tissue in vertebrates: (1) *skeletal muscle,* responsible for the animal's ability to move; (2) *cardiac muscle,* a muscle type found only in the heart but structurally similar to skeletal muscle; and (3) *smooth muscle,* which surrounds hollow organs such as blood vessels, gut, and uterus. All three produce tensile force by contracting and shortening the length of the muscle. All muscle contraction occurs by the binding and the allosteric properties of two proteins, actin and myosin. Starting and stopping the contraction process depends on two additional proteins in skeletal and cardiac muscle, troponin and tropomyosin. Contraction initiation and cessation in smooth muscle depend on a different system with different proteins and are discussed later in this chapter.

Myosin is a large protein whose shape resembles a two-headed golf club. The elongated tail of the myosin molecule corresponds to the shaft of the golf club, and there are two knobs at one end of the tail that, as with golf clubs, are called *heads.* Myosin tails bind specifically to other myosin tails, forming bipolar aggregates called *thick filaments* (Fig. 1.3). Myosin heads specifically bind ATP and another muscle protein, actin. Actin binds to itself to form long, thin filaments called *thin filaments* in muscle and *F-actin* (filamentous actin) in other cell types. Actin filaments play an important architectural role in all animal cells. Although actin is best understood in muscle cells, all animal cells depend on actin filaments for their shape and for their capacity to migrate in their environment. Actin filaments can be "woven" in various ways to produce different structures, such as ropelike bundles and clothlike networks. These actin bundles and actin networks are used to support the cell in particular shapes, similar to ropes holding up the woven cloth of a tent.

In muscle, the interaction of myosin, ATP, and actin produces contraction and force, as shown in Fig. 1.4:

Step A: ATP binds to a myosin head; in this conformation, myosin has little ability to bind to actin.

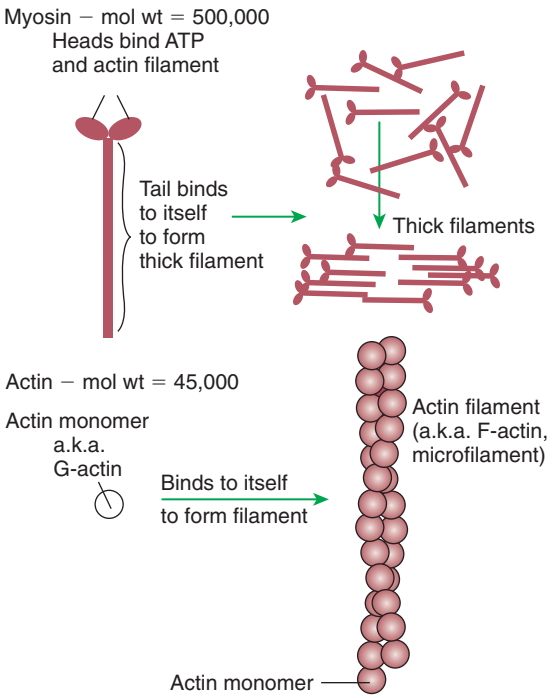
Step B: Enzymatic activity associated with the myosin head, an *adenosinetriphosphatase* (ATPase), rapidly causes a partial hydrolysis of ATP to *adenosine diphosphate* (ADP) and *inorganic phosphate* (P_i), both of which stay bound to the myosin. With ADP and P_i bound, myosin has a slightly different shape that binds avidly to nearby actin filaments.

Step C: When myosin binds to actin, called *cross-bridging,* the myosin head couples the complete hydrolysis of ATP to a forceful flexing of the myosin head. This allosteric change causes the actin filament to slide past the thick filament. This sliding puts the actin filament under tension, which in turn causes the muscle to contract (shorten) against the load of the muscle (i.e., lifting a weight or pumping out blood). All muscle contraction depends on this *sliding filament mechanism* of actin- and myosin-based filaments. This same

Vet-ebooks.com

allosteric change of myosin also alters myosin-binding properties so that it releases the ADP and P_i.

Step D: The binding of a new ATP molecule to the myosin head again causes myosin to change shape; the head unflexes and loses its affinity for actin, releasing the cross-bridge, and the cycle can start over. *Rigor mortis* of dead animals is caused by a lack of new ATP to bind to myosin heads. In the absence of ATP, myosin heads remain in Step C (i.e., bound to actin). The muscle is stiff because it is completely cross-bridged together.



• **Fig. 1.3** Assembly of myosin and actin to form filamentous structure. Myosin tails aggregate with one another to form a thick filament, a substructure of striated muscle. Actin monomers (G-actin) are a single polypeptide chain forming a globular protein that can bind to other actin monomers to form actin filament, also called microfilaments. The actin filament is the basic structure of striated muscle thin filaments; thin filaments also have troponin and tropomyosin as part of their structure. *ATP,* Adenosine triphosphate.

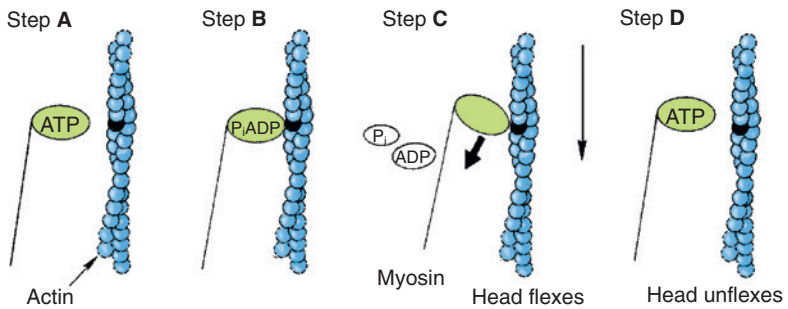
This *actomyosin motor* uses the binding and allosteric properties of proteins to (1) create structural filaments capable of withstanding and transmitting mechanical force, (2) catalyze the hydrolysis of ATP, and (3) couple the "downhill" ATP hydrolysis to the "uphill" contraction to produce force. For just the one protein, myosin, there are a number of examples of the characteristic sequence described earlier: specific binding → protein shape change → change in protein-binding properties and protein function → this change makes a difference.

This system of contractile proteins requires some control so that, for example, the heart beats rhythmically and skeletal muscle contraction is coordinated. At the organismal level, skeletal and cardiac muscle contraction is primarily under control by electrical stimulation from nerves or other electrically active cells (see Chapter 6). The transmission of electrical excitation to the actomyosin system is called *excitation-contraction coupling. Excitation-contraction coupling in all types of muscle depends on changes in intracellular calcium ion (Ca^{2+}) concentration.* In skeletal and cardiac muscle, but not smooth muscle, two additional thin-filament proteins, troponin and tropomyosin, are required for this coupling. (Excitation-contraction coupling for smooth muscle is discussed later in this chapter.) In striated muscles, *troponin* binds to tropomyosin and to Ca^{2+}. *Tropomyosin* is a long, thin protein that binds in the groove of the actin filament in such a way that its positions, high in the groove or snuggled down deep in the groove, allow or prevent the myosin head access to the thin filament (Fig. 1.5). Excitation-contraction coupling of striated muscle works as follows:

Step A: Electrical excitation of a striated muscle cell causes an increase in the intracellular concentration of Ca^{2+}.

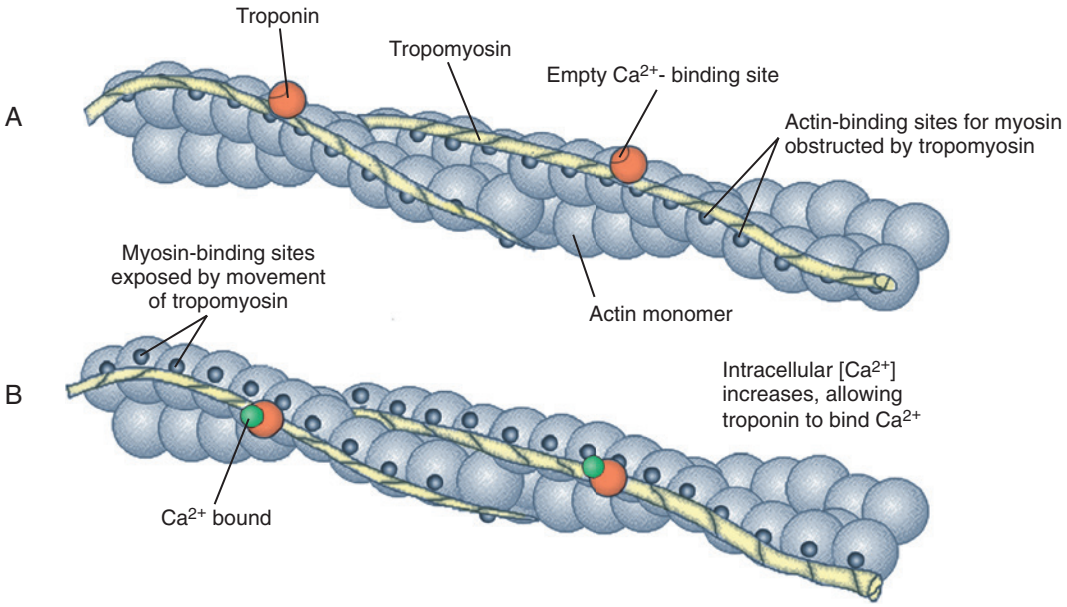
Step B: The additional Ca^{2+} binds to troponin, causing an allosteric change in troponin.

Step C: Because Ca^{2+} is bound to troponin, which in turn is bound to tropomyosin, the Ca^{2+}-induced change in troponin conformation is transmitted to the tropomyosin molecule. When troponin binds Ca^{2+}, tropomyosin changes its binding to actin in such a way that it exposes the actin site for myosin



• **Fig. 1.4** Power stroke of actomyosin. (A) The myosin head has bound to adenosine triphosphate (ATP). In this conformation, myosin has little affinity to bind to actin. (B) ATP is partially hydrolyzed to adenosine diphosphate (ADP) and inorganic phosphate (P_i); the hydrolysis is partial because the products remain bound to the myosin head. The change in what is bound to the myosin (ADP and P_i, not ATP) changes the conformation of myosin so that it binds to actin with high affinity. (C) Hydrolysis is complete; myosin releases ADP and P_i. This change in what is bound at the myosin head causes an allosteric change in the head; it flexes. Because the myosin head is still bound to the thin filament, the flexion causes the thin filament to slide past the thick filament. (D) New ATP molecule binds to the myosin head; as for step A, myosin had little affinity for actin in this state, and the head releases from the thin filament and unflexes.

Vet-ebooks.com



• **Fig. 1.5** Regulation of the actomyosin adenosinetriphosphatase (ATPase) and striated muscle contraction by Ca^{2+}. (A) In the absence of high concentrations of Ca^{2+}, tropomyosin sits in the groove of the actin filament to obstruct the binding sites on actin for myosin. (B) In the presence of higher Ca^{2+} concentrations, the ion binds to troponin, causing an allosteric change in the interaction of troponin with tropomyosin. This allosteric change in turn changes the interaction of tropomyosin with the actin filament to expose the myosin-binding sites on actin.

cross-bridging. (Tropomyosin snuggles down deeper in its actin groove, revealing actin to the myosin head.) As long as troponin binds Ca^{2+}, the muscle contracts by the actomyosin cycle outlined earlier.

Step D: When the Ca^{2+} concentration drops to normal, however, troponin no longer binds Ca^{2+}. This causes tropomyosin to move up in the thin filament groove so that it again blocks the myosin-binding sites on actin. Myosin heads can no longer cross-bridge, and muscle contraction stops.

As with the actomyosin force generation itself, its regulation also shows many examples of the specific binding function. The specific binding of Ca^{2+} to troponin is a purely informational use of protein binding and shape change; that is, troponin has no catalytic, transport, or structural function, but transmits the "on" signal to the next protein. The binding of tropomyosin to actin serves not only a regulatory role but also a structural role; the actin filament is stabilized by tropomyosin, making it less likely to disassemble into actin subunits. The change in the binding geometry of tropomyosin that directly regulates myosin access to actin is a good example of the importance of allosteric change and the following sequence: specific binding (troponin to tropomyosin) → protein (tropomyosin) shape change → change in protein-binding properties (tropomyosin to actin) → a difference in the position of tropomyosin, which in turn regulates the actomyosin motor.

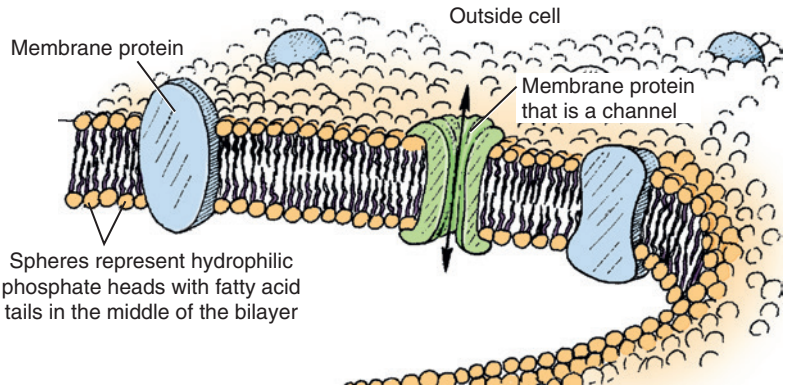
Biological Membranes Are a Mosaic of Proteins Embedded in a Phospholipid Bilayer

Before continuing the discussion of the cellular basis of physiological control, an additional basic structure must be introduced. This is the phospholipid bilayer of the biomembranes of cells. *Phospholipids* are molecules that have two long tails of hydrophobic fatty acid and a head containing a charged, hydrophilic phosphate group. Under appropriate aqueous conditions, these molecules spontaneously form an organized membrane structure, similar to the film of a soap bubble. This filmy layer is composed of two layers (a bilayer) of phospholipid molecules. In both layers, the hydrophilic heads point outward to hydrogen bond with water, and the oily, fatty-acid tails point inward, toward one another and away from the water. Proteins embedded in this lipid bilayer, called *intrinsic membrane proteins* or just *membrane proteins,* produce the fluid mosaic structure of biomembranes shown in Fig. 1.6. All biological membranes share this *fluid mosaic* structure, whether the membrane is the outer plasma membrane separating cytoplasm from extracellular fluid or the membrane surrounding intracellular membranous organelles such as endoplasmic reticulum or lysosomes. It is called a fluid mosaic because of the mosaic of proteins among phospholipids, and because the phospholipid layer is fluid; proteins can move around and diffuse within the plane of the bilayer "like icebergs floating in a phospholipid sea" (the apt phrase of S.J. Singer, one of the originators of the model).

Biological membranes are another crucial molecular structure underlying physiological control. The basic fluid mosaic structure serves four broad functions: (1) compartmentation, (2) selective transport, (3) information processing and transmission, and (4) organizing biochemical reactions in space.

Compartmentation is the ability to separate and segregate different regions by composition and function. For example, the lysosome is a membranous organelle within cells that contains hydrolytic (digestive) enzymes that can potentially digest the cell. The lysosomal membrane compartmentalizes these potentially harmful enzymes, segregating them from the bulk cytoplasm. The rigor mortis, mentioned earlier, that begins shortly after death is transitory because, on death, the lysosomes begin to break open, releasing their enzymes, and the actomyosin cross-bridges are eventually digested apart.

Vet-ebooks.com



• **Fig. 1.6** Fluid mosaic model for biomembranes. Biomembranes consist of a lipid bilayer in which membrane proteins are embedded.

Clearly, the membrane cannot keep a compartment perfectly sealed; material must enter and leave the cell and its internal compartments. *Selective transport* results partly from the properties of the phospholipid bilayer but mostly from transport proteins embedded in the membrane. These proteins are characteristically selective in their transport functions; for example, the protein that is the specialized ion channel underlying neuronal signaling is 15 times more permeable to sodium ions (Na^+) than to potassium ions (K^+). Transport is a major topic of cell physiology and is discussed in more detail later.

If the cells of an organism are to respond to external changes, they must receive information about the state of the outside world. Just as we higher animals have our sensory organs—eyes, ears, nose, and so forth—arrayed on our outside surface, so too do cells have most of their environmental information processing and transmission apparatus on their external surfaces. These are intrinsic membrane proteins of the plasma membrane, called *membrane receptors,* that serve a purely informational function, as discussed earlier.

At first glance it might seem odd that a fluid membrane could provide spatial organization for biochemical reactions. However, returning to the "icebergs in a phospholipid sea" analogy, random collisions are much more likely for material in the two-dimensional membrane surface than for material moving through the three-dimensional volume of the cytoplasm. (If the Titanic had been able to dive or fly, it would have had additional ways to avoid the iceberg!) This much larger collision probability is exploited by the cell in a number of physiological processes. Membranes can also be fenced off into distinct regions, across which there is limited diffusion of membrane proteins. For example, certain cells in the kidney have two membrane regions that are quite distinct with respect to transport proteins, which is important in the regulation of salt and water balance by the animal.

Transport

Only Small, Uncharged Molecules and Oily Molecules Can Penetrate Biomembranes Without the Aid of Proteins

Charged particles (i.e., ions) do not pass through a pure phospholipid bilayer because of the inner, hydrophobic region of bilayer. *Polar molecules* (molecules with no net charge but with electrical imbalances) with a molecular weight greater than about 100 daltons are also unable to pass readily through a pure lipid bilayer, thus excluding all sugar molecules (monosaccharides), amino acids, nucleosides, as well as their polymers (polysaccharide, proteins, nucleic acids). On the other hand, some crucially important polar molecules (e.g., water, urea) are small enough to pass through the lipid bilayer. Small, moderate-sized, and large molecules that are soluble in oily solvents readily pass through a pure lipid bilayer. Physiologically important molecules in this class include O_2, N_2, and the steroid hormones (see Chapters 33 and 34). However, many toxic, synthetic molecules, such as insecticides, are also in this category.

Molecules Move Spontaneously From Regions of High Free Energy to Regions of Lower Free Energy

The majority of biochemicals do not pass readily through a phospholipid bilayer. Transport of this molecular majority requires a protein pathway across the biomembrane. Also needed is a force causing movement along the pathway. Before elaborating on membrane proteins as pathways through the lipid bilayer, the energy factors that drive the transport are considered.

Objects fall spontaneously because of gravity. This is a manifestation of the principle that movement occurs to minimize the potential energy of the object. Indeed, all change in the universe (at scales greater than the subatomic particles) occurs to minimize the potential energy, also called the *free energy,* of the system. The movement of molecules is strongly affected by forces such as concentration, pressure (both part of chemical potential), and voltage (electrical potential). Molecules move spontaneously from a region of higher concentration to lower concentration, from higher to lower pressure, and from higher to lower electrical potential. Each of these factors—concentration, pressure, and electrical potential—is a source of free energy. The transport of a molecule does not depend necessarily on any one factor; rather, the sum of all the free energy contributions is the determinant of transport. The sum of all the free energy contributions on a substance is usually expressed on a per-mole basis as the electrochemical potential. The *electrochemical potential* is the free energy of the substance, from all sources, per mole of the substance.

For spontaneous transport to occur, there must be a difference in the electrochemical potential of the substance between two regions. The two regions are usually two compartments separated by a membrane. This difference in electrochemical potential is called the *driving force.* Typically, students have little difficulty

Vet-ebooks.com

understanding the direction of spontaneous flow as long as only one factor contributes to the electrochemical potential, pressure, or concentration or the voltage. However, understanding physiological transport, both across cells and across tissues, requires an understanding of the contribution of each factor to the driving force. For example, the flow of fluid from the capillaries of the vascular system depends on the balance between both the hydrostatic pressure difference and the concentration difference of solutes (osmotic pressure) across the capillary. Similarly, movement of Na^+ and K^+ ions across the plasma membrane of nerve cells depends on the driving forces contributed by both voltage differences and ion concentration differences across the membrane.

Material moves spontaneously from regions of high electrochemical potential to low electrochemical potential. Such transport is called *diffusion* or *passive transport.* Net movement of material (i.e., diffusion) stops when the electrochemical difference between regions equals zero. The state at which the free energy or the electrochemical potential difference is zero is called *equilibrium.* Equilibrium means "balance," not equality. Equilibrium is reached when the free energy (electrochemical potential) is balanced; the value on one side is the same as the other. In most cases the source of the free energies on the two sides never becomes equal; the concentrations, the pressure, and the voltages remain different, but their differences "balance out" so that the sum of the free energy differences is zero.

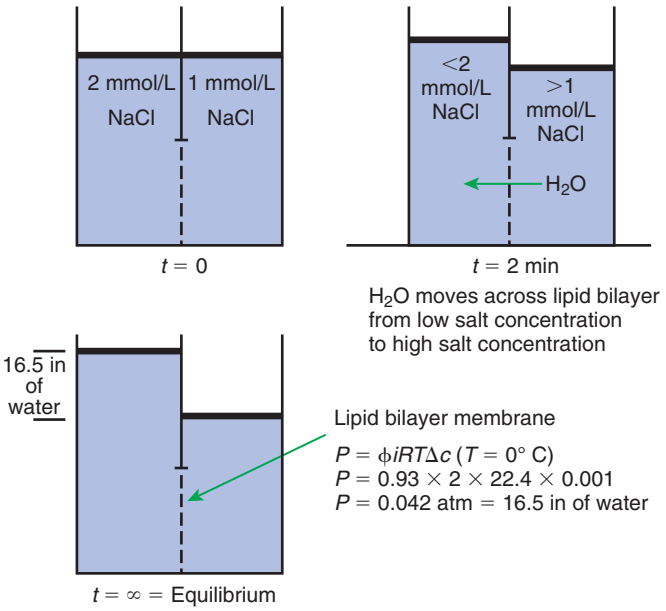
Equilibrium is a particularly important concept because it describes the state toward which change occurs if no work is put into the system. When the system reaches equilibrium, no further net change occurs unless some work is done on the system. The words *net change* are important. Molecules at equilibrium still move and exchange places, but as much goes in one direction as in the other, so there is no net flow of material.

If the cell requires material to move from low to high electrochemical potential (i.e., in the direction away from equilibrium), thus increasing the difference in free energy between two regions, then some driving force, some work, must be provided by some other decrease in free energy. This type of transport is *active transport.* Active transport uses proteins that combine transport and reaction coupling functions; the protein couples the "uphill" movement of material to a "downhill" reaction such as ATP hydrolysis.

Important Transport Equations Summarize the Contributions of the Various Driving Forces

It is worthwhile developing some quantitative aspects of transport, beginning with simple examples and developing equations for the effect of more than one driving force. These equations can be seen as summaries of the physical laws. In most cases the equations describe phenomena with which we have experience by living in a technological society. In these equations, c stands for concentration, V for volume, P for pressure, and so forth; these are common concepts. It is important, however, to think about these equations in real-life terms, not as abstract symbols.

One of these equations relates a hydrostatic (pressure) driving force for water movement that just balances a driving force caused by a chemical potential difference. *Osmosis* is the movement of water across a semipermeable membrane in response to the difference in the electrochemical potential of water on the two sides of the membrane (Fig. 1.7). The chemical potential of water is lower in 1 liter (L) of water (H_2O) in which is dissolved 2 millimoles (mmol) of sodium chloride (NaCl) than in 1 L of H_2O in which is dissolved 1 mmol of NaCl. If these two solutions are separated by a pure lipid bilayer, Na^+ and Cl^- ions cannot move to equilibrate the concentration. Rather, the freely permeable water moves from the side with the higher water potential (low concentration of solute) to the side with the lower water potential (higher concentration of solute). Thus *water follows solute* (a good summary of osmosis), and this water movement dilutes the 2 mmol solution. However, water movement never produces equal concentrations of salt. Rather, another driving force appears as the water moves. The hydrostatic pressure of water increases on the side to which the water moves, increasing the electrochemical potential of the water on that side. Net water movement stops when the increase in water potential from hydrostatic pressure exactly balances the decrease in water potential from the dissolved salt, so that the electrochemical potential becomes equal on both sides of the membrane.



• **Fig. 1.7** Osmosis. At time *(t)* = 0, two compartments are separated by a lipid bilayer membrane (no transport proteins) that contains salt solutions of differing concentrations. At t = 2 minutes, the salt ions cannot move across the membrane to equilibrate their concentration, but water can move. Water moves from the region of higher water potential (low salt) to the region of lower water potential (high salt). Water continues to pass the lipid bilayer until at t = equilibrium; the difference in the height of water between the two sides creates a difference in pressure equal but opposite to the difference in the water potential between the two sides. That is, the free energy difference resulting from differing salt concentrations is equilibrated by an equal but opposite free energy difference caused by pressure.

The initial potential difference of water shown in Fig. 1.7 is caused by the difference in the concentration of material dissolved in the water. A proper explanation of why the water in a solution has a lower chemical potential than pure water (and why water in a concentrated solution has a lower potential than in a dilute solution) is beyond the scope of this chapter. However, readers familiar with the concept of *entropy* will realize that the disorder of a system increases with the introduction of different particles into a pure substance and with the number of different particles introduced. An analogy would be that a canister with mixed sugar and salt is more disordered, and therefore at higher entropy, than a canister with only pure salt or pure sugar. Also, the disorder of the system increases as more sugar is added to salt (up to 50:50); a pinch of sugar in a canister of salt only increases the disorder

Vet-ebooks.com

slightly. Because an increase in entropy causes a decrease in free energy, the free energy of a solution is decreased as the mole fraction of solute increases.

Osmosis is important to cells and tissues because, generally, water can move freely across them, whereas much of the dissolved material cannot. Given a concentration difference of some nonpermeable substances, the *van't Hoff equation* relates how much water pressure is required to bring the system to equilibrium, that is, the free energy contributed by a pressure difference across the membrane that exactly balances an opposing free energy contribution caused by a concentration difference:

$$\Pi = iRT\Delta c$$

Π = Osmotic pressure, the driving force for water movement expressed as an equivalent hydrostatic pressure in atmospheres (1 atm = 15.2 lb/in^2 = 760 mm Hg). Osmotic pressure is symbolized by Π to distinguish it from other types of pressure terms.
i = Number of ions formed by dissociating solutes (e.g., 2 for NaCl, 3 for $CaCl_2$).
R = Gas constant = 0.082 L atm/mol degree.
T = Temperature on the Kelvin scale; 0° C = 273° K. (RT is a measure of the free energy of 1 mol of material because of its temperature. At 0° C, RT = 22.4 L atm/mol.)
Δc = Difference in the molar concentration of the *impermeable* substance across the membrane.

This equation summarizes a balance of driving forces; P amount of hydrostatic (osmotic) pressure is the same driving force as a particular concentration difference, Δc. The osmotic pressure depends only on the concentration difference of the substance; no other property of the substance need be taken into account. (Those phenomena that depend only on concentration, such as osmotic pressure, freezing-point depression, and boiling-point elevation, are called *colligative properties.*) The van't Hoff equation is strictly true only for ideal solutions that are approximated in our less-than-ideal world only by very dilute solutions. Real solutions require a "fudge factor," called the *osmotic coefficient,* symbolized by Φ (phi). The osmotic coefficient can be looked up in a table, and then plugged into the equation as follows:

$$\Pi = \Phi iRT\Delta c$$

The term Φic for a given substance represents the osmotically effective concentration of that substance and is often called the *osmolar* or *osmotic concentration,* measured in osmoles per liter (Osm/L). In general, the osmolar concentration of a substance is approximated by the usual concentration times the number of ions formed by the substance; the osmotic coefficient provides a small correction. The osmolarity of a 100 mmol NaCl solution (0.1 mol) is then 0.93 (Φ for NaCl) × 2(NaCl → $Na^+ + Cl^-$) × 0.1 mol = 0.186 Osm = 186 mOsm.

The previous equation summarizes a phenomenon crucial for physiological function. The greater the concentration difference of an impermeable substance across a membrane, the greater is the tendency for water to move to the side of high concentration. (Water follows solute.) Indeed, if you plug some numbers into this equation, you may be surprised at the large pressures required to balance modest concentration differences. For example, an NaCl concentration difference of 0.1 mol (5.8 g/L) is equilibrated by a pressure (4.2 atm) equal to a column of water 141 feet high (divers must be wary of the bends when ascending from below 70 feet of water). The importance of this is that a small concentration difference can produce a strong force for moving water. The body makes effective use of this to transport water in many tissues: ions/molecules are transported into or out of a compartment → and water follows by osmosis.

Starling's Hypothesis Relates Fluid Flow Across the Capillaries to Hydrostatic Pressure and Osmotic Pressure

An excellent practical example of how a balance of driving forces is responsible for the flow of water and permeable substances across a semipermeable membrane is the movement of water and ions across the single layer of cells (endothelial cells) that compose blood capillaries. The single cell layer composes, in effect, a semipermeable membrane with different transport qualities than that of a simple lipid-bilayer membrane. The junctions between cells are sufficiently permeable to allow small molecules and ions to diffuse between compartments. Only large molecules, most importantly proteins, are unable to move through the holes. The difference in protein concentration between the blood and the water solution surrounding tissue cells, called the *extracellular fluid (ECF)* or *interstitial fluid (ISF),* creates an osmotic pressure for the movement of water with all its dissolved small molecules and ions. This osmotic pressure resulting from dissolved proteins has a special name: c*olloid osmotic pressure* or *oncotic pressure.* Protein is more concentrated in the blood than in the ISF, producing an oncotic pressure of about 0.02 to 0.03 atm = 15 to 25 mm Hg, driving water into the capillary. On the basis of this driving force alone, one would expect the capillaries to fill up with water, thus dehydrating the tissue spaces. However, the heart is a pump that exerts a true hydrostatic pressure on the blood, tending to drive the water (and other permeable molecules) out of the capillaries. The net driving force is the algebraic sum of the oncotic pressure difference and hydrostatic pressure difference between the capillaries and the ISF, as follows:

$$\text{Net driving force in capillary} = (P_c - P_i) - (\pi_c - \pi_i)$$

P_c = Hydrostatic pressure in the capillary.
P_i = Hydrostatic pressure in the interstitial space (usually near 0).
π_c = Oncotic pressure of blood plasma in capillary (~28 mm Hg).
π_i = Oncotic pressure of ISF (~5 mm Hg, but depends on the particular tissue).

This equation has enormous relevance to the function of the circulatory system. On the arterial end of capillaries, the hydrostatic pressure (P_c) is high, about 35 mm Hg. Plugging this number into the equation along with the others, the net pressure in the capillary is +12 mm Hg; fluid is being driven out of the capillary on the arterial side *(capillary filtration).* The flow of fluid through the resistance of the capillary causes a decline in pressure so that the hydrostatic pressure on the venous side is low, P_c = 15 mm Hg. The oncotic pressures have not changed, so the net driving force on the venous side is −8 mm Hg; there is a net absorption of fluid into the capillary on the venous side *(capillary reabsorption).* This arrangement achieves a major function of the circulatory system; in this way the fluid of the blood circulates among the cells and is then recycled back into the circulatory system.

Pathological alterations in this system emphasize the physiological importance of balance of driving forces for transport. Chronic liver disease occurs with some frequency in horses and dogs, among other mammals. The liver is compromised in its ability to synthesize and secrete a major blood protein, serum albumin. The decline in

Vet-ebooks.com

TABLE 1.1 Concentrations of Various Substances in the Intracellular, Extracellular, and Plasma Fluids

	CONCENTRATION (MMOL/L)		
	Intracellular	Extracellular	Blood Plasma
Na^+	15	140	142
K^+	150	5	4
Ca^{2+}	0.0001	1	2.5
Mg^{2+}	12	1.5	1.5
Cl^-	10	110	103
HCO_3^-	10	30	27
Phosphate	40	2	1
Glucose	1	5.6	5.6
Protein	4.0	0.2	2.5

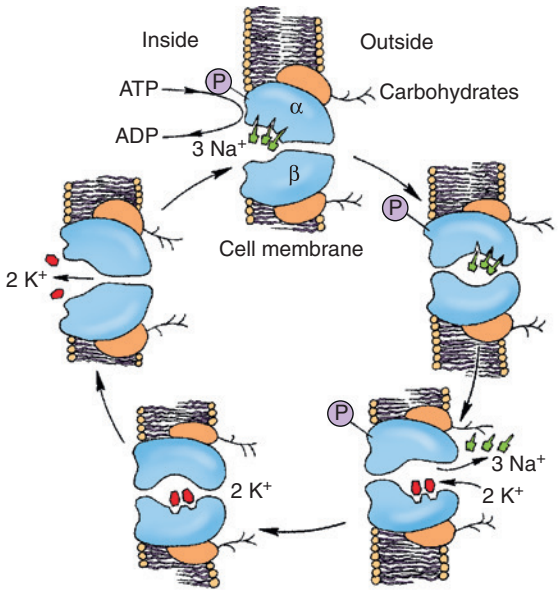
the concentration of serum albumin lowers the oncotic pressure of the blood. As a result, there is more force to drive fluid out of the capillaries on the arterial side and less driving force for net absorption of fluid on the venous side of capillaries. This causes the tissue spaces of the diseased animals to fill with fluid, a painful and visually obvious symptom called *edema.* The Clinical Correlations section at the end of the chapter provides another example of edema in which increased hydrostatic pressure in the veins and capillaries causes increased capillary filtration and less capillary reabsorption.

Membrane Proteins That Serve the Triple Functions of Selective Transport, Catalysis, and Coupling Can Pump Ions and Molecules to Regions of Higher Free Energy

The van't Hoff equation and Starling's hypothesis deal with passive transport (i.e., movement of material in the direction of lower electrochemical potential). However, the cell moves many ions/molecules against their electrochemical potential. That is, this selective transport requires the expenditure of energy by the cell. Transport in a direction requiring an expenditure of energy (i.e., input of work) is called *active transport.* Active transport depends on intrinsic membrane proteins that use specific binding and allostery to achieve the dual functions of selective transport and reaction coupling. Many (but not all) active transport proteins obtain the energy for transport from ATP hydrolysis; these proteins must function also as enzymes (ATPases).

An important example of active transport is the Na^+, K^+ pump (also known as Na^+, K^+-ATPase). This intrinsic membrane protein consists of four polypeptide chains (2 α + 2 β) and has a mass of approximately 300,000 daltons. This molecule catalyzes the hydrolysis of ATP and couples the hydrolysis energy to the movement of Na^+ out of the cell and K^+ into the cell. This ion pump creates and maintains a considerable concentration gradient across the cell membrane for both ions (Table 1.1).

Fig. 1.8 shows our current understanding of this protein's structure and outlines the cycle of binding and conformational changes underlying its transport function. The Na^+, K^+-ATPase

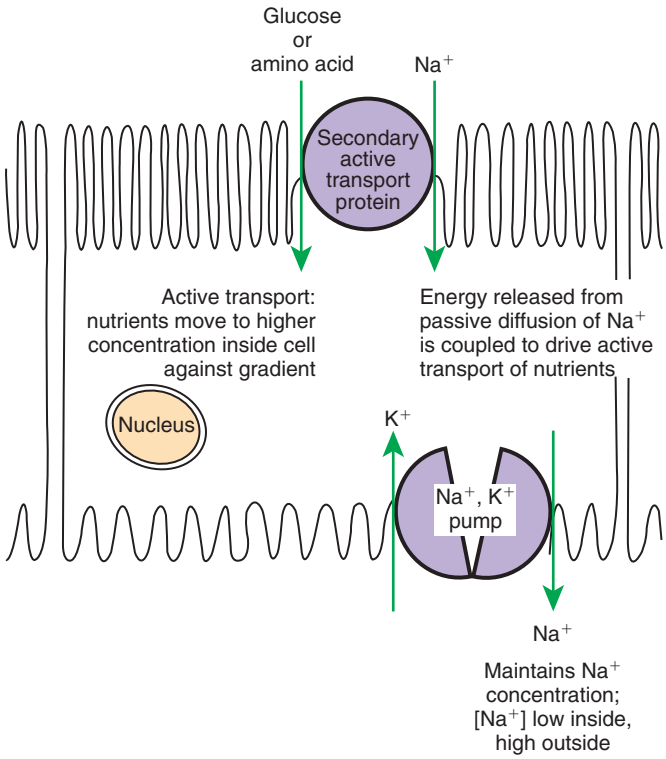


• **Fig. 1.8** Hypothetical transport cycle for Na^+, K^+-ATPase. Changes in the conformation of this transport protein driven by ATP hydrolysis and ion-binding events cause three Na^+ ions to be moved out of the cell against a concentration gradient and two K^+ ions to be moved into the cell, also against a concentration gradient, for each ATP hydrolyzed. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate. (Redrawn from a diagram by Dr. Seth Hootman.)

pumps three Na^+ ions *out* of the cell and two K^+ ions *into* the cell for each ATP molecule hydrolyzed. These directions of ion pumping cause a high Na^+ concentration outside the cell and a low concentration inside, whereas K^+ concentration is high inside and low outside the cell. The different directions of pumping for the two ions depend on differing binding specificity of the pump protein in the two conformational states. The ability of the protein to couple this transport to the enzymatic breakdown of ATP allows the transport to occur against the concentration gradients, from lower to higher electrochemical potentials for both ions. In the particular case of the Na^+, K^+ pump, the number of transported electrical charges is asymmetrical; three positive charges leave for each two positive charges that enter. This asymmetry of electrical charge transport means that the Na^+, K^+ pump is *electrogenic,* making a minor contribution to the electrical potential (voltage) across cell membranes, as discussed later.

Many different intrinsic membrane proteins actively transport a wide variety of ions and molecules against the transported molecules' electrochemical gradient. Many, such as the Na^+, K^+ pump, couple the energy-requiring "uphill" transport with the "downhill" hydrolysis of ATP. However, any potential source of free energy can be coupled to the energy-requiring transport. Indeed, the gradient of Na^+ set up by the Na^+, K^+ pump is itself used frequently as a source of energy. That is, the "downhill" flow of Na^+ from outside the cell to the inside is a spontaneous reaction whose energy can be coupled to some "uphill" reaction (Fig. 1.9). For example, the transport of glucose and many amino acids from the food mass in the small intestine into the cells lining the gut is an active transport process and requires an Na^+ concentration gradient. Transport proteins in the plasma membrane of intestinal epithelial cells couple the spontaneous diffusion of Na^+ into the cell to the inward, energy-requiring transport of the sugar or amino acids. These nutrients are at higher concentration inside the cell

Vet-ebooks.com



• **Fig. 1.9** Secondary active transport as exemplified by uptake of nutrients by gut epithelia. Nutrients such as glucose and amino acids must be actively transported from relatively low concentration in the gut lumen toward higher concentrations within the cells lining the gut. This active transport process uses the concentration gradient of Na^+ ions set up by Na^+, K^+-ATPase (see Fig. 1.8) as the source of energy for the active transport process. In other words, the energy released by the passive diffusion of Na^+ into the cell along its concentration gradient is coupled to the energy-requiring transport of glucose or amino acids against their concentration gradients. Thus the secondary active transport protein both serves a transport function and couples the "downhill" transport of Na^+ to the "uphill" transport of nutrients. There are many such secondary active transport processes in the body. For example, the same mechanism shown here is used to reabsorb nutrients from blood filtrate in the kidney. *ATPase,* Adenosinetriphosphatase.

than outside, so they must be actively transported into the cell at the expense of the energy stored in the Na^+ electrochemical gradient. That is, the energy from the "downhill" diffusion of Na^+ into the cell is coupled to the "uphill" transport of the nutrient into the cell. Such active transport coupled to Na^+ diffusion across the cell membrane is called *secondary active transport* because of its dependence on the Na^+ concentration gradient established by the primary active transport of the Na^+, K^+ pump.

Examples of transport can be identified in a number of ways. Our examples have been instances in which two ions/molecules must be transported together or not at all, and such transport is called *cotransport.* Cotransport can involve one process of passive transport (diffusion) with an active transport process, as in the two previous examples; it can involve two active transport processes (e.g., Na^+, K^+-ATPase) or two diffusion processes. In the first case, the need for cotransport is energetic; the flow of one ion is needed to drive the other. In the two latter cases, the need for cotransport is a restriction based on the binding properties of the transport protein; it cannot bind one without the other. Cotransport proteins that transport both substances in the same direction are called *symports* or *symporters.* The Na^+/sugar cotransporter in the gut is a symport. Cotransport proteins that transport the two substances in opposite directions (e.g., Na^+, K^+-ATPase) are called *antiports* or *antiporters.* Parenthetically, proteins that transport only a single ion or molecule are called *uniports* or *uniporters.*

Many Membrane Proteins Selectively Facilitate the Transport of Ions/Molecules From High to Low Electrochemical Potential

The movement of ions and of medium and large polar molecules requires a protein molecule to serve as a pathway through the obstruction of the phospholipid bilayer. If the movement of the substance is in the natural direction of its electrochemical gradient (movement from high to low), the transport process is called *facilitated diffusion.* The membrane proteins mediating this transport process through the phospholipid bilayer are *channels* or *carriers* (Fig. 1.10). These are distinguished by the extent to which the protein interacts with the transported substance.

Carriers bind the transported substance in the lock-and-key manner, so there is a site-specific binding of the transported substance to the transport protein (see Fig. 1.10A). Carrier-mediated transport is typically much slower than channel-mediated diffusion because of the relatively slow binding and unbinding processes. The Na^+, K^+ pump and the Na^+/glucose symport are both examples of carriers.

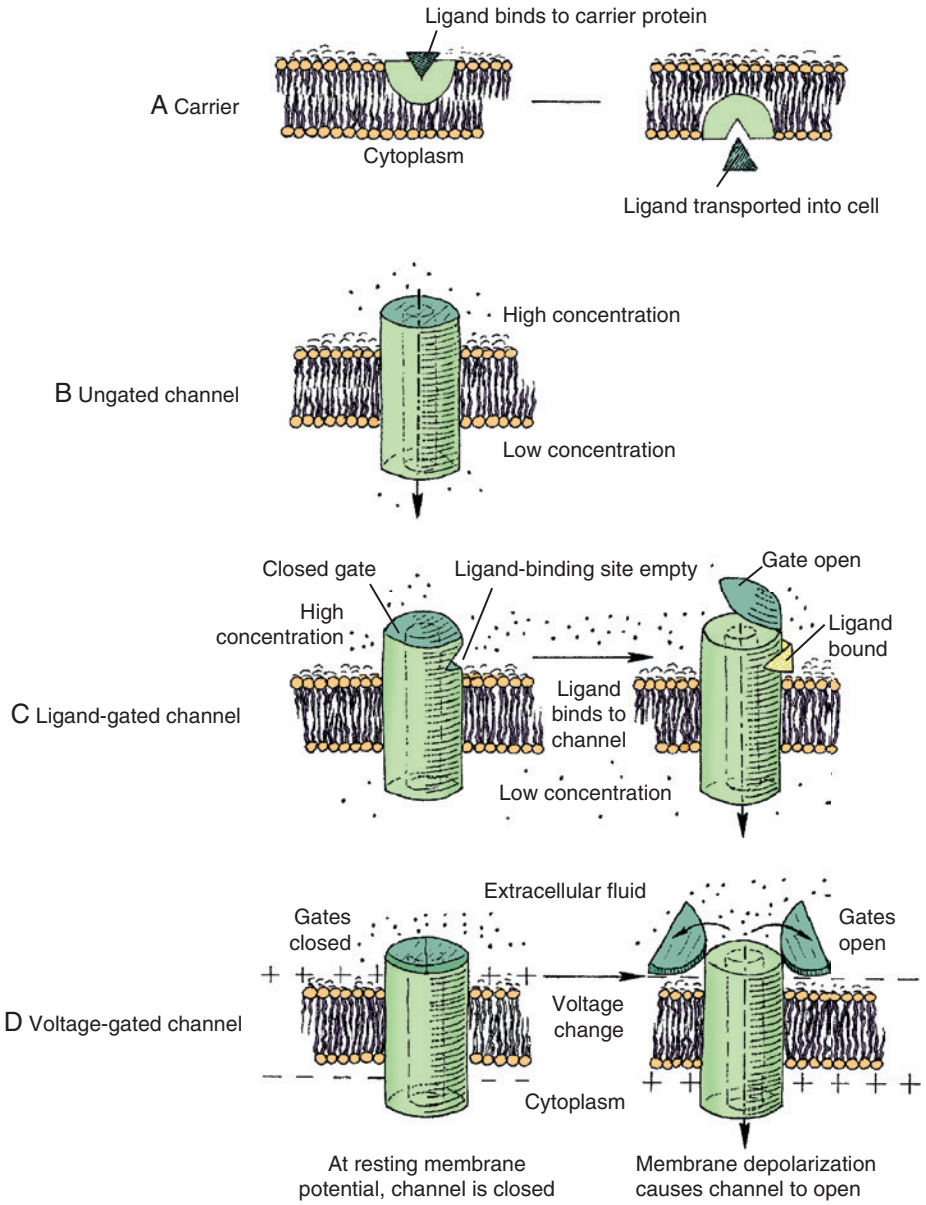
Channels can be thought of as "protein donuts" embedded in the phospholipid bilayer. The hole in the donut is a pore in the membrane through which small ions such as Na^+, K^+, Ca^+, Cl^-, and H^+ are transported. Although most channels transport ions, a class of channels called *aquaporins* comprises channels for water flow. (Although water can flow through a pure lipid bilayer, this transport is too slow for some functions. Kidney cells, for example, are particularly rich in aquaporins, which are required for the water balance function of the kidney.) For all channels, the pore size and the interaction of the transported material with the amino acid side groups lining the pore allow membrane channels to be selective. Only specific molecules or ions can move through a particular channel. Movement of material through channels is almost as rapid as simple diffusion through a water-filled space of the same area as the channel pore.

The plasma membranes of most cells have passive leaks of ions, particularly K^+. These ionic leaks are typically ascribed to *leak channels,* which are open at all times (see Fig. 1.10B). However, most ion channels open or close in response to signals. These latter types are called *gated channels.* The opening and closing of the gates are examples of the allosteric property of proteins. The same signals responsible for allosteric changes in general—ligand binding, phosphorylation, and voltage differences—also control the opening and closing of gated channels, as shown in Fig. 1.10. (Because mechanically gated channels are so poorly understood, these are not discussed here.)

Channels that open in response to ligand binding are called *ligand-gated channels* (see Fig. 1.10C). The nicotinic acetylcholine receptor is a ligand-gated channel found in skeletal muscle membrane directly beneath incoming neurons (nerve cells). This channel is found also in the membrane of neurons in autonomic ganglia and in the brain. As the name implies, the nicotinic acetylcholine receptor binds to the drug nicotine and the neurotransmitter acetylcholine. In both cases, the channel opens in response to ligand binding.

This nicotinic acetylcholine channel plays a key role in transmitting electrical stimulation from neurons to skeletal muscle cells. Briefly, motor neurons release the neurotransmitter acetylcholine

Vet-ebooks.com



• **Fig. 1.10** Types of transport proteins mediating facilitated diffusion. In all cases the ion moves from a region of high potential (shown here as high concentration) to a region of low potential. (A) Carriers. In a few cases, material is carried by a transport protein that binds tightly to the material, and the complex moves through the lipid bilayer. (B) Leak channels. These channels are thought not to open and close as do gated channels, and thus they support a small but persistent leak of a particular ion through the pore. Although their existence was long postulated, distinct, ungated leak channels have only recently been identified and isolated, as opposed to leaks through normally gated channels. Selectivity of these and other channels is based on the size of the pore and the weak interactions of ions with the atoms lining the pore. (C) Ligand-gated channels. The transport protein again forms a pore through the membrane. In the case of gated channels, access of the ion to the pore is controlled by a gate, a substructure of the transport protein that can open and close the pore. In ligand-gated channels the opening and closing of the gate are controlled by the binding of a ligand to the channel. (D) Voltage-gated channels are similar to ligand-gated channels, except the opening and closing of the gate are controlled by the electrical field around the channel.

in response to the electrical signal coming down the neuron. This acetylcholine binds to and opens the ligand-gated channel on the skeletal muscle. The influx of Na^+ into the muscle cell initiates an electrical response in the muscle, causing the release of Ca^{2+} (through gated channels in the endoplasmic reticulum), in turn causing contraction. (This brief account of neuromuscular transmission, presented only to provide orientation to the function of the acetylcholine channel, is expanded in Chapters 5 and 6). In the case of the nicotinic acetylcholine receptor/channel, the specific binding and allosteric properties of the protein serve the dual functions of selective transport across the membrane and information reception and transmission to the muscle cell.

Vet-ebooks.com

Channels that open in response to voltage changes across the membrane are called *voltage-gated* or *voltage-dependent channels* (see Fig. 1.10D). This type of channel is largely responsible for the neurons' ability to transmit information along their length and to release neurotransmitter. All voltage-gated channels have a range of membrane potentials that cause them to open; this is the channel's activation range. The minimum membrane potential that causes opening is the channel's *threshold.* The activation range and threshold vary from channel to channel, depending on the conformation of the protein and the electrical properties of the amino acid side groups that form the gate of the channel. In addition to an open and closed configuration, many voltage-dependent channels have a third conformation, called *inactivated.* Like the closed configuration, the inactivated conformation prohibits the diffusion of ions through the channel. Unlike the closed configuration, it does not open immediately in response to changes in membrane potential. Inactivation can be regarded as an enforced rest period for the channel. Voltage-dependent channels that do not inactivate have only open and closed conformations, and they take up one or the other conformation, depending on the membrane potential.

As previously discussed, any of the functions of proteins can be used to transmit information if a difference in the protein function changed the cell. Gated channels, both ligand and voltage gated, are ideal candidates for information transmission because they change their function: opening and closing, permitting or stopping transport. Indeed, the sole physiological function of the nicotinic acetylcholine receptor/channel, as described earlier, is the transmission of information: turning the chemical stimulation by the neuron of the muscle into electrical stimulation (see following discussion) of the muscle membrane, leading to muscle contraction.

Passive Transport of K^+ Across the Plasma Membrane Creates an Electrical Potential

As just discussed, gated ion channels can convert chemical information into electrical information. Electrical signaling in the animal body is the result of electrical imbalances maintained across the plasma membrane of virtually all cells: cells maintain an electrical potential difference across their plasma membrane. That is, the cell membrane is a battery; if one attaches electrodes to the two ends of a battery or to the inside and outside of a cell, one finds a voltage difference between the two ends or sides. If one provides a path for electrical charges to move—a metal wire containing free electrons in the case of a battery or a membrane channel through which ions can move in the case of the cell—an electrical current flows from higher to lower electrical potential. The diversity of battery-powered devices in our society suggests how many ways this electrical potential can be exploited. The physiology of animals also exploits the baseline electrical potential across the plasma membrane, called the *resting membrane potential.* The word "resting" is added to distinguish the baseline potential from the instantaneous values of membrane potential during the passage of membrane currents.

The resting membrane potential is the indirect result of the concentration gradients of ions across the plasma membrane caused by the activity of the Na^+, K^+-ATPase. Partly, this membrane potential is a result of the asymmetry in numbers of ions pumped by the Na^+, K^+-ATPase. However, most of the membrane potential is caused by the passive flow of K^+ through *K^+ leak channels* in response to the concentration gradient of K^+ (high inside, low outside). This concentration gradient sets up an electrical driving force (voltage) that exactly balances the concentration driving force. The concentration of K^+ inside a mammalian cell is about 150 mmol; outside in the ISF, it is about 5 mmol. As a result, K^+ tends to diffuse from the cytoplasm through the leak channel to the ISF. However, when K^+ alone leaves the cytoplasm without an accompanying negative ion, it causes an electrical imbalance. The exit of K^+ ions leaves the inside of the cell with negative charges not neutralized by positive potassium ions, and the ISF now has positive K^+ ions not balanced by negative charges. The cell is building an electrical potential difference across the plasma membrane with the cytoplasm being negative relative to the ISF.

This electrical potential driving force increases until it balances the concentration driving force for K^+. This situation is analogous to osmosis; the concentration-driven flow of water across a semipermeable membrane creates a different driving force, pressure, that eventually balances the concentration-driving force. Similarly, for the resting membrane potential, the concentration-driven flow of K^+ across the semipermeable membrane (semipermeable in the sense that negative ions do not accompany the K^+) creates a different driving force, an electrical voltage, that eventually balances the concentration force. As in the case of osmosis, an equation is used to relate the size of the concentration gradient to the size of the electrical potential that provides an exact balance. This equation is called the *Nernst equation,* as follows:

$$E_X = RT/zF \ln[X_{outside}]/[X_{inside}]$$

E_X = Equilibrium potential for ion X
RT = Gas constant × Absolute temperature
z = Electrical valence for the ion, +1 for Na^+ and K^+, −1 for Cl^-, and so forth
F = Faraday constant = number of coulombs of electrical charge in a mole of ions = 96,500 coulombs/mol
ln = Natural logarithm (i.e., log to base e)
[X] = Concentration of ion X

A simpler form of this equation can be written by taking advantage of the fact that R and F are constants, T is almost constant under physiological conditions, and the natural log of a number is 2.3 times the common log ($\log_{10}$), as follows (mV, millivolts):

$$E_X = -60 \text{ mV}/z \log[X_{inside}]/[X_{outside}]$$

Because the state of balance between the electrical driving force and the concentration driving force is equilibrium the value of the electrical potential is called the *equilibrium potential* of the ion. Given the previous concentrations for K^+ inside (150 mmol) and outside (5 mmol) the cell, the equilibrium potential for K^+ is

$$\begin{aligned} E_{K+} &= -60 \text{ mV}/+1 \times \log 150/5 \\ &= -60 \text{ mV} \log 30 \\ &= -60 \text{ mV} \times 1.47 \\ &= -88.2 \text{ mV} \end{aligned}$$

Indeed, the measured resting membrane potential across a human muscle cell is −90 mV.

Several aspects of this important equation are worth discussing. If the equilibrium potential for a particular ion is the same as the measured membrane potential, the net driving force for the ion is zero. In this case there is no net movement, even in the presence

Vet-ebooks.com

of wide-open channels, to provide a path through the membrane. However, for any gradient of a specific ion, if the measured membrane potential is *not* the equilibrium potential of that ion, there is a driving force for the transport of that ion. That is, when the membrane potential is anything other than the equilibrium potential, that ion will flow across the membrane if an appropriate channel is open. Thus the equilibrium potential for an ion provides a "baseline" for comparison with the actual membrane potential to determine whether an ion will tend to move across the plasma membrane. If the measured membrane potential has the same sign but is larger in magnitude than the equilibrium potential, the ion flows in the direction of the electrical potential. If the sign is the same but the magnitude lower, the concentration-driving force determines the direction of flow of the ion. If the measured potential has the opposite sign of the equilibrium potential, both electrical and concentration forces are acting on the ion in the same direction. Flow of ions across the plasma membrane (i.e., electrical current) in response to the balance of force between concentration and voltage produces the electrical changes in neurons that underlie the nervous system, as discussed in Chapter 4.

It would be reasonable but incorrect to assume that the transport of ions required to set up the electrical potential measurably alters the concentration gradient. This is untrue because of the large amount of energy required to separate electrical charges. The separation of charge arising from the transport of a few ions balances the energy of quite substantial concentration gradients. Indeed, so few ions move that they cannot be measured by chemical means. Thus electrical, not chemical, measurements are used routinely to assess transport of ions in cells. The measurable voltage changes caused by immeasurably small concentration changes of ions means also that the electrical phenomena at the membrane persist for many hours, even if the Na^+, K^+-ATPase is inactivated by a toxin. That is, an existing concentration gradient of K^+ would require hours to dissipate at the rate of K^+ leakage characteristic of the plasma membrane. Using the membrane battery analogy, the Na^+, K^+-ATPase is a battery recharger. A smartphone does not require the minute-to-minute services of a battery recharger. Enough energy is stored in the battery to operate the phone for an appreciable period, although the battery recharger is needed eventually. Similarly, enough energy is stored in the K^+ concentration gradient to maintain the membrane potential for a period of time. The Na^+, K^+-ATPase is not required on a minute-to-minute basis, although it is needed ultimately to maintain the concentration gradient on which the resting membrane potential depends.

Spatial Organization of Active and Passive Transport Proteins Enables Material to Pass Completely Through the Cell

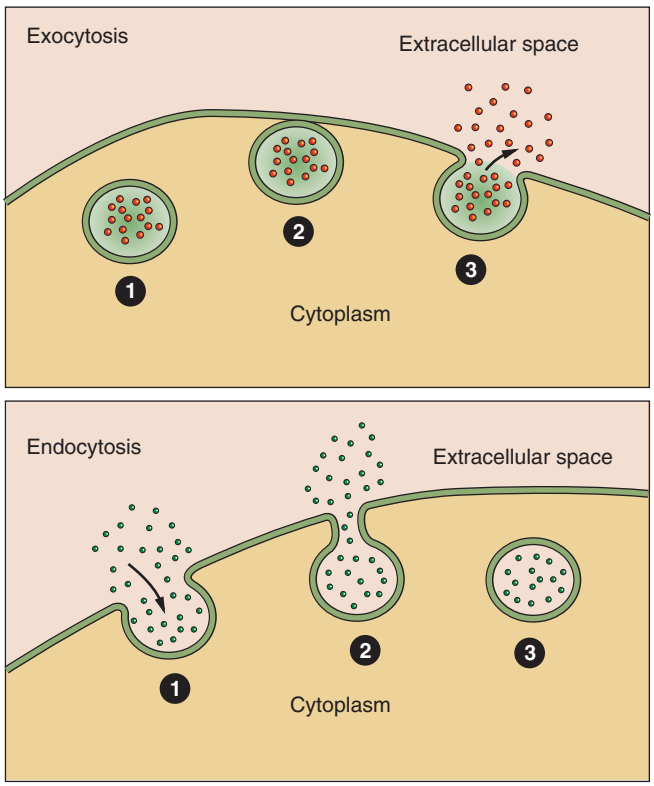
Although macromolecules and biomembranes clearly underlie physiological function, many phenomena of the intact animal emerge that are not initially apparent as a simple sum of parts. One interesting example is the spatial organization of plasma membrane transport proteins so that ions move across the cell from one ECF compartment to another. This is called *transcellular transport* or, because it typically occurs across a layer of epithelial cells, *epithelial transport.* This epithelial transport is important in the kidney (see Chapter 42). The plasma membrane of the epithelial cells in the proximal tubules of the kidney contains two distinct regions. The apical membrane regions face the lumen of the tubule and the fluid that will become urine, and the basolateral regions are near the capillaries and the blood. The apical surface contains ungated leak channels for Na^+, whereas the basolateral surface contains Na^+, K^+-ATPase molecules. (The membrane proteins in one region are prevented from diffusing into the other by membrane protein "fences" called *tight junctions.*) Na^+ diffuses into the cell on the apical surface from the urinelike fluid driven by both the concentration gradient and the resting membrane potential. When inside the cell, the Na^+ is pumped from the basolateral surface, essentially into the blood, by the Na^+, K^+-ATPase. This allows the kidney to reabsorb and thus conserve Na^+. As long as the Na^+, K^+-ATPase remains restricted to the basolateral surface and the passive channel to the apical membrane, Na^+ can move through the cell from the urinelike fluid in the tubule to the blood in the capillaries. If either protein should lose its spatial restriction, Na^+ would be transported into and out of the cell on the same surface, merely consuming ATP, with no net transport of Na^+ from lumen to capillary.

Membrane Fusion Allows for a Combination of Compartmentalization and Transport of Material

Impermeable molecules can be transported across the cell membrane by means other than membrane proteins. This method involves using membrane itself as a carrier compartment. The lipid bilayer of biological membranes has a structure similar to soap bubbles. As with soap bubbles, small vesicles of biomembrane (essentially "membrane bubbles") can fuse to form larger membrane surfaces. A large membrane surface can also pinch off (requiring fusion of two membrane surfaces) into small vesicles. When these processes occur at the plasma membrane, they are called *exocytosis* and *endocytosis,* respectively (Fig. 1.11). More generally, pinching off of membrane or fusion of two membrane vesicles (e.g., for internal membranes) is referred to as *membrane fusion,* whatever the direction. Membrane fusion underlies much of membrane vesicle traffic around the cell. This traffic creates intracellular vesicles, renews plasma membrane by adding newly synthesized membrane, and transports material within the cell and across the plasma membrane. Because the transport is compartmentalized within a membrane bubble, the transported material can be targeted specifically to one or another region of the cell. Also, changes to the "cargo" can occur within a particular membrane compartment, as seen in cholesterol transport.

Exocytosis and endocytosis are crucial in the transport of cholesterol (Fig. 1.12). *Cholesterol* is an essential lipid component of many animal biomembranes; the plasma membrane lipids of animals are about 15% cholesterol and 60% phospholipids. Cholesterol is also the starting material for the synthesis of the entire group of hormones called *steroids* (see Chapters 33 and 34). Cholesterol can be synthesized by animals and is also absorbed by meat-eating animals from their diet. Because cholesterol is soluble in oil, it passes from food through the plasma membrane without protein mediation into the cells of the gut lining. However, transport of dietary cholesterol through the circulatory system requires that cholesterol molecules form a complex with a protein molecule to create low-density lipoproteins (LDLs). To take up cholesterol from the circulation, cells bind the LDLs to intrinsic membrane proteins that act as LDL receptors, as shown in Fig. 1.12. The receptor/LDL complex then diffuses in the plane of the membrane into specific regions to form coated pits. The coated pit is taken into the cytoplasm by endocytosis. In addition to the transport function, *receptor-mediated endocytosis* functions also to concentrate extracellular material before internalization. The coated pit is not

Vet-ebooks.com



• **Fig. 1.11** Two membrane fusion processes: exocytosis and endocytosis. (*Top*) In exocytosis, a membrane-bound vesicle from the cytoplasm (*1*) makes contact and fuses with the plasma membrane (*2*). As the vesicle membrane becomes continuous with the plasma membrane, the contents of the vesicle are released to the extracellular space (*3*). (*Bottom*) In endocytosis, some material from the extracellular space is surrounded by plasma membrane (*1*), which continues to invaginate until the edges are able to fuse (*2*), thus pinching off a vesicle from the plasma membrane (*3*). Membrane fusion can occur between any two compartments within cells separated by lipid bilayer membrane, not only between the cytoplasm and extracellular space, as shown here.

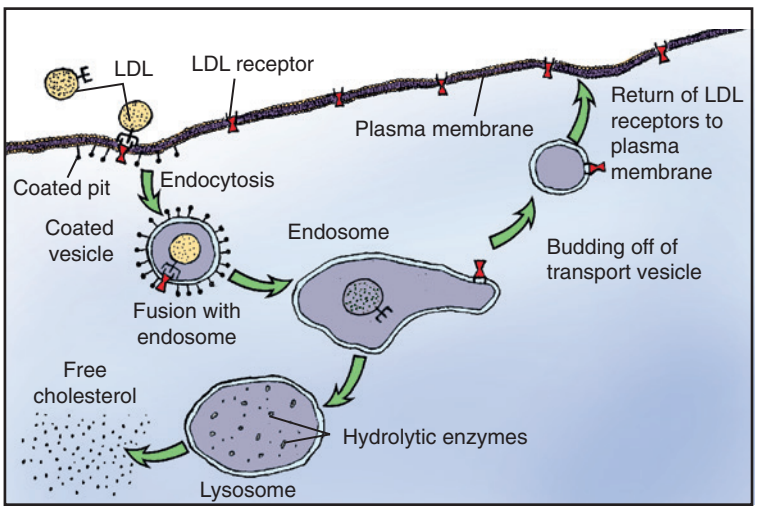
taken into the cell until it has collected the LDLs from a much larger volume of ECF than the cell could "drink." The membrane vesicles formed by this endocytosis fuse subsequently to become an *endosome*. The endosome compartment becomes acidic, which causes dissociation of the LDL and the receptor. Through unknown means, the endosome is then able to further separate and compartmentalize the receptor from the LDL. Membrane vesicles containing the now-vacated LDL receptors return to the plasma membrane and fuse by exocytosis. The LDL receptor is recycled to the plasma membrane to pick up more LDL. Experimental evidence suggests that a single LDL receptor molecule can cycle between the plasma membrane and endosomal vesicles more than 100 times before losing its activity. Meanwhile, the LDL moiety is segregated to another endosomal vesicle, which fuses with the lysosome. The lysosome contains hydrolytic enzymes, thus allowing the internalized LDL to be digested. The cholesterol is now available to the cell for steroid synthesis or incorporation into membrane.

Other molecules are also recycled by endocytosis. As with the LDL receptor, for example, many signal receptors, discussed in the next section, are endocytosed back into the cell that released them, saving the cell the effort of synthesizing new receptors. Not all endocytosed molecules are recycled. Many are broken down after their endosome fuses with a lysosome. Indeed, as described later, this is one method of regulating receptor number on the plasma membrane.

Information Transmission and Transduction

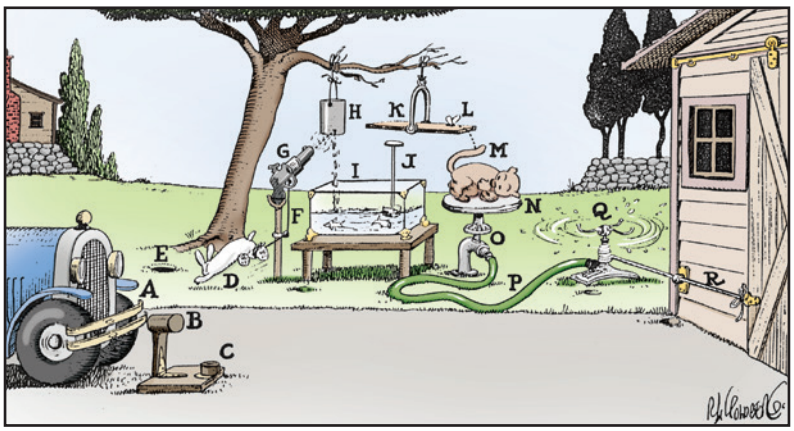
Cell Signaling Usually Occurs by a Lengthy Chain of Sequential Molecular Interactions

One of the areas of most rapid progress in cellular physiology has been our understanding of the mechanisms by which extracellular signals, such as hormones, growth factors, and neurotransmitters, alter cellular function, which in turn alters tissue, organ, and animal function. At the molecular level, almost all chemical signaling



• **Fig. 1.12** Processes of membrane fusion involved in cholesterol uptake by cells. Starting at the left, a low-density lipoprotein (LDL)–containing cholesterol binds to an LDL receptor protein of the plasma membrane and undergoes endocytosis, forming an endosome. The receptor is detached from its LDL ligand in the endosome. The LDL portion of the endosome fuses with a lysosome to digest the LDL and produce free cholesterol, while the receptor-containing portion of the endosome pinches off a vesicle to return to the plasma membrane, thus recycling the receptor. (Redrawn from Alberts B, Bray D, Lewis J, et al. *Molecular Biology of the Cell*, New York: Garland; 1983.)

Vet-ebooks.com



• **Fig. 1.13** Rube Goldberg device (garage door opener, circa 1928) as an analogy for the complex cause-and-effect sequence characteristic of cellular chemical signaling. Automobile (*A*) drives into driveway, causing hammer (*B*) to ignite toy cap (*C*), frightening rabbit (*D*) into its burrow (*E*), and causing pistol (*G*) to fire and so forth, ultimately leading to the opening of the garage door (*R*). As explained in the text, this whimsical "machine" serves as an analogy for chemical signaling within cells because of the multiple control elements, their connection as a cause-and-effect sequence, and the use of household items, similar to the use of evolutionarily conserved proteins of cells in signaling.

shares a common "strategy" of mechanism; signals are sent as a long chain of chemical cause-and-effect interactions transmitted between many sequential chemical steps. Indeed, chemical signaling pathways are structured similar to the whimsical "machines" in the cartoons of Rube Goldberg (1883–1970), a famous American newspaper cartoonist. Fig. 1.13 shows one of his cartoons from 1928 illustrating an outlandish contraption (a *Rube Goldberg device*) to serve as an automatic garage door opener, realistic versions of which had not yet been invented. The automobile (A) drives in, causing a hammer (B) to ignite a toy cap gun (C), which frightens a rabbit (D) with a string (F) tied to its leg, thus firing a pistol (G), and so forth, until a connection to a rotating water sprinkler causes the carriage-house door to slide open (overhead doors also had not yet been invented). Although much of the humor of this parody of a machine is lost on us (our attitudes about machines have changed greatly since Goldberg's heyday), Rube Goldberg devices are a surprisingly useful analogy to the overall mechanism of cellular chemical signaling.

Just as the garage door opener of Fig. 1.13 depends on a series of sequential cause-and-effect interactions, so chemical signaling occurs by a series of cause-and-effect changes in protein shape and binding. Just as the complex events of Goldberg's device are linked to signal and to actuate a garage door opening, so a cascade of changes in protein shape and function are linked to signal and actuate physiological events. Our earlier example of muscle contraction illustrates such a pathway of cause and effect and the analogy to Rube Goldberg devices. Electrical excitation (A) of a muscle cell increases intracellular Ca^{2+} concentration (B), causing Ca^{2+} to bind to troponin (C). This in turn alters the binding of tropomyosin (D) to actin (E) allowing the myosin heads (F) to bind to actin, thus leading to cross-bridging (G) and hydrolysis of ATP and contraction.

As this example indicates, the sequence of cause and effect for both chemical signaling and Rube Goldberg devices is complex. Both involve many different elements, none of which can be identified as *the* controller; all the elements are involved in control. Importantly, this creates multiple sites for regulation and for therapeutic drug action. Just as increasing the caliber of the pistol in the garage door opener would change the response time for opening, so a drug that bound to an element in the middle of a signaling pathway in a cell could increase or decrease the final physiological change in response to a particular hormone, for example. Also related to complexity is that the chain of cause and effect is not obvious; the particular sequence connecting a particular signal (adrenaline binding to a receptor on heart muscle) to a particular outcome (increased cardiac output) must be memorized. However, when the sequence is understood, you can predict from the state of one element in the chain what should happen next. Finally, Rube Goldberg devices were cobbled together from reasonably common household items, such as the bucket, fish tank, sprinkler, and even pistols. Similarly, the elements of chemical-signaling pathways are often highly conserved, and you will see throughout your studies that the same molecules or same basic types of molecules are used in a wide variety of different stimulus-response pathways.

Signaling Pathways Begin With the Binding of an Extracellular Molecule to a Receptor

In addition to the Rube Goldberg–like sequence of signal pathways, another aspect of the overall "strategy" of cellular information transmission is that signaling pathways almost always begin with the environmental signal molecule binding to a protein molecule specialized for information transfer, called a *receptor.* The LDL receptor discussed earlier is involved in the transport of material into cells (see Fig. 1.12). However, most other receptors are proteins whose task is to transmit and transduce information to the cell from the extracellular environment. Receptors distinguish among the large number of external signaling molecules (e.g., various hormones, neurotransmitters, growth factors) through the usual protein mechanism of highly specific binding.

Three broad classes of receptors, called *receptor families,* are particularly important in physiological function and are discussed in this chapter and Chapter 2. Two of these families, the *G-protein–coupled receptors* (GPCRs) and the *receptor tyrosine kinases* (RTKs), are intrinsic membrane proteins of the plasma membrane. These membrane receptors bind the signal molecule in the extracellular environment, and the signal is then communicated intracellularly

Vet-ebooks.com

through a Rube Goldberg sequence of "differences that make a difference." The third class of receptors is the *nuclear receptors.* These are not membrane proteins but rather intracellular proteins that transduce signals from oily, lipidic molecules that can easily enter the cell. Signaling molecules that bind and activate nuclear receptors include steroid and thyroid hormones, fat molecules in the diet, and derivatives of vitamins A and D. The information transduction pathway of nuclear receptors is simpler than that of the membrane receptors in that nuclear receptors are themselves direct regulators of gene transcription; that is, nuclear receptors are transcription factors. Binding of the signal molecule activates the nuclear receptor so that it is then able to bind directly to specific regions of deoxyribonucleic acid (DNA) and stimulate the binding of ribonucleic acid (RNA) polymerase to, and thus production of, messenger RNA from the particular gene or genes in that region of DNA. An example involving the female-specific production of egg protein in hens is discussed later in the chapter.

Specific Physiological Information Is Inherent in the Receptor/Ligand Complex, Not in the Hormone/Neurotransmitter Molecule

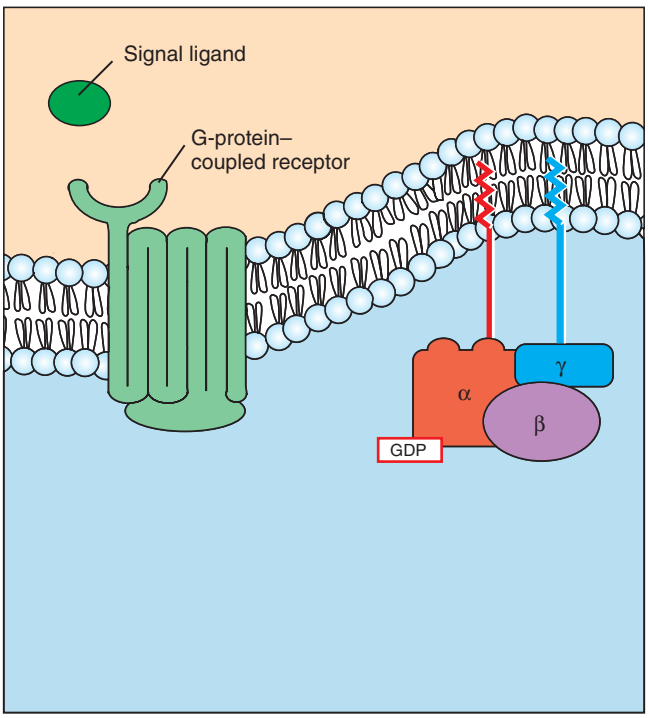
Before discussing specific receptors, it is useful to elaborate on some important points about the nature and regulation of the information transfer between the external signal molecule and receptor. This text provides ample evidence that the same hormone and especially the same neurotransmitter molecule can bind to different receptors. These different receptor-binding events send different information to the cell from the same external signal molecule. For example, acetylcholine is bound by two different receptors, the nicotinic ion channel described earlier and the muscarinic receptor, which is a GPCR, not an ion channel, and sends completely different information to the cell. The hormone/neurotransmitter itself does not contain any specific information; rather, it is a simple signal, such as a phone ringing. One must answer the phone to receive the information. The information content of the hormone/neurotransmitter is really contained in the three-dimensional shape of the receptor molecule. The change in the shape of the receptor on binding the hormone/neurotransmitter is the specific message to the cell.

Cells can make themselves more or less sensitive to the signal of the hormone/neurotransmitter. For example, most cells respond to a prolonged period of exposure to hormone/neurotransmitter by reducing their sensitivity to that molecule. For membrane receptors, one way is to internalize the receptors by endocytosis, fuse the endosome with a lysosome, and digest the receptor. Typically, membrane receptor number is decreased by endocytosis in response to a sustained high concentration of ligand. This is called *downregulation* of the receptor. This process allows the cell to adapt to high ligand concentrations. Receptor-ligand interaction is a true chemical equilibrium, so the proportion of receptor-ligand complexes, which determines physiological response, depends on the concentration of both receptors and ligands. In the presence of a high ligand concentration, a decrease in receptor number returns the binding equilibrium to the normal proportion of bound/unbound receptors. This allows the cell to respond to increases and decreases in ligand, even at high concentrations of ligand. Another way of regulating the response to a hormone/neurotransmitter is to alter the binding function of the receptor, e.g., by phosphorylating it, so that its affinity for the ligand is reduced (*desensitization*) or increased (*hypersensitization*). Nuclear receptors appear to be less subject to short-term regulation of responsiveness, but at least some nuclear receptors require constant turnover by proteolytic breakdown and new synthesis to function.

G-Protein–Coupled Receptors Are the Largest Family (a *Superfamily*) of Receptors and Help Regulate Almost All Physiological Processes

It would be difficult to exaggerate the importance and versatility of information processing that begins with a signal molecule binding to GPCRs. There are approximately 900 GPCRs in humans (Table 1.2). There are an even greater number in animals that depend more on olfaction, with about 1300 in rodents, because smell is mediated by different odorants binding to different GPCRs. An estimated 40% to 50% of all commercial drugs act in a GPCR pathway, exemplifying the importance of GPCRs to medicine. All GPCRs share a similar molecular shape; they are an integral membrane protein composed of a single polypeptide chain that passes in and out of the plasma membrane seven times, resembling a snake (Fig. 1.14). As a result, two other names for GPCRs are *seven-transmembrane receptors* and *serpentine receptors.* However, the name GPCRs is more revealing about their mechanism because all also share the same "next step" in their Rube Goldberg signal sequence; they activate a molecular "on-off switch" known as a *G protein,* so called because it is a *guanosinetriphosphatase* (GTPase).

GPCRs bind to a particular type of G protein (another of the many "families" of informational proteins), which is a membrane-associated trimeric protein composed of α, β, and γ subunits.



• **Fig. 1.14** G-protein–coupled receptor (GPCR) and the heterotrimeric G protein. The hundreds of GPCRs share a similar protein shape, snaking in and out of the membrane seven times. Thus GPCRs are also called *serpentine receptors* and *heptahelical receptors.* These receptors interact with a membrane-associated guanosinetriphosphatase (GTPase) molecule composed of three different polypeptide subunits *(heterotrimeric).* The heterotrimeric G protein is not an intrinsic membrane protein, but rather associates with the membrane through lipid tails inserted into the membrane. *GDP,* Guanosine diphosphate.

Vet-ebooks.com

TABLE 1.2 Partial List of G-Protein–Coupled Receptors (GPCRs)

Receptor/Receptor Family[a]	Example of Function	Drug Ligands
α-Adrenergic	Regulates vasculature	Phenylephrine, oxymetazoline
β-Adrenergic	Regulates heart and vasculature	Atenolol, propranolol
Angiotensin	Principal regulator of blood pressure	Losartan
Calcitonin	Regulates bone resorption	[b]
Cannabinoid	Unknown but found widely in brain	Marijuana and derivatives
Dopamine	Movement, cognition, and emotions	Chlorpromazine, bromocriptine
Frizzled	Regulates proliferation and differentiation, particularly in stem cells	[b]
Gastrin	Regulates acid secretion in stomach	Pentagastrin
Glucagon	Regulates "starvation" response	Exendin-4
Histamine	Mediates inflammation and allergy	Diphenhydramine, chlorpheniramine
Muscarinic	Secretion of hormones and neurotransmitters	Atropine, carbachol
Olfactory	Mediates smell	[b]
Opioid	Mediates analgesia	Morphine, codeine, heroin
Opsins	Mediates light transduction in retina	[b]
Prostaglandin	Vasodilation	Sulprostone
Serotonin[c]	Regulates gut motility, behavioral arousal, feeding, circadian rhythms	Sumatriptan, ketanserin
Vasopressin	Regulates water balance of body	Terlipressin, desmopressin

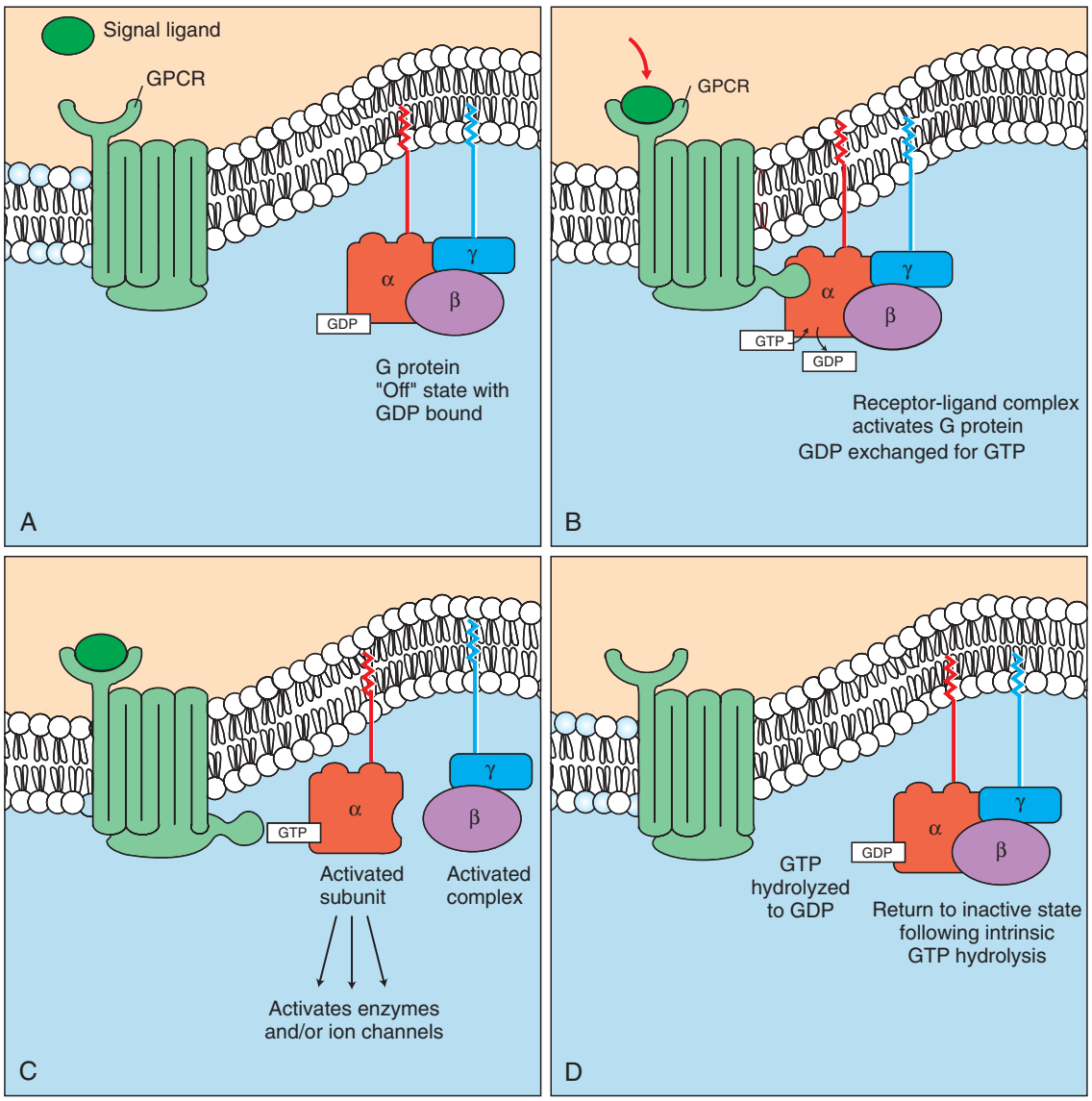
[a]In most cases, receptor is named for its ligand.
[b]None commonly known.
[c]One member of the serotonin receptor family is not G-protein–coupled.

Thus this type of G protein is called the *heterotrimeric G protein* ("three different subunits"). The heterotrimeric G proteins bind directly to the cytoplasmic domain of a GPCR. Although not intrinsic membrane proteins, heterotrimeric G proteins are closely associated with the plasma membrane through lipid molecules that are posttranslationally added to the subunits and insert into the lipid bilayer (see Fig. 1.14).

As noted, G proteins are molecular "on-off switches" that are also GTPases activated by the binding of a signal molecule to its cognate GPCR. That is, in addition to binding to GPCRs, G proteins also bind *guanosine triphosphate* (GTP) and hydrolyze it to *guanosine diphosphate* (GDP). The binding and hydrolysis of the GTP to GDP underlie the biochemical mechanism of the on-off switch. In Fig. 1.15A, the unstimulated GPCR does not bind to the heterotrimeric G protein, which is in its "off" state by virtue of the α subunit having GDP and the β and γ subunits bound to it. In Fig. 1.15B, a signal ligand binds to its GPCR, activating the receptor and the G protein. The activation of the G protein takes the form of dissociation of the β/γ complex from the α subunit, which allows the α subunit to exchange GDP for GTP. The principal "on" activity of the G protein is represented by the G_α subunit with GTP bound to it. GTP-bound G_α stimulates a variety of enzymes and ion channels that send the signal into the cytoplasm (see Fig. 1.15C), as discussed in the next section. However, the $G_{\beta\gamma}$ complex, once thought to be only an inhibitory factor for the G_α subunit, is now known to activate certain K^+ channels itself and inhibit certain voltage-dependent Ca^{2+} channels. After stimulating the next element in the signal pathway, the activated GTP-bound G_α subunit returns to an inactive, "off" state as a result of its intrinsic GTPase activity (see Fig. 1.15D). That is, the bound GTP is hydrolyzed to GDP, and the $G_{\beta\gamma}$ complex rebinds to the G_α subunit, returning it (and the $G_{\beta\gamma}$ complex) to its inactive state, awaiting the next ligand-receptor–binding event.

As noted earlier, one of the aspects of the Rube Goldberg analogy is that the same conserved types of molecules are often used in many different pathways. Among the many protein "differences that make a difference" to transmit information, one of the most widely used is a GTPase that has on-off states based on whether GTP or GDP is bound to it. Thus the heterotrimeric G proteins that couple to GPCRs are only one type of GTPase protein acting as an on-off switch in signaling pathways. Most other members of the G-protein (GTPase) superfamily are simpler and resemble the G_α subunit alone. For example, one such class of these *small G proteins,* called *Rabs,* helps mediate the membrane fusion processes that underlie exocytosis and endocytosis, as discussed earlier. All G proteins share evolutionarily conserved GTP binding and enzymatic hydrolysis sites and a similar on-off mechanism; when GTP is bound, the protein is "on," and when GDP is bound, the protein is "off." Chapter 2 discusses a particular small G protein, Ras, that plays a crucial role in regulating cell division and whose dysfunction plays a major role in cancer. Consequently, G proteins in general are discussed in Chapter 2, and this discussion focuses on signaling mechanisms linked to the heterotrimeric G protein specifically.

Vet-ebooks.com



• **Fig. 1.15** Duty cycle of heterotrimeric G-protein, a GTPase that acts as a molecular "on-off switch." See text for further details. (A) Unstimulated GPCR is not bound to the heterotrimeric G protein. (B) Signal ligand binds to its GPCR, activating receptor and G protein. (C) GTP-bound G_α subunit stimulates a variety of enzymes and ion channels that send the signal into the cytoplasm. (D) After stimulating next element in signal pathway, the activated GTP-bound G_α subunit returns to inactive, "off" state because of its intrinsic GTPase activity. *GDP,* Guanosine diphosphate; *GPCR,* G-protein-coupled receptor; *GTP,* guanosine triphosphate.

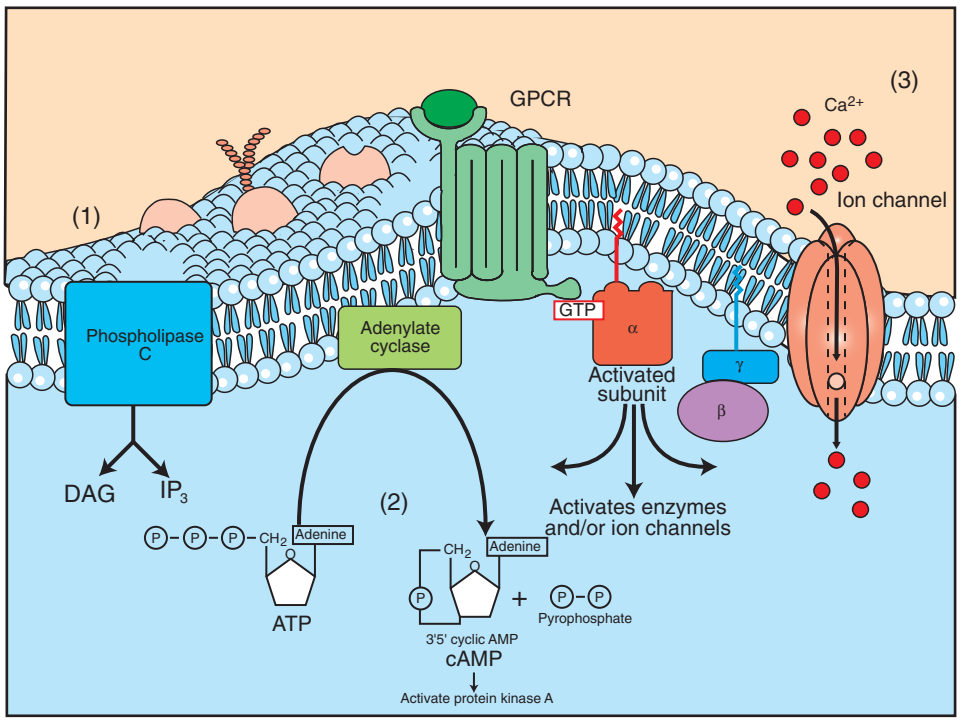
Most G-Protein–Linked Information Is Sent to the Cytoplasm by Second Messengers

As previously noted, the active (heterotrimeric) G protein stimulates an enzyme or ion channel associated with the plasma membrane. The ensuing change in ion channel or enzyme function can alter the membrane potential or cause certain molecules/ions to change their concentration in the cytoplasm. Those cytoplasmic ions and molecules linked to receptor-ligand binding are called *second messengers.* A second messenger is an ion or molecule that carries the information within the cytoplasm of a cell in response to a signal on the outside surface of a cell (the first message), such as the binding of a hormone or neurotransmitter, or to an electrical event. Most G-protein–linked information is transduced into the cytoplasm in this manner. One of the major advances in our understanding of the molecular basis of physiological signaling is the realization that there are only a few second-messenger systems within animal cells. The most important include the following (Fig. 1.16):

1. Two second messengers, inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG), are produced by G-protein activation of an enzyme phospholipase C (PLC) (see Fig. 1.19 and later discussion).
2. Changes in the concentration of *cyclic adenosine monophosphate* (cAMP).
3. Changes in Ca^{2+} concentration within the cytoplasm.

Clearly, there are many more GPCRs than second messengers. This means that several receptor-mediated events are converted into the same intracellular signal. How does the cell sort out this information? Different cells respond differently to the same

Vet-ebooks.com



• **Fig. 1.16** Activated α subunit of the G protein (G_α) can activate enzymes and ion channels, leading to second-messenger signaling within the cytoplasm. Three principal second messengers send the G-protein–coupled receptor (GPCR) information to the cytoplasm. These arise from the activation of ion channels and enzymes stimulated by G_α. The second messengers are (*1*) increases in the concentration of inositol 1,4,5-trisphosphate (IP_3) in the cytoplasm and increases in the concentration of diacylglycerol (DAG) in the plasma membrane, both as a result of the breakdown of a rare membrane phospholipid, phosphatidylinositol 4,5-bisphosphate (PIP_2) by phospholipase C, another G_α-stimulated enzyme; (*2*) changes in the concentration of cyclic adenosine monophosphate (cAMP), a special hydrolytic breakdown product of adenosine triphosphate (ATP) created by the enzyme adenylyl cyclase, which can be activated or inhibited by α subunits; and (*3*) changes in Ca^{2+} concentration within the cytoplasm resulting from transport of Ca^{2+} through gated channels stimulated by G_α. *GTP,* Guanosine triphosphate.

second-messenger ion/molecule as a result of the specialized function and makeup of that cell (i.e., the differentiated state it achieved during the development of the animal). For example, smooth muscle cells respond differently to activation of muscarinic acetylcholine receptors (see Table 1.2) than do nerve cells because the two cells have different proteins that are responsible for their specialized tasks. However, this is only part of the answer, and the specificity of response to the same second messenger and to activation of similar or identical receptors remains an important open question in physiology.

Ca^{2+} Transport Across Plasma and Intracellular Membranes Is an Important Second Messenger

The transport of Ca^{2+} ions through gated channels across the plasma membrane and across intracellular membranes (e.g., endoplasmic reticulum) is a major second-messenger system for physiological information transfer. The available evidence suggests that the major role of Ca^{2+} *within* cells is as a physiological signal. In the extracellular compartment, the major physiological function of Ca^{2+} is as the principal mineral of bone. Ca^{2+} is an excellent ion for use as a second messenger because the cytoplasmic concentration of Ca^{2+} is extremely low, about 10^{-7} mol/L in a resting cell. Increases in intracellular Ca^{2+} concentration can be (1) detected easily because the background noise is so low and (2) achieved easily because the Ca^{2+} concentration [Ca^{2+}] in the ECF and in some cellular compartments, such as the endoplasmic reticulum and mitochondria, is 10^4 higher than in the cytoplasm (see Table 1.1). Thus there is an enormous driving force for Ca^{2+} diffusion into the cytoplasm under most conditions.

Although many GPCRs use Ca^{2+} as one part of their intracellular pathway, the interaction is more complex than usual, as discussed shortly. Thus we focus here on a simpler and very important example of Ca^{2+} as a second messenger that has already been discussed: the role of Ca^{2+} in regulating the actomyosin ATPase of muscle.

Increased [Ca^{2+}] in the cytoplasm alters cellular function by binding to any of several Ca^{2+}-binding proteins that serve as control proteins. Troponin is one Ca^{2+}-binding protein already mentioned. Reviewing the example of striated muscle contraction from the Ca^{2+} point of view, Ca^{2+} (second messenger) diffuses through gated channels in the endoplasmic reticulum (sarcoplasmic reticulum) of muscle in response to electrical events (first message) on the plasma membrane of the muscle cell. The diffusion of Ca^{2+} from the concentrated storehouse of the sarcoplasmic reticulum increases [Ca^{2+}] in the cytoplasm of the muscle cell, where it binds to troponin. On binding Ca^{2+}, troponin changes its interaction with tropomyosin, which now moves to allow myosin heads access to the actin of the thin filament. The actomyosin ATPase is activated, and muscle contraction ensues.

Calmodulin is a Ca^{2+}-binding protein that plays an important control function in almost all animal cells. As with troponin, calmodulin binds Ca^{2+} when the cytoplasmic [Ca^{2+}] increases. The

Vet-ebooks.com

Ca^{2+}/calmodulin complex activates a large number of different cellular processes. In most such cases, but not all, the Ca^{2+}/calmodulin complex binds to and activates an enzyme. One such enzyme, a protein kinase, is involved in the excitation-contraction coupling in smooth muscle (Fig. 1.17), not discussed earlier with the striated muscle types. Protein kinases in general catalyze the hydrolysis of ATP and couple it to the simultaneous phosphorylation of other proteins, as follows:

$$\text{ATP} + \text{Protein} \xrightarrow[\text{Protein kinase}]{Ca^{2+}/\text{calmodulin activated}} \text{Protein phosphate} + \text{ADP}$$

In the case of smooth muscle, the particular protein kinase is *myosin kinase,* which, as its name implies, specifically phosphorylates myosin. This phosphorylation increases the affinity of the myosin heads for actin filaments, thus allowing cross-bridging to actin. On formation of the cross-bridge, myosin strokes past the thin filament, producing filament sliding, contraction, and force production by smooth muscle. Cessation of contraction is achieved by cleavage of the phosphate from the myosin by another enzyme, *myosin phosphatase.*

Thus initiating smooth muscle contraction involves a Rube Goldberg sequence in which environmental stimulation of a smooth muscle cell causes an increase in the intracellular [Ca^{2+}], the second messenger. This in turn leads to a cascade of cause and effect. Increased intracellular [Ca^{2+}] causes calmodulin to bind Ca^{2+}. The Ca^{2+}/calmodulin complex activates the myosin kinase. This enzyme phosphorylates the myosin head, allowing it to cross-bridge to actin. Cross-bridging leads to actomyosin activation, causing filament sliding observed as muscle contraction at the tissue level.

This discussion of Ca^{2+} as a second messenger emphasizes one of its major physiological functions: as the second messenger responsible for mediating contraction for all types of muscle (skeletal, cardiac, and striated), although the details of each pathway differ.

Cyclic AMP Is Produced by Activation of a Membrane-Bound Enzyme in Response to Hormone/Neurotransmitter Binding to Receptors

Changes in the activity of membrane-associated enzyme activities are an important mechanism of transmitting information across the cell membrane and are used by most GPCRs. Binding of a signaling molecule to receptors on the extracellular face of the plasma membrane changes the activity of an enzyme located on the cytoplasmic face. The enzyme catalyzes a breakdown reaction; one or more of the breakdown products released into the cytoplasm are second messengers. One important such second-messenger system, and the first to have been discovered, is the hydrolytic breakdown of ATP to 3′,5′-adenosine monophosphate, or cAMP, by the enzyme *adenylyl cyclase* (previously called *adenyl cyclase* and *adenylate cyclase*). Cyclic AMP is the second messenger, and adenylyl cyclase is turned on or off as a result of the binding of various hormones and neurotransmitters to cell surface receptors.

As summarized in Fig. 1.18, three distinct membrane proteins interact to produce cAMP: (1) any of several receptors, including many GPCRs; (2) the heterotrimeric G protein; and (3) the catalytic protein that actually hydrolyzes ATP to cAMP. Their interaction provides an example of the ability of biomembranes to organize biochemical reactions in space. The likelihood of three proteins colliding and thus being able to interact is much greater in the two-dimensional "phospholipid sea" than in the three-dimensional cytoplasm.

A large number of different hormones/neurotransmitters that bind to different membrane receptors use cAMP to transmit information across the membrane. Among the GPCRs (see Table 1.2) and their hormones/neurotransmitters that use cAMP as their second messenger are *β-adrenergic receptors* that bind epinephrine or norepinephrine, increasing cAMP production and providing important regulation to almost all tissues. The starvation message carried by the binding of glucagon to its receptor (see Chapter 34) is carried to the cytoplasm by an increase in cAMP. *Vasopressin* (also called *antidiuretic hormone,* ADH) binding to its receptors in kidney cells uses cAMP to regulate urine production (see Chapter 33). A number of therapeutic drugs bind to these same receptors and mimic or prevent the physiological action of the hormone/neurotransmitter that normally binds to the receptor.

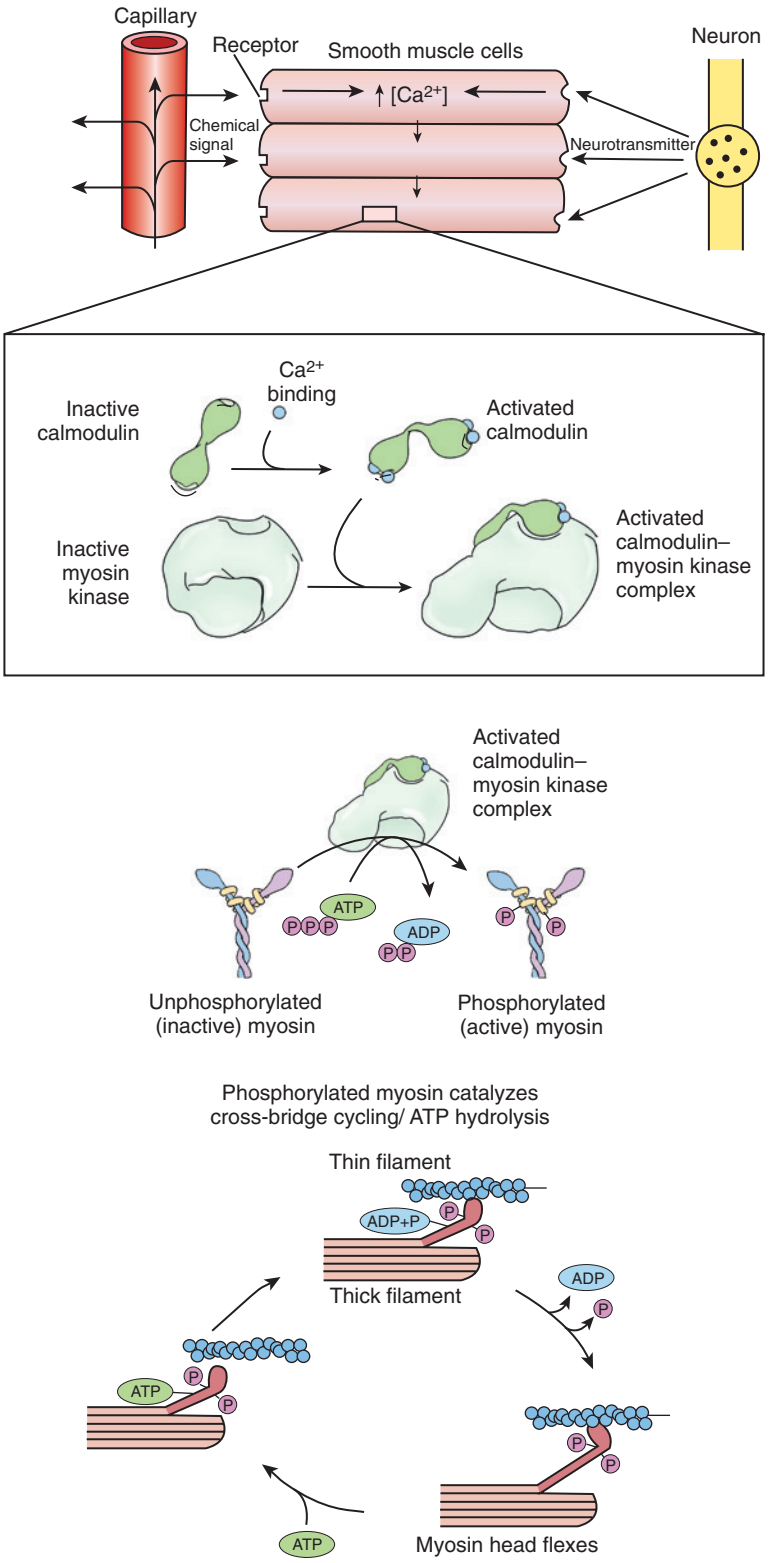
After ligand binding, the ligand-receptor complex is able to bind to and activate the regulatory G protein (see Fig. 1.15B). The G protein in turn changes shape and binds to the catalytic subunit, altering its shape and regulating its ability to bind ATP, and hydrolyzes the catalytic subunit to cAMP (see Fig. 1.18). There are two types of G proteins in the adenylyl cyclase system, which differ in their α subunit. The G_s (more specifically, $G_{\alpha s}$, s for stimulatory) activates the catalytic subunit; this is the G protein shown in Fig. 1.18. A different G protein, the α subunit of G_i, inhibits adenylyl cyclase when activated. Some diseases are the result of the binding of bacterial toxins to the G proteins. Cholera symptoms result in part from the binding of the toxin of the bacteria *Vibrio cholerae* to the G_s protein, and the irreversible activation of the G_s protein, which in turn irreversibly activates the catalytic subunit. Pertussis (whooping cough) toxin binds irreversibly to and activates G_i, thus inactivating the enzymatic activity.

As suggested by the inhibitory G protein (G_i), regulated decreases in cAMP concentrations are an important part of the cAMP second-messenger system. There are two mechanisms for such decreases: decreasing the rate of cAMP production and eliminating cAMP after formation. The former is achieved by G_i inhibiting the catalytic subunit. Certain inhibitory receptors specifically interact with G_i. Opium and drugs derived from it, such as codeine and morphine, are examples of signaling molecules that bind to inhibitory GPCR (opioid) receptors, activate G_i, and inhibit production of cAMP. Other examples are norepinephrine and epinephrine acting through α_2-adrenergic receptors. Recall that these same neurotransmitters activate adenylyl cyclase when bound to β-adrenergic receptors. This is another example of the principle that the receptor/ligand complex contains the information, not the hormone/neurotransmitter itself.

The other control on cAMP levels is elimination of cAMP after formation. This is regulated by the enzyme *cyclic nucleotide phosphodiesterase* (PDE). This enzyme hydrolyzes the 3′ ester bond of the phosphate to the sugar to produce "plain" 5′ AMP (see Fig. 1.18). As with myosin kinase discussed earlier, PDE is a Ca^{2+}/calmodulin-activated enzyme, so in many cells the activities of the Ca^{2+} and cAMP second-messenger systems antagonize one another.

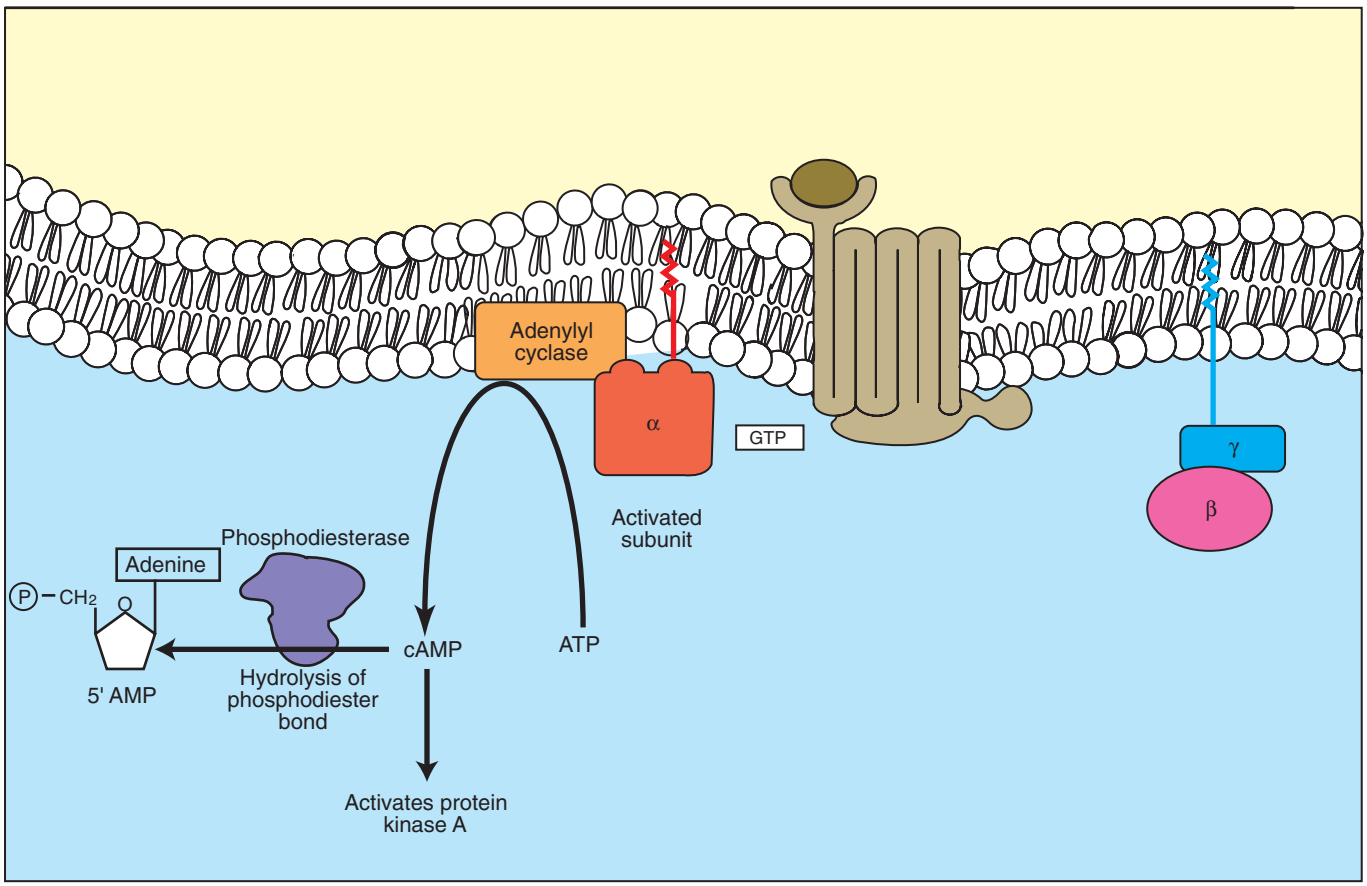
The increase or decrease in cAMP concentrations most often affects cell function through cAMP's interaction with a particular protein kinase. This protein kinase is called *cAMP-dependent protein kinase,* or *protein kinase A* (PKA). This protein kinase is distinct from the Ca^{2+}/calmodulin-dependent protein kinase discussed earlier, although the basic outline of action is similar. PKA is activated by binding cAMP. The higher the concentration of cAMP

Vet-ebooks.com



• **Fig. 1.17** Role of Ca^{2+} and calmodulin in the regulation of smooth muscle contraction. Smooth muscle regulation is more complex than regulation of striated muscle, and the account here is a simplification. Smooth muscle can be stimulated to contract by a variety of stimuli, including neural signals and soluble chemical signals, as shown here. These external signals all stimulate increased intracellular [Ca^{2+}], which leads to smooth muscle contraction. In the presence of increased intracellular [Ca^{2+}], the Ca^{2+} ions bind to calmodulin, activating it by causing a conformational change. In smooth muscle cytoplasm, the activated Ca^{2+}/calmodulin complex activates myosin kinase, which catalyzes the phosphorylation of myosin. Phosphorylated, activated myosin in turn catalyzes actin-dependent ATP hydrolysis (cross-bridge cycling). Thus smooth muscle contraction is thick filament regulated, because changes in myosin activate cross-bridging, whereas striated muscle contraction is thin filament controlled, because changes in troponin and tropomyosin of the thin filament activate cross-bridging. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate.

Vet-ebooks.com



• **Fig. 1.18** Activity of cyclic adenosine monophosphate (cAMP) as a second messenger. Cyclic AMP is generated through G-protein–coupled receptor-linked activation of adenylyl cyclase, causing hydrolysis of adenosine triphosphate (ATP) to cAMP. The cAMP thus generated binds to and activates a specific protein kinase, protein kinase A, which in turn can phosphorylate and change the activity of various cellular substrates. When generated, cAMP is broken down by phosphodiesterase, which hydrolyzes cAMP to "normal" AMP (i.e., 5′ AMP). *GTP,* Guanosine triphosphate.

in a cell, the greater is the number of active PKA molecules. The activated kinase binds to proteins and ATP, hydrolyzing the ATP and phosphorylating the protein. As shown in previous examples, this phosphorylation alters the activity of the target protein, altering its particular characteristic function: catalysis, transport, coupling, and so forth.

Mammals respond to a stressful stimulus by increasing the force and rate of heart contraction, among other physiological effects. This increase in force demonstrates the role of cAMP as a second messenger and the role of Ca^{2+} in GPCR signaling, and it is another example of physiological Rube Goldberg devices based on allosteric changes in proteins. The stressful stimulus causes the adrenal medulla to release epinephrine to the blood, and sympathetic nerves release norepinephrine to the heart. Both catecholamines bind to β-adrenergic GPCRs on the cardiac muscle cells. The receptor-ligand interaction stimulates adenylyl cyclase by way of G_s, increasing intracellular [cAMP], thus increasing PKA activity. PKA phosphorylates a number of substrates in the cardiac muscle cells, including voltage-dependent Ca^{2+} channels in the plasma membrane. In the phosphorylated state, these channels remain open somewhat longer in response to membrane potentials above threshold. Consequently, more Ca^{2+} enters the cell for a given electrical stimulation than at lower levels of cAMP. The increase in Ca^{2+} allows more troponin to bind Ca^{2+}; more tropomyosin moves out of the way of myosin heads, causing more cross-bridging and more force production. (Rube Goldberg would have loved modern physiology!)

Another cyclic nucleotide, *cyclic guanosine monophosphate* (cGMP), also serves as a second messenger but is not nearly as widely used as cAMP. Cyclic GMP is the second messenger stimulated by opsins (see Table 1.2) in the rod cells of the retina underlying vision and also causes relaxation of some vascular smooth muscle, including that responsible for penile erection and clitoral engorgement (i.e., blood flow into the corpus cavernosum of both tissues). The role of cGMP in erections is mediated by its activation of cGMP protein kinases, similar to cAMP action via PKA. Activation of cGMP-dependent protein kinase causes relaxation of certain smooth muscles, including those responsible for blood flow to the corpus cavernosum. This has an important clinical correlate; the drug Viagra (sildenafil) inhibits the breakdown of cGMP by a cyclic nucleotide phosphodiesterase, thus increasing blood flow to the penis, and aids erection, but only if neural signals (i.e., sexual stimulation) have stimulated cGMP production initially. This is a good example of how the multistep pathway of cell signaling provides multiple potential sites for appropriate therapeutic intervention; a drug that simply stimulated cGMP production

Vet-ebooks.com

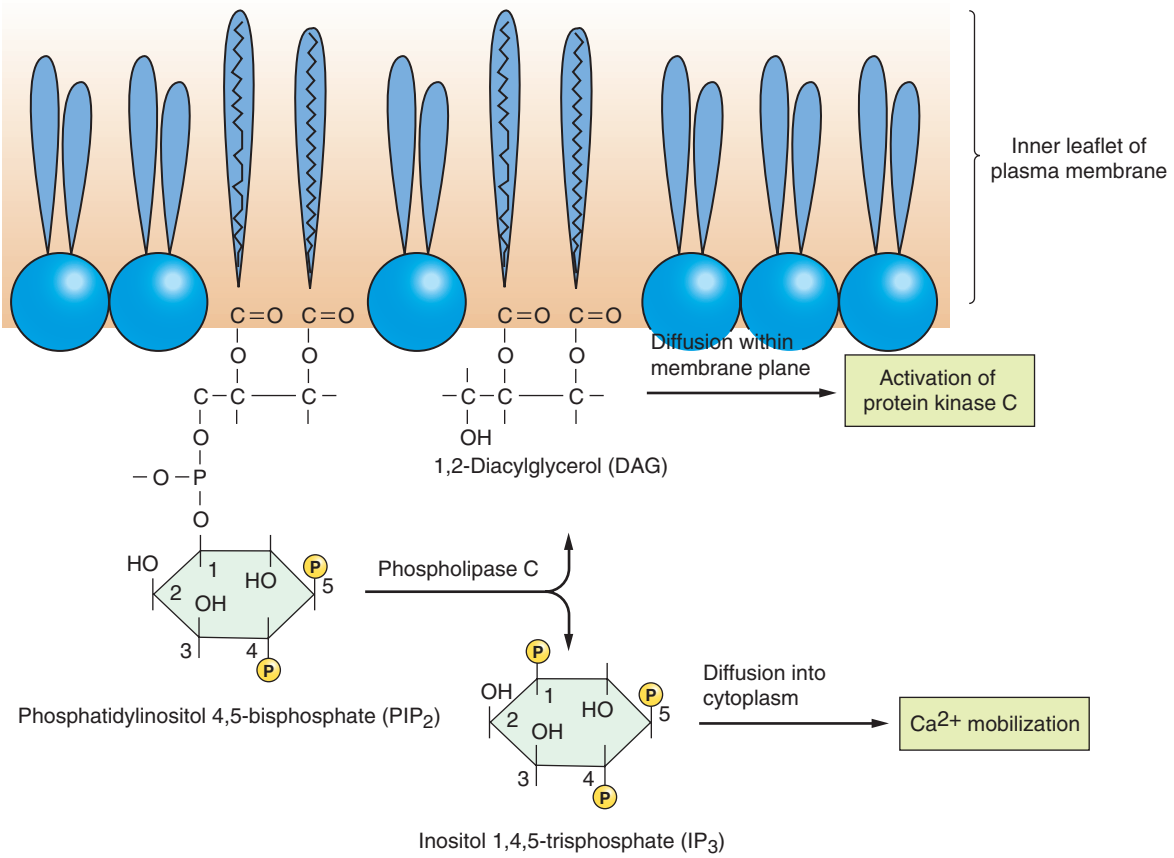
would cause inappropriate erections, whereas inhibiting its breakdown aids timely erections. Although used mostly by men, sildenafil is also occasionally used for stallions to assist them in "covering" a mare.

In addition to activating protein kinases, cAMP and cGMP can also bind directly to and cause opening of a class of ligand-gated ion channels, cyclic nucleotide–gated ion channels. These ion channels are atypical in that their structure resembles voltage-gated K^+ channels, but they open by directly binding a cyclic nucleotide. These channels play an important role in smell, for which cAMP is the relevant second messenger. In vision, as noted earlier, cGMP is the second messenger, and mutations in the cyclic nucleotide–gated ion channels of cones are responsible for most forms of complete color blindness (but which is rare).

The examples of physiological control by second messengers discussed thus far are short time-scale changes (seconds to hours), which historically have been the purview of "physiologists." It has become increasingly clear, however, that most, if not all, major signals have longer-term (days and weeks) effects based on changes of gene transcription, which in turn mediates changes in growth, differentiation, and long-term behavior. For example, cAMP is now known to be an important regulator of gene transcription controlling learning, production of gametes, and cell division. The effect of cAMP on gene expression is the result of PKA phosphorylating a specific transcription factor associated with cAMP signaling ("cyclic AMP response element binding protein," or CREB). Although space does not allow further discussion of the transcriptional roles of "classic" physiological signal pathways, when dealing with signal pathways it is worthwhile to keep in mind the disclaimer in the first paragraph: only a highly simplified account of cell function is presented here!

The Receptor-Mediated Hydrolysis of a Rare Phospholipid of the Plasma Membrane Produces Two Different Second Messengers With Different Actions

Another second-messenger system differs from both Ca^{2+} and cAMP in that *two* distinct second-messenger molecules are produced as a result of an enzymatic activation by a single receptor/ligand complex. *Phosphatidylinositol* (PI) is a membrane phospholipid that can accept additional phosphate groups by reaction with the –OH groups on the inositol (Fig. 1.19). *Phosphatidylinositol 4,*



• **Fig. 1.19** Hydrolysis of a membrane lipid to produce two second messengers. After appropriate receptor and G-protein activation, the rare membrane phospholipid shown to the left, phosphatidylinositol 4, 5-bisphosphate (PIP_2) is hydrolyzed into two separate second messengers by phospholipase C. The phosphate "head" of the PIP_2 molecule is cleaved to produce the soluble messenger inositol 1,4,5-trisphosphate (IP_3), which mobilizes intracellular Ca^{2+}, as well as the lipidic messenger diacylglycerol (DAG), which remains in the membrane and activates protein kinase C.

Vet-ebooks.com

5-bisphosphate (PIP_2) is the membrane phospholipid that is broken down to produce two important second messengers. PIP_2 is hydrolyzed to DAG and IP_3 by a receptor-mediated enzyme called PLC or phosphoinositidase. Although many distinct processes are controlled through the PIP_2 path, it plays a particularly important role in control of growth and of receptor-mediated secretion. The effect of acetylcholine acting through muscarinic receptors (*not* the nicotinic receptor/ion channel of the nerve-muscle synapse) is often transmitted and transduced through activation of the PIP_2 pathway.

The events involved in the receptor-mediated production of IP_3 and DAG from PIP_2 are similar to those in the production of cAMP. The membrane system appears to consist of three distinct intrinsic membrane proteins: (1) any of several different GPCRs, including the muscarinic acetylcholine receptor and the receptors for some growth factors; (2) a heterotrimeric G protein, similar but not identical to G_s of the cAMP pathway; and (3) the hydrolytic enzyme PLC. A hormone/neurotransmitter or growth factor binds to the receptor, forming a receptor/ligand complex. This complex activates the G protein, which in turn activates the hydrolytic enzyme. At present, only a stimulatory G activity on PLC is known; there is no evidence for an inhibitory G activity in this system.

The activation of the hydrolytic enzyme increases the concentration of IP_3, which is water soluble and thus diffuses through the cytoplasm. IP_3 binds to and opens ligand-gated Ca^{2+} channels in the endoplasmic reticulum. This releases Ca^{2+} from that high [Ca^{2+}] compartment into the cytoplasm. Ca^{2+} thus becomes the "third messenger" in this system (although this term is not in widespread use) and is another example of a role of Ca^{2+} in GPCR signaling. The ensuing increase in cytoplasmic [Ca^{2+}] affects cellular function by the same mechanisms outlined earlier for Ca^{2+} as a second messenger (e.g., binding to calmodulin), with the Ca^{2+}/calmodulin complex in turn activating various enzyme activities. In receptor-mediated secretion, for example, the binding of acetylcholine to muscarinic receptors in the pancreas (the organ that secretes digestive enzymes) causes an increase in PIP_2 breakdown and an increase in cytoplasmic IP_3. The IP_3 opens ligand-gated Ca^{2+} channels in the endoplasmic reticulum, and intracellular [Ca^{2+}] increases. The process then becomes similar to smooth muscle contraction. Calmodulin binds Ca^{2+}, and the complex activates a protein kinase. However, rather than activating myosin, as for smooth muscle, activation of this protein kinase causes exocytosis of secretory vesicles (membrane bubbles full of secretory product) with the plasma membrane, releasing the enzymes into an extracellular space contiguous with the gut.

DAG is also produced on activation of PLC, but it is not at all soluble. DAG diffuses in the plasma membrane, binding to and activating membrane-associated protein kinase, *protein kinase C* (PKC). PKC is not an intrinsic membrane protein and can bind reversibly to the cytoplasmic face of the plasma membrane. PKC phosphorylates other proteins and changes their activity. Because of the membrane-bound character of the enzyme, most evidence indicates that PKC phosphorylates membrane proteins such as receptors and ion channels, regulating their function. In the case of the secretory response to some hormone/neurotransmitter stimulus, PKC generally acts separately but additively with IP_3 to produce the response. As with cAMP and PKA, however, much interest focuses on longer-term effects of DAG activation of PKC, particularly its role in growth control and cancer. A class of chemicals long known to promote the onset of cancer, phorbol esters, is a potent substitute for DAG at activating PKC. It is now known that PKC indirectly activates an important transcription factor involved in cell proliferation, nuclear factor kappa B (NF-κB). Thus, as a second messenger and as with cAMP, DAG has both short-term effects and longer-term transcriptional effects.

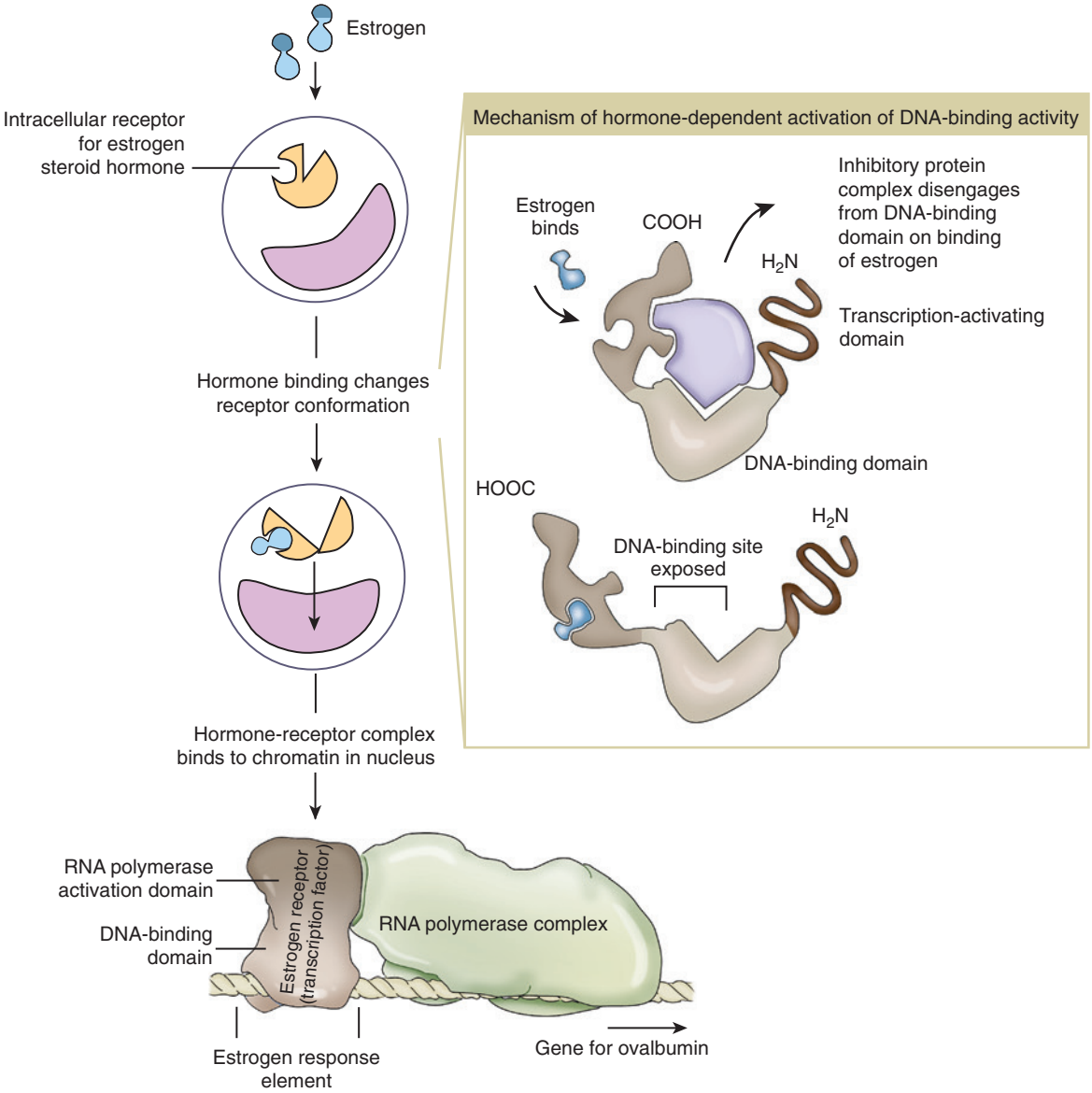
Steroid Hormones and Other Lipid Signals Interact With Nuclear Receptors, Which Are Transcription Factors Within the Cell

Nuclear receptors are another large class of protein molecules specialized for information transmission and transduction. Nuclear receptors are sufficiently numerous and diverse that they compose a superfamily of evolutionarily conserved and related receptors, as with the GPCRs. All nuclear receptors are transcription factors that respond to the binding of their cognate lipid signal by regulating which genes are expressed within particular cells under particular conditions. Accordingly, one of the conserved features of nuclear receptors is their DNA-binding domain, which can bind directly to specific sequences of DNA (promoter regions) that control the expression of the neighboring gene(s) (Fig. 1.20). As with all other proteins, the DNA-binding function of nuclear receptors is based on their shape. The DNA-binding domain, for example, is a part of the protein shaped into "fingers" by a zinc ion. These *zinc fingers,* also found in many other transcription factors, fit into the grooves of the double helix of DNA at the appropriate base-pair sequence.

Recall that steroid hormones are soluble in oily solvents and are able to diffuse through the lipid bilayer without the mediation of transport proteins. Thyroid hormones are also lipophilic and diffuse through the lipid bilayer. Additionally, several lipid-soluble nutrients are also signaling molecules, including vitamins A and D. Vitamin A is required for vision because it is the covalently bound cofactor for the opsin GPCRs, but it also plays a role in embryonic development. Vitamin D controls Ca^{2+} metabolism. Similarly, saturated and unsaturated fats in the diet are also known to provide signals that control their own breakdown and metabolism and to regulate the differentiation of fat cells (adipose tissue). Consequently, the receptors for these lipid signals are soluble proteins within the target cell. The cellular location of the nuclear receptors varies. Some receptors can be found in the cytoplasm before ligand binding, whereas others are largely restricted to the nucleus (after their initial synthesis in the cytoplasm), but all are functional as transcription factors in the nucleus after activation. The lipid-soluble hormone/nutrient diffuses from the blood into the cell and binds to its receptor, and the hormone/receptor complex is, as in previous examples, the physiologically active entity that ultimately triggers a cellular response. As noted earlier, because the nuclear receptor complex is itself a transcription factor, steroid and thyroid hormones do not require a second messenger; the hormone/receptor complex is itself active within the cell, altering gene expression.

A well-studied example of nuclear receptor action as a regulated transcription factor, with some relevance to veterinary medicine, is the action of estrogen on the reproductive tracts of female chickens (see Fig. 1.20). Estrogen is the principal female sex hormone of birds and mammals, and, of course, hens lay eggs in which the embryo and yolk are surrounded by eggwhite. The principal protein of eggwhite is ovalbumin, which is secreted by the epithelial cells of the avian oviduct as the ovum slides by. Thus one of estrogen's targets in female chickens is oviduct epithelial cells. Estrogen enters the cytoplasm of these cells and binds to its receptor, the *estrogen*

Vet-ebooks.com



• **Fig. 1.20** Steroid hormone action as illustrated by control of ovalbumin expression by estrogen in hens. The steroid hormone estrogen penetrates the lipid bilayer passively because of the oil solubility of the steroid. Inside the cell, the estrogen binds to a cytoplasmic receptor, the estrogen receptor. The binding of estrogen to its receptor causes the receptor protein to change conformation, which in turn changes the DNA-binding activity of the receptor. The hormone/receptor complex enters the nucleus and binds to regulatory sequences of deoxyribonucleic acid (DNA), the estrogen response element. This binding, in turn, activates ribonucleic acid (RNA) polymerase. This initiates transcription of the ovalbumin gene, an estrogen-responsive gene, to produce messenger RNA (mRNA), which is ultimately translated into the ovalbumin protein for secretion.

receptor. The hormone/receptor complex, but not the ligand-free receptor, is able to mediate estrogen-specific, essentially female-specific, gene transcription. The estrogen receptor complex binds to a sequence of DNA, called an *estrogen response element,* that controls the transcription of a neighboring gene, for ovalbumin in this case. In other cells of the female, binding of the estrogen receptor to the estrogen response elements of other genes would cause these other female-specific genes to be transcribed and ultimately expressed as a protein (e.g., proteins in yolk of egg). Different steroids bind to different receptors (e.g., male sex hormone testosterone binds to testosterone receptor), which bind to different response elements, leading to different genes expressed (e.g., male-specific gene expression).

Differential gene expression and its regulation were initially pursued primarily by molecular biologists. It rapidly gained importance in physiology, however, and will do so soon in veterinary medicine. Humankind will have fewer scruples about controlling gene expression in animals other than their own species (a fact well illustrated by studies of cancer in mice as discussed in the next chapter). Indeed, understanding control of gene expression may prove more important to veterinary students in the near term than for students of human medicine.

Vet-ebooks.com

CLINICAL CORRELATIONS

Peripheral Edema

History

You examine a 2-year-old cow that has been grazing on poor-quality pasture. The owner states that the cow seems to have a poor appetite, walks slowly, and stands apart from the rest of the herd. The cow has developed swelling beneath the skin of her brisket and ventral thorax.

Clinical Examination

On clinical examination, you find a listless cow standing in a pasture littered with various metal objects. Examination of the cardiovascular system reveals distended jugular veins and abnormal heart sounds characterized by irregular sloshing sounds throughout the cardiac cycle that drastically muffle the first and second heart sounds. Subcutaneous edema (swelling) can be seen throughout the chest and abdomen but most prominently in the dependent ventral areas of the thorax. Pushing on these swollen areas leaves a dent (pitting edema).

Comment

This is a characteristic history of a cow with hardware disease. The cow, grazing on a pasture littered with metal debris, swallows nails, wire, and so forth. Because these objects are heavier than the feed, they drop into the reticulum, a stomach chamber located just caudal to the diaphragm and heart. With the contractions of the reticulum, a metal object migrates through the reticular wall, diaphragm, and pericardium, leading to an inflammatory response in the pericardium (pericarditis). The resulting process is caused by both inflammation and possible secondary bacterial infections from a contaminated metal object traversing regions of the gastrointestinal tract that contain numerous microorganisms, before the object penetrates the pericardium. An inflammatory exudate fills the pericardial sac; it muffles the heart sounds, and a sloshing sound may be heard on auscultation. As this exudative fluid fills the pericardial sac, it limits the pumping efficiency of the heart by limiting its filling during diastole and by obstructing venous return to the heart (see Chapter 21). The result is left-sided heart failure because the heart cannot circulate (pump) the blood throughout the body. This causes the blood to accumulate initially, leading to an increased hydrostatic pressure in the veins and capillaries. As the capillary hydrostatic pressure rises, capillary filtration is favored over reabsorption, and water leaves the capillary and accumulates in the interstitial space. This accumulated interstitial fluid, primarily as the result of increased capillary filtration, is seen clinically as edema. The other common cause of edema is decreased capillary colloidal osmotic pressure from low serum protein. However, this does not usually play a part in hardware disease.

Treatment

Treatment includes surgical removal of the foreign object or objects, antiinflammatory agents, and antibiotic treatment for the pericarditis. Even though considerable inflammation is present, a secondary bacterial infection often contributes to the response. In such an advanced case, however, treatment often is not completely successful.

Horse With Hyperkalemic Periodic Paralysis

History

A 3-year-old Quarter Horse mare, showing in her third halter class of the evening, began having spontaneous muscle twitches (fasciculations). She collapsed in the ring and was unable to rise. She had a long trailer ride the day before the show, and the rider has been feeding her a slightly different diet because less of her normal hay was available. She has never had these signs before.

Clinical Examination

On examination, her temperature, pulse, and respiratory rate are all increased. She is sweating with muscle fasciculations, and she appears to have muscle cramps. She is not very responsive, and she is unable to rise when she is placed lying on her breastbone (sternal recumbency).

Comment

This mare likely has hyperkalemic periodic paralysis (HYPP). It is the result of an autosomal dominant mutation with a single DNA base-pair substitution. This produces a population of abnormal voltage-gated sodium channels at the muscle cell membrane in which the inactivation gate is thought to have trouble docking into its shut position. It is likely that some combination of multiple shows (increased exercise), the change in diet, and the long trailer ride may have initiated this episode. Different feeds, both natural and commercial, can vary widely in their potassium (K^+) content, and the change in diet may have produced increased serum and extracellular K^+. Increased exercise has been shown to have a similar effect. This excess extracellular K^+ causes a slight depolarization in membrane potential, which then results in an opening of sodium (Na^+) voltage-gated channels, along with a failure of the defective Na^+ voltage-gated channels to inactivate. This leads to a persistent flux of Na^+ into the cell, abnormally prolonging the change in membrane potential. This abnormal voltage change can be associated with prolonged inactivation of normal Na^+ voltage gated channels, resulting in loss of electrical excitability from normal stimulation and thus causing paralysis. Stress, such as that associated with the preceding long trailer ride, has been shown to exacerbate the onset of such episodes in HYPP.

Treatment

Initial treatment for horses experiencing an HYPP episode includes feeding corn syrup to promote insulin-mediated K^+ transport out of the extracellular fluid by the Na^+, K^+ pump. More severe cases can be treated with calcium gluconate; the calcium increases the threshold for muscle stimulation. Also, intravenous dextrose can be used to promote intracellular movement of K^+ by stimulating ATP production, in turn activating the Na^+, K^+ pump. Long-term management includes diet and management modifications to limit stress and K^+ intake. Acetazolamide can also be used to increase renal excretion of K^+, as well as stimulate insulin secretion, which stabilizes blood glucose and K^+. Horses can die of severe episodes. Horses can be tested for this mutation.

Practice Questions

1. Increasing the extracellular K^+ concentration will:
- **a.** Have no effect on the resting membrane potential.
- **b.** Cause the resting membrane potential to decrease (i.e., cause the inside to become less negative with respect to the outside).
- **c.** Cause the resting membrane potential to increase (i.e., cause the inside to become more negative with respect to the outside).
- **d.** Increase the concentration potential for K^+ across the plasma membrane.
- **e.** Require the Na^+, K^+ pump to work harder to pump K^+.

2. G proteins are similar to receptors in that both:
- **a.** Bind extracellular signaling molecules.
- **b.** Interact directly with adenylyl cyclase catalytic subunits.
- **c.** Have activated and inactivated states dependent on ligand binding.
- **d.** Are extracellular protein molecules.
- **e.** Directly activate a protein kinase activity.

Vet-ebooks.com

3. Which of the following statements concerning intracellular Ca^{2+} is *false*?
- **a.** It is a second messenger for hormones and neurotransmitters.
- **b.** It is responsible for excitation-contraction coupling in smooth muscle.
- **c.** An increase in its concentration in a nerve terminal stimulates the release of a neurotransmitter.
- **d.** It activates protein kinase A.
- **e.** Its concentration is increased in the presence of IP_3.

4. If, in a particular capillary bed, the plasma oncotic pressure were to increase and hydrostatic pressure remained constant:
- **a.** More blood plasma would filter from the capillaries.
- **b.** The transport effect would be similar to decreasing hydrostatic pressure.
- **c.** One would suspect a deficiency in blood protein levels.
- **d.** One would suspect an increase in extracellular fluid protein concentrations.
- **e.** Fluid reabsorption on the venous side of the capillary bed would decline.

5. Substance X is found to be at a much higher concentration on the outside of a cell than in the cytoplasm, but no transport of X from the extracellular fluid to the cytoplasm occurs. Which of the following statements is inconsistent with this situation?
- **a.** Substance X has the same electrochemical potential outside and inside the cell.
- **b.** Substance X is large, is poorly soluble in oil, and has no transport proteins in the membrane.
- **c.** Substance X is an ion, and the measured membrane potential is the equilibrium potential calculated by the Nernst equation.
- **d.** Substance X is a steroid molecule.
- **e.** Substance X is actively transported from the cell to the extracellular fluid.

Bibliography

Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*. 6th ed. New York: Garland Science; 2014.

Lodish H, Berk A, Kaiser CA, et al. *Molecular Cell Biology*. 8th ed. New York: Macmillan Learning; 2016.

Luttrell LM. Reviews in molecular biology and biotechnology: transmembrane signaling by G protein-coupled receptors. *Mol Biotechnol*. 2008;39(3):239–264.

Novac N, Heinzel T. Nuclear receptors: overview and classification. *Curr Drug Targets Inflamm Allergy*. 2004;3(4):335–346.

Valberg SJ. Diseases of muscles. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby; 2014.

Valberg SJ, Carlson G. Muscle cramping. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby; 2014.

Vet-ebooks.com

2

Cancer: A Disease of Cellular Proliferation, Life Span, and Death

STEVEN R. HEIDEMANN

KEY POINTS

1. Cancer arises from genetic dysfunction in the regulation of the cell cycle, cell life span, and cell suicide.

Control of the Cell Cycle (Proliferation)

1. Cell division is the result of a clocklike cell cycle.
2. Cyclin-dependent kinases (CDKs) are the "engines" driving the cell cycle.
3. The CDK "engines" are controlled by both throttle (oncogene) and brake (tumor suppressor) controls.

Growth Factor Pathway: Principal Stimulator of Cell Proliferation

1. The cell cycle is stimulated by growth factors that bind to and activate receptor tyrosine kinases.
2. The *ras* oncogene contributes to many cancers and serves as a model for understanding *small G proteins.*
3. The MAP kinase pathway leads to the expression of cyclins and other stimulators of the cell cycle.
4. The MAP kinase pathway also mediates the stimulation of the cell cycle by cell adhesion.

Tumor Suppressors: Inhibitors of Cell Cycle

1. Checkpoints in the cell cycle are manned by tumor suppressors.
2. The retinoblastoma and P53 proteins are the main gatekeepers for the cell cycle.

Mechanisms Regulating Cell Suicide and Cell Life Span

1. Apoptosis is the process of cell suicide.
2. Resistance to apoptosis via the intrinsic pathway is a hallmark of cancer.
3. Many types of cancer cells suppress immune attack and so avoid apoptosis via the extrinsic pathway
4. Cellular life span is determined by DNA sequences at the ends of chromosomes.

Tumor Origin and the Spread of Cancer

1. Cancer cells may be related to stem cells.
2. Death by cancer is usually the result of its spread, not the original tumor.
3. Growth of solid tumors depends on development of new blood vessels.

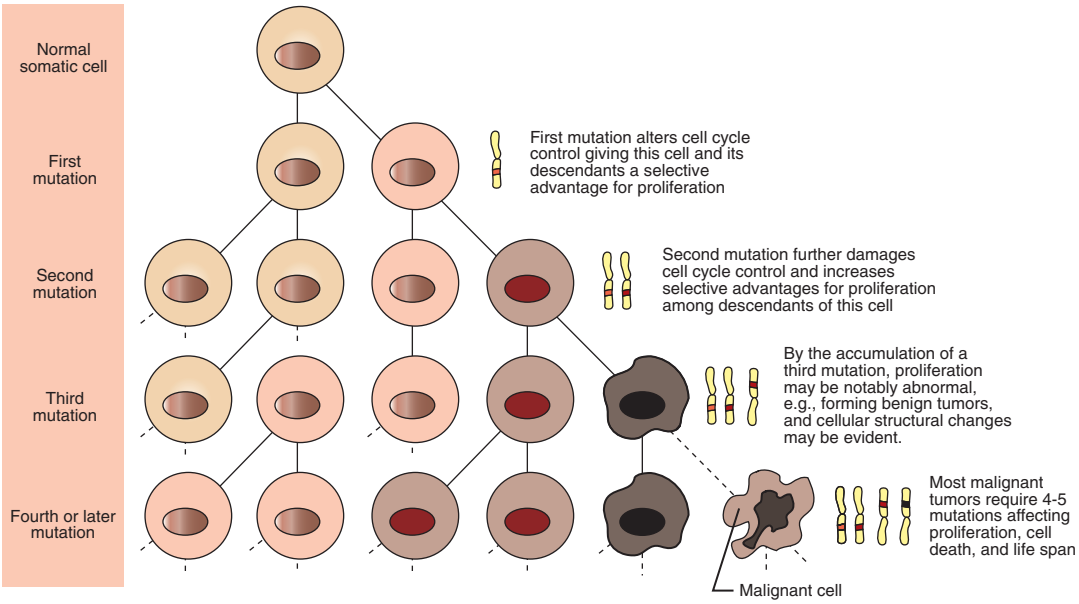
Prospective Cancer Therapy

1. Cancer therapy has a hopeful but challenging future.

Traditionally, cancer was (and often still is) first detected in humans and domestic animals by clinicians feeling for an unusual mass of cells, tumor cells. Thus cancer is quite intuitively a disease affecting cellular growth. In the last 30 years, enormous progress has been made in understanding several normal control pathways that regulate cell growth, as well as how these *Rube Goldberg* pathways (see Chapter 1) go wrong in cancer.

The first path to be unraveled, long thought to play a major role in cancer, was the pathway controlling cellular proliferation. By the early twentieth century, cellular proliferation was known to occur by a regular clocklike cycle of chromosomal doubling followed by mitotic division, called the *cell cycle.* However, progress on understanding the molecular control of the cell cycle did not begin until the 1980s from the study of cancer cells but importantly also from the study of the proteins synthesized by fertilized sea urchin eggs; how frogs ovulate; and how yeast cells divide. Cell number depends not only on new cells being formed by cell division but also on cells dying. As a result of studying in detail the history and fate of every cell that arises during embryonic development to form a soil roundworm (a nematode), it was discovered that cells are programmed to commit "suicide." That is, cells can actively kill themselves using metabolic machinery during normal development (e.g., a tadpole losing its tail) and also if the cell has internal damage, such as mutations or oxidative stress. This surprising discovery quickly led to the realization that not only do cancer cells divide inappropriately, but they are also resistant to programmed death and thus continue to divide despite internal damage. The final general process affecting cellular growth is that normal cells, like the organisms they are part of, have a characteristic life span. However, cancer cells were long known to be "immortal," being able to divide indefinitely. How cells age, or become immortal, was not understood until the process of chromosomal duplication was studied in a ciliated protozoan, similar to the familiar *Paramecium* of college biology laboratories.

Vet-ebooks.com



• **Fig. 2.1** Clonal basis of cancer. Cancer is the result of the accumulation of mutations in a cell lineage of somatic (nongamete) cells of the body. Beginning with a normal cell, mutations occur by chance or by environmental inputs, such as radiation or cancer-causing chemicals, and accumulate to cause cancer.

As these examples illustrate, our understanding of cellular proliferation, cellular life span, and cell suicide came in large part from the study of problems that first seemed distant from the cancer seen in the clinic. As such, the recent progress on cancer is an unusually dramatic example of the importance of understanding basic biology to understand medicine. The vast majority of cancer studies are conducted on humans and in mice, *the* preeminent animal model for cancer, and using cultured cells derived from human and mouse tumors. The much smaller number of studies on domestic animals strongly indicates that the principles derived from humans and mice are generally applicable. However, it is also clear that humans and mice differ in a few aspects of cancer, and thus there are likely to be "special" aspects of cancer for each species. In the case of domestic animals, different species and breeds are known to have differing frequencies of various cancers. For example, the Bibliography at the end of this chapter includes two reviews on the cancer biology of dogs. Veterinary practitioners will need to carefully evaluate the application of knowledge about human and mouse tumors for their patients.

Cancer Arises from Genetic Dysfunction in the Regulation of the Cell Cycle, Cell Life Span, and Cell Suicide

Cancer is a genetic disease (but not usually a hereditary disease) and a uniquely cellular disease. As shown in Fig. 2.1, tumors and other cancers arise from the division of a single mutant cell whose descendants accumulate several additional mutations to become increasingly damaged with respect to control of cellular proliferation, life span, and cell death. That is, cancer is a genetic disease caused by the accumulation of mutations in body cells, such as those of the epithelial cells lining the lungs or the secretory epithelial cells of the mammary glands.

All the cells of a tumor can trace their ancestry back to a single cell that developed an initial deleterious mutation. This first mutation usually occurs in a gene controlling proliferation, such that the cell produces a mutant protein[1] that is a dysfunctional, more permissive regulator of the cell cycle. This greater "permissiveness" provides the mutant cell with more opportunity to proliferate, and it thus has a selective advantage compared with its normal neighbors. Perhaps because of this selective advantage, or because of continued exposure to mutagens (e.g., cigarette smoke, agricultural chemicals), a descendant of this cell accumulates another mutation that also affects some aspect of the cell cycle or cell death. This increases the doubly mutant cell's selective advantage further still, and the downward spiral of increasingly abnormal, dividing cells begins to spin out of control. Scientists agree that this accumulation of mutations in individual genes is necessary for cancer to develop, but some think it is not sufficient. Rather, they argue that cancer only results when the accumulation of mutations eventually leads to large-scale genetic instability, such that whole chromosomes are gained and lost. The majority of spontaneous tumors do have cells with abnormal sets of chromosomes, a phenomenon called *aneuploidy.* Whether aneuploidy is necessary for cancer remains to be seen, but there is no disagreement that cancer cells are in some way badly damaged with respect to genes controlling growth.

The mutations leading to cancer are the same type as those that underlie Mendel's familiar laws of heredity. These include base-pair changes, deletions or additions of nucleotides in the gene, and translocation of one piece of a chromosome to another. However, it is important to understand that the *cells* in which the mutations are occurring are different than those underlying Mendel's laws of inheritance. Mendelian inheritance results from mutations occurring in the *germ line* of the organism. These are

[1]For the many instances when a gene and protein share the same name, this chapter adopts a widely used, but by no means universal, convention for distinguishing genes and their cognate proteins. Gene names are in italics and all lowercase (e.g., *ras*), whereas the protein name is not italicized has one or more capital letters (e.g., Ras). This convention is used throughout in preference to the various species-dependent conventions also used in the literature.

Vet-ebooks.com

the cells that will produce gametes, either sperm or eggs, and whose deoxyribonucleic acid (DNA) will be passed down to every cell of the offspring following fertilization. The mutations leading to cancer are occurring in nonreproductive cells throughout the body, called *somatic cells.* These are passed down only to a limited number of other somatic cells by cell division, not to offspring through sexual reproduction. Thus although cancer is a genetic disease, only about 10% of the time is it a "hereditary disease," that is, the result of mutation inherited from a parent. In general, cancer appears to be the result of the accumulation of mutations leading to genetic instability in a particular lineage of somatic cells.

Traditionally, cancers are divided into categories based on the cell type involved. Carcinomas are cancers of epithelial cells; sarcomas are derived from connective tissue or muscle; and leukemias are cancers of blood-forming cells. There are many subdivisions based on specific cell types and location of the tumors. However, these names are traditional only; they do not reflect any fundamental differences in the biology of the cancer. Rather, it is now clear that cancers of all types share broadly similar types of dysfunctions controlling cell proliferation, cell suicide, and cell life span.

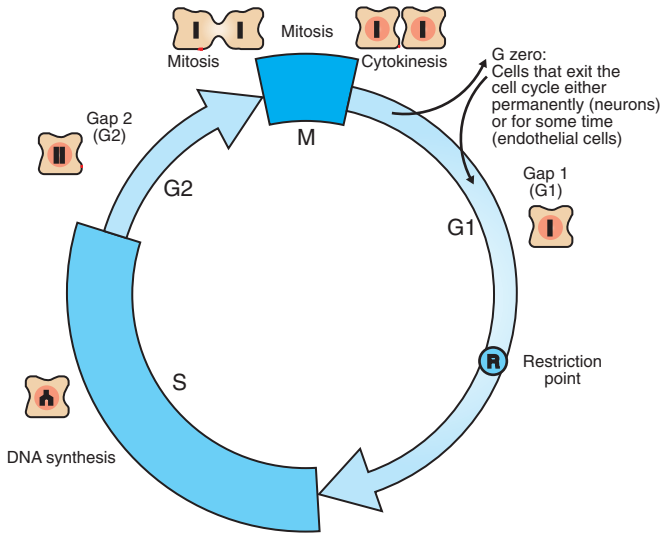
Control of the Cell Cycle (Proliferation)

Cell Division Is the Result of a Clocklike Cell Cycle

The *Rube Goldberg device* that controls cell growth is particularly complex, with many, many more components than the "garage door opener" of Fig. 1.13. To explain these pathways, we begin with the cell cycle that, like the carriage house door, is near the end of the system of control. That is, most of the control elements feed "downstream" to control the cell cycle or intersect with some aspect of cell cycle control.

Fig. 2.2 shows the classic diagram of the cell cycle in which the cell changes its state toward division, progressively going around the diagram, like the hands of a clock. For most mammalian cells in culture, the duration of one cell cycle varies between 18 and 30 hours. Two phases of the cell cycle were identified first and seemed to be where the most important events of the cell cycle occurred. One is *synthesis (S) phase,* during which the DNA is duplicated. The second is *mitosis (M) phase,* during which the duplicated chromosomes are separated to opposite sides of the cell and the cytoplasm divides. In addition to the obvious need for such events if cells are to reproduce, note that both phases must be highly precise. It is crucial for the cell that DNA synthesis produces *exactly* twice the original amount of DNA, no more and no less. Otherwise, there will not be two identical copies of the genetic material to pass on to two identical cells. Similarly, the machinery segregating the duplicated chromosomes during mitosis must partition exactly equal numbers and types of chromosomes to daughter cells, or the cells will be aneuploid. If DNA is not precisely replicated, or if the chromosomes are not properly aligned, the cell cycle is halted by *checkpoints,* as described later.

However, the events during G1 ("gee-one") and G2 phases remained a mystery. The "G" stands for *gap,* because of the decades-long gap in our understanding of what was happening during this time. Although it was suspected that the cell was preparing itself for DNA synthesis during G1 and preparing for mitosis during G2, the nature of these "preparations" proved difficult to determine. In the mid-1980s, work initially conducted on frog oocytes revealed that specialized protein kinases were activated during G1 and G2 to drive the cell into S phase and M phase, respectively.



• **Fig. 2.2** The mammalian cell cycle. Cell proliferation occurs by a clock-like progression of phases in which characteristic events occur. The most familiar is M phase (mitosis), during which the cytoplasm and replicated chromosomes are distributed to the daughter cells. Cells then enter G1, during which a "decision" is made whether or not to go forward with the cell cycle; this is the R (restriction) point. The events in G1 then allow S (synthesis) phase to proceed, during which the DNA is replicated to produce exactly two copies. After DNA synthesis, the cell prepares for mitosis during G1, and the cycle is complete. Although cells in culture typically go around the cycle continuously, most cells in the body divide only occasionally. These quiescent cells, as well as cells such as neurons that never divide after differentiation, are in G0, a nondividing phase. Under appropriate stimulation, cells can then exit G0 and are said to reenter the cell cycle.

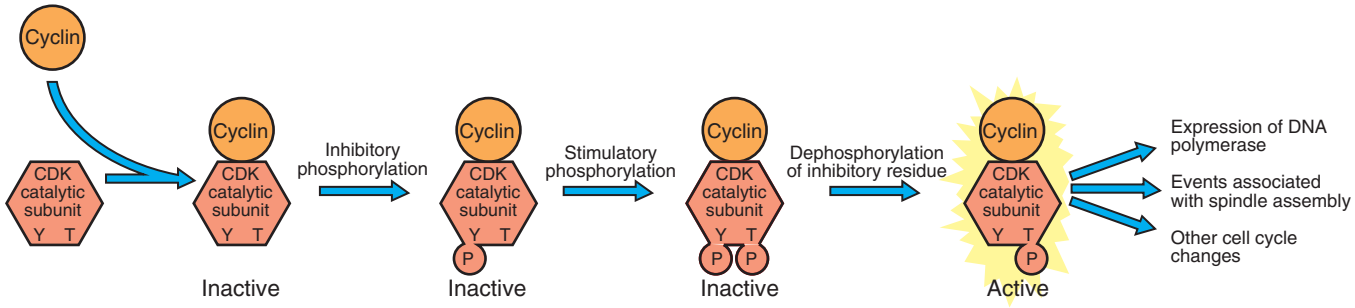
These special protein kinases are now called *cyclin-dependent kinases* (CDKs).

Cyclin-Dependent Kinases Are the "Engines" Driving the Cell Cycle

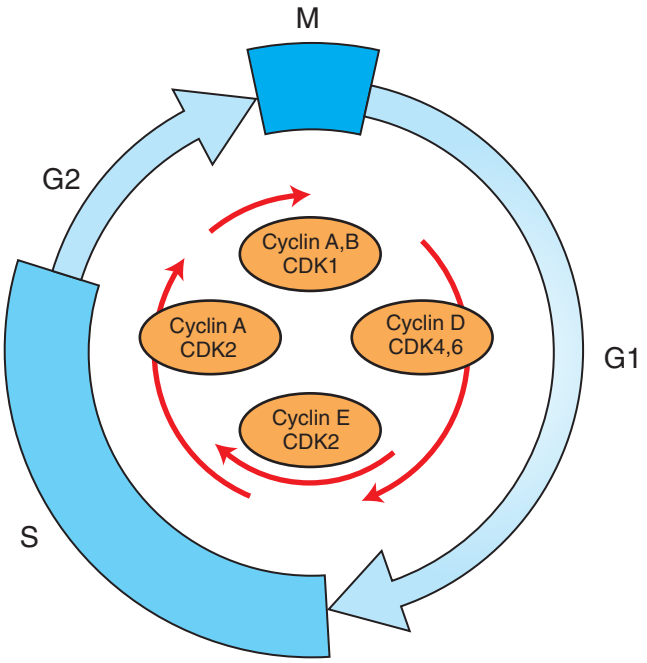
Recall from Chapter 1 that protein kinases, which are enzymes that phosphorylate other proteins, are important as elements of signaling pathways. For example, the second messenger cyclic adenosine monophosphate (cAMP) acts by activating protein kinase A (see Fig. 1.18), and diacylglycerol as a second messenger activates protein kinase C (see Fig. 1.19). Protein kinases play a major role in many aspects of control of the cell cycle; most importantly, CDKs, when activated, can directly cause a cell to enter either S phase or mitosis, whether the cell is ready or not.

Active CDKs are composed of two different types of protein subunits (Fig. 2.3). The catalytic subunits (numbered CDK1, CDK2, etc.) are the subunits that have enzymatic activity for hydrolyzing adenosine triphosphate (ATP) and transferring the phosphate group to a protein substrate. The other subunit is an activator of the catalytic subunit and is called a *cyclin;* the abundance of this protein increases and decreases during the cell cycle (i.e., the protein concentration cycles up and down during the cell cycle, hence the name). Different cyclins are specific for various CDKs and for the different phases of the cell cycle. The various cyclins are identified by letters, such as cyclin A and cyclin B. Cyclins must reach a threshold concentration to activate the catalytic subunit, and the threshold is achieved as a result of protein accumulation from new synthesis during the G phases.

Vet-ebooks.com



• **Fig. 2.3** Activation of the cyclin-CDK "engines" of the cell cycle. Activation of CDKs depends on the association of a cyclin with a catalytic subunit and then an appropriate pattern of inhibitory and stimulatory phosphorylations on the catalytic subunit. *CDK,* Cyclin-dependent kinase.



• **Fig. 2.4** Cyclins and CDKs around the cell cycle. Different phases of the cell cycle are associated with and driven by different cyclin-CDK pairs, as shown here. *CDK,* Cyclin-dependent kinase.

When the cyclins have bound to their appropriate catalytic subunit, the cyclin-CDK complex as a whole is activated by achieving a particular state of *phosphorylation.* There are inhibitory sites of phosphorylation around amino acid 15 of the catalytic subunit, and these must be dephosphorylated. There is also a stimulatory phosphorylation site at amino acid 167, and this must be phosphorylated for cyclin-CDK activity. When activated, the CDK phosphorylates various substrates associated with either S phase or M phase. For example, the cyclin-CDK complex responsible for mitosis directly phosphorylates the protein filaments that strengthen the nuclear membrane (lamins). This phosphorylation causes the filaments to disassemble, in turn allowing the nuclear membrane to dissolve, which is an early event of mitosis.

The different phases of the cell cycle are controlled by different cyclin-CDK pairs, as shown in Fig. 2.4. Thus the complex of CDK1 with either cyclin B or cyclin A is the particular CDK pair responsible for driving the cell into mitosis. Cyclins E and A interacting with CDK2 play important roles in initiating and maintaining DNA synthesis in S phase. Cyclin D interacting with either CDK4 or CDK6 functions in late G1 in a "decision" by the cell to commit to DNA synthesis. This decision is called the *restriction (R) point* and is discussed in the later section on tumor suppressors.

Given the importance of cyclins and CDKs in driving the cell cycle, one would expect they would have some connection to cancer. Overexpression of cyclin D is associated with human and mouse breast cancer, and ablation of cyclin D provides some protection against breast cancer in mice. Virtually all multiple myelomas, a type of leukemia, show overexpression of cyclin D. Overexpression of cyclin A is strongly associated with some lung cancers and with testicular cancer of humans, and overexpression of cyclin E is associated with certain human leukemias. Curiously, in contrast to the cyclin subunit, the CDK enzymatic subunit is not known to be mutated in any common cancer.

The CDK "Engines" Are Controlled by Both Throttle (Oncogene) and Brake (Tumor Suppressor) Controls

The CDK-cyclin pairs are controlled by both stimulatory and inhibitory pathways, analogous to an automobile controlled by throttle and brake mechanisms. The throttle mechanisms are largely the result of the cell's environmental inputs. That is, various environmental cues, both soluble signal molecules and insoluble signaling molecules from (or on) neighboring cells, are required for cells to divide. However, the pathways sending inhibitory signals to the cell cycle, the "brakes" for cell division, are largely internal and activated by damage or stress to the cell. In general, these inhibitory signals are like the safety interlocks on an automobile. Just as one cannot start a car in gear, so the cell should not divide if DNA synthesis has not exactly duplicated all the genes and chromosomes, or if something is wrong with the mitotic spindle.

The environmental stimulatory signals for cell division can be as simple and nonspecific as availability of nutrients, to the extent that cells only divide when they have approximately doubled in size through synthetic growth. However, two more specific stimulators of the cell cycle are primarily implicated in cancer. One is the response to soluble growth factors found in the circulation and in the extracellular fluid surrounding cells (see Chapter 1). Growth factors are proteins secreted by a variety of other cell types that are required for the division, and indeed survival, of normal, noncancerous cells. Cancer cells, however, can divide and survive with little or no stimulation from growth factors because of the

Vet-ebooks.com

acquired ability to synthesize growth factors of their own or the inappropriate activation of downstream elements in the signaling pathway.

The second stimulatory pathway of general importance in cancer is cell attachment. From just a simple mechanical point of view, it is obvious that the cells of multicellular organisms must be tightly attached to one another and to their surrounding matrix (similar to tendon) or we would otherwise be jelly, juice, and bubbles on the floor. Also, however, attachment of cells to their surroundings is a source of specific and complex information to the physiology of the cell. One of the most important such messages is a "permissive" signal to divide. Normal cells must be anchored to some substrate to respond to other signals to divide. That is, most normal animal cells show *anchorage dependence of growth.* For this reason, vertebrate cells in culture are grown on the surface of a dish or flask, not in suspension the way bacteria are cultured. Again, cancer cells have lost this normal restriction on proliferation, and many cancer cells can divide and survive in suspension. The common test for the absence of anchorage dependence is growth in soft agar; cancer cells will, but normal cells will not, divide and form colonies when suspended in soft agar. Thus cancer cells can survive unattached while riding the circulation to relocate in a different tissue than that of the original tumor. In this way, cancer cells "go walkabout" and are able to spread through the body, a process called *metastasis,* which is ultimately the cause of death in most cases of cancer.

If attachment among normal cells had only a stimulatory effect on cell division, one might imagine that it would favor unlimited proliferation, overgrown organs, and tumor formation. It is not surprising then that normal cells with "enough" cell attachment are inhibited in cell division. When normal cells are, indeed, well attached and completely surrounded by their neighbors, proliferation is inhibited. Normal cells detect their surrounding cell density and at a threshold "take their foot off the accelerator," and proliferation ceases. This is called *contact inhibition of growth* and is in part regulated at the level of CDK activation. As a result, normal cells in culture only grow to form a "monolayer," one cell thick. Normal cells will not "pile on" and form multilayered clumps, basically minitumors, on the dish. You will not be surprised to learn that, yet again, cancer cells lose this "good neighbor" regulation of the cell cycle and are thus able to form tumors. Correspondingly in culture, cancer cells "overgrow the monolayer" and, like growth in soft agar, this is a common adhesion-based test to distinguish cancer cells from normal cells.

The Rube Goldberg pathways that underlie the proliferative signals of growth factors and adhesion, both stimulatory and inhibitory, are similar and intersect. These "throttle" contraptions begin with a soluble signal binding to a growth factor receptor and a "solid-state" signal about attachment to the surrounding tissue. However, both pathways quickly converge on the same stimulation pathway for conserved cell division. These stimulatory pathways are driven by proteins that were originally identified as being encoded by genes in viruses that caused cancer in animals. Thus these were named *oncogenes,* literally "cancer genes." A major breakthrough came with the discovery that these oncogenes were actually derived from the host genome, not genes normally encoded in the virus. That is, viruses had stolen cell cycle control genes from their animal host cell. Being viruses, they did not take good care of the animal cell cycle genes they stole. The stolen genes mutated into deranged cell cycle regulators. Subsequently, the same mutant genes that were found in viruses were found to explain many spontaneous cancers in humans and in the long-used experimental tumors of mice. The finding that cancer was caused by abnormal host genes helped confirm that cancer was a somatic genetic disease due to mutations in the tumor cells.

Further analysis revealed that these oncogenes often encode normal stimulators of the cell cycle, and the mutations involved had the effect of permanently activating an element in the cell cycle pathway. You can see how this would work based on the Rube Goldberg cartoon of Fig. 1.13. Note that all the elements in the garage door opener are stimulatory; if any one turns "on," a signal is sent "downstream" to cause the garage door to open. If the fish tank of the cartoon were to "mutate" by developing a leak, an "on" signal would be sent downstream of the fish tank, regardless of whether a car had pulled into the driveway. So it is with the oncogene elements controlling the cell cycle. If one of the elements mutates to turn itself "on," that is, acquired a *gain-of-function mutation,* it will stimulate cell division and contribute to cancer. To return to the automobile analogy, oncogenes represent a stuck throttle or accelerator pedal. The normal, well-behaved versions of the oncogene (a watertight fish tank before the bullet, Fig. 1.13) are called *proto-oncogenes.* Thus strictly speaking, oncogenes have their normal equivalent as proto-oncogenes. However, given this awkward usage, increasingly the normal versions are also informally called oncogenes, and it is usually clear from the context whether the mutant or normal version is being discussed. The molecules and molecular events of the oncogene pathway (also called the *growth factor* or *MAP kinase pathway*) are discussed later.

The mechanisms to stop the cell cycle, the "brakes," are called *checkpoints.* Progress through the cell cycle depends on appropriate conditions being reached within the cell before a "decision" is made to go ahead with division. The first such checkpoint occurs before S phase. During G1, the cell checks itself over particularly with respect to DNA damage. The cell has sophisticated pathways to detect and repair DNA damage, such as mismatched bases detected in the double helix. For needed repairs to take place, however, DNA synthesis is delayed; the checkpoint is "engaged." If the DNA is properly repaired, the checkpoint is disengaged, and after the delay, the cell goes ahead into S phase. However, if the DNA damage cannot be repaired, the checkpoint machinery is supposed to signal a more serious consequence. If the checkpoint is not disengaged after about a day, the cell "commits suicide." Thus the checkpoint (or braking machinery) is tied into both the CDK engines and the processes of cell suicide, as described later. Similarly, the second checkpoint is in mitosis and checks for proper mitotic spindle assembly and correct chromosome alignment. Here again, if damage is detected, there are repair mechanisms, and a properly repaired cell will go into M phase after a delay for repair. If no repair can be made, the cell commits suicide.

The molecules and their interactions that underlie both oncogene ("throttle") pathways and checkpoint ("brake") pathways are now covered in greater detail, beginning with the role of growth factors.

Growth Factor Pathways: Stimulators of Cell Proliferation

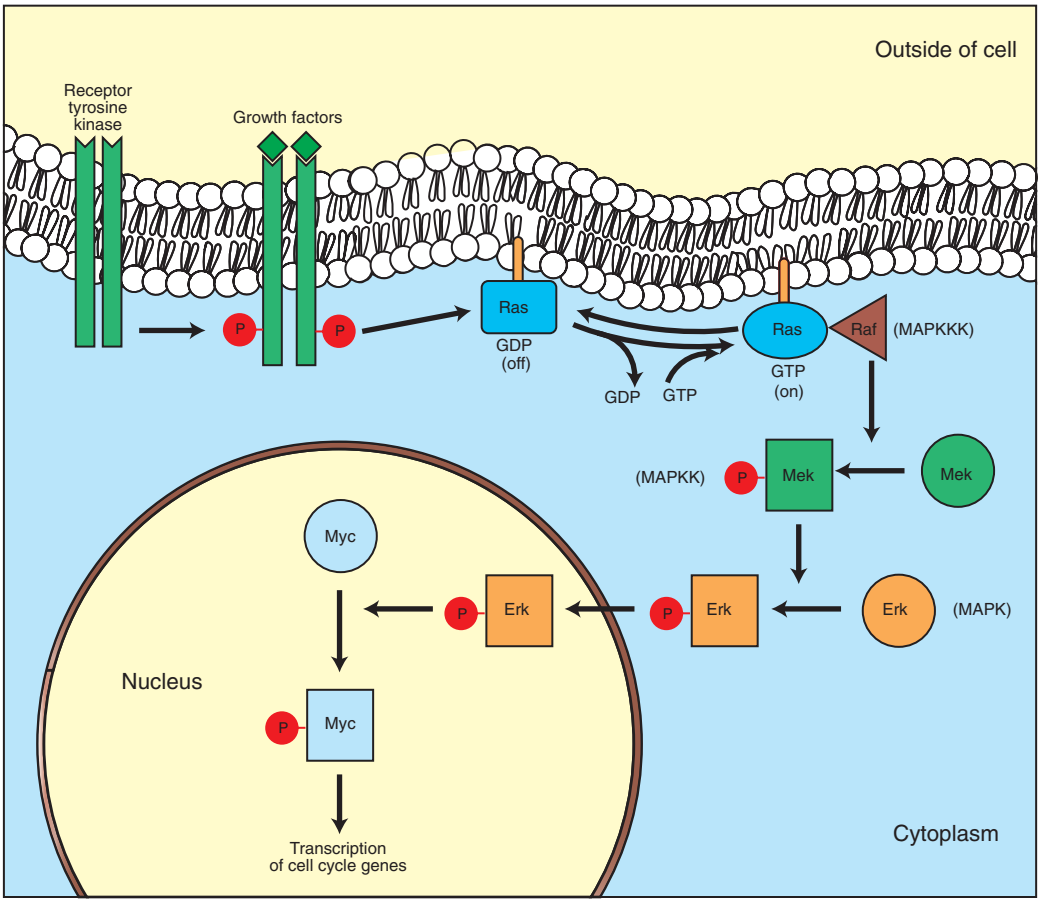
The Cell Cycle Is Stimulated by Growth Factors that Bind to and Activate Tyrosine Kinases

The growth factor/oncogene pathway begins with growth factors that function in a familiar way, as discussed in Chapter 1; they bind to and activate an integral membrane protein receptor. Indeed, growth factor receptors belong to the third large receptor

Vet-ebooks.com

family for environmental signals, the *receptor tyrosine kinase* family. This family of signal transducers has some similarities with the G-protein–coupled receptors (GPCRs), but also some important differences. Receptor tyrosine kinases (RTKs) do not require second messengers, but they do function through protein kinase activity (as many GPCRs do). The structure of RTKs is such that binding of ligand (a growth factor) by the extracellular portion of the receptor directly activates protein kinase activity by the cytoplasmic portion of the protein. The receptor itself is an enzyme (Fig. 2.5). Thus the RTK activates a cytoplasmic signal without the need for a second message. RTKs specifically add a phosphate group to a tyrosine residue of the substrate protein. This differs from the protein kinases discussed in Chapter 1 (PKA and PKC), which add the phosphate to serine or threonine residues. Phosphorylation of tyrosine residues within a protein is largely (but not exclusively) specialized to control cell growth pathways, and therefore tyrosine kinase activity generally is associated with stimulation of proliferation.

The growth factors that bind to the RTKs are too diverse to be discussed at length in this chapter. Rather, one important similarity for introductory professional students is that these factors are all poorly named, so do not judge the factor by its name. Sometimes growth factors have "growth factor" in their name; some are referred to as *cytokines,* and some are called *colony-stimulating factors* (for growth of colonies in soft agar, as previously mentioned). Further confusion arises because their names always reflect their history but rarely their broader function. Thus "epidermal growth factor" stimulates cell division in many more types of cells than only skin cells, but it was discovered using skin cells. The other, more important similarity among growth factors is that, whatever their name, they share a conserved basic pathway and "strategy" for control, as with the numerous ligands binding GPCRs and nuclear receptors, of their downstream effectors, in this case the CDK engines of the cell cycle. Growth factor activation of RTKs stimulates a pathway involving a G-protein "on-off" molecular switch, the *Ras protein* introduced in Chapter 1, and uses a cascade of protein kinases, both tyrosine and serine-threonine, called the *MAP kinase pathway.* Ultimately, the MAP kinase pathway activates transcription factors, in turn controlling the expression of cyclins, and other direct regulators of CDKs (see Fig. 2.5).

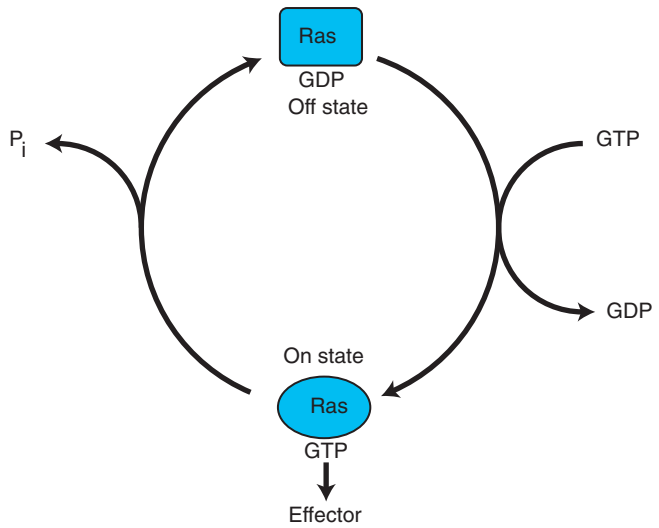


• **Fig. 2.5** Growth factor/oncogene pathway. This diagram shows the normal stimulatory pathway by which growth factors lead to cell division. Growth factors bind to membrane receptors (receptor tyrosine kinases, RTKs) that are themselves protein kinases. As shown here, after activation by binding a growth factor, the first protein to be phosphorylated at tyrosine residues is the receptor protein itself. This in turn causes a small G protein, Ras, to exchange GDP for GTP and thus be "turned on." The activated Ras then activates the first protein kinase in a conserved pathway of three kinases, called the MAP kinase pathway. For more detail on Ras and the MAP kinase pathway, see the text. Finally, this series of activating phosphorylations leads to the activation of transcription factors, such as Myc, in turn leading to the expression of genes directly involved in driving the cell cycle (e.g., expression of cyclin D). In this pathway, gain-of-function mutations of the RTKs, Ras, and Myc are particularly important in human cancers. *GDP,* Guanosine diphosphate; *GTP,* guanosine triphosphate.

Vet-ebooks.com

The *ras* Oncogene Contributes to Many Cancers and Serves As a Model for Understanding Small G Proteins

After activation of the RTK, the next major step in the growth factor/oncogene pathway in normal cells is activation of the protein product of the *ras* proto-oncogene. Investigations of how it worked revealed that the Ras protein was an important member of the small G-protein family of molecular regulators, all of which have intrinsic guanosinetriphosphatase (GTPase) activity and serve as molecular "on-off switches." These proteins control many basic cellular functions, and the heterotrimeric G protein evolved from Ras-like ancestor proteins (see Chapter 1). Indeed, in yeast it is Ras, not a heterotrimeric G protein, that controls adenyl cyclase and phospholipase C (see Fig. 1.16). Fig. 2.6 illustrates the duty cycle of this on-off switch and its basic similarity to the alpha subunit (G_α) of the heterotrimeric G proteins. Ras, other small G proteins, and G_α all are in the "on" state when they have guanosine triphosphate (GTP) bound to them (because of receptor activation). All are in the "off" state when the G protein hydrolyzes its GTP so that guanosine diphosphate (GDP) is now bound. You can see how this gene could be discovered as an oncogene, that is, a gene in which a gain-of-function mutation contributes to the development of cancer. If the GTPase activity is lost by mutation, this simple, enzymatic on-off switch remains trapped in the "on" position (the accelerator pedal is stuck). It continues to send an activating signal to the downstream cell cycle machinery without the presence of growth factors or the activation of RTKs. In fact, such mutations in Ras underlie its oncogenic function, and it is estimated that 30% of human cancers have activating mutations in their *ras* gene.



• **Fig. 2.6** Duty cycle of the Ras molecular "on-off switch." Ras serves as a model for the activity of small G proteins, of which there are hundreds in the cell. The molecular mechanism of Ras is similar to the alpha subunit of the heterotrimeric G protein, discussed in Chapter 1 and which evolved from Ras-like proteins. As shown here, Ras is in the "off" state when bound to GDP. Activation of receptor tyrosine kinases leads to nucleotide exchange: GDP is lost and GTP is bound. In the GTP-bound form, Ras is in the "on" state and sends a stimulatory signal downstream, in this case to Raf in the MAP kinase pathway (see Fig. 2.4). Normally, Ras rapidly returns to the off state because an intrinsic GTPase activity of the Ras protein hydrolyzes the GTP to GDP. This nucleotide-dependent on-off cycle is characteristic of all normal small G proteins. *GDP,* Guanosine diphosphate; *GTP,* guanosine triphosphate.

Other small G proteins control a myriad of cellular functions, including others involved in cancer. Thus the Rho subfamily of small G proteins is directly involved in the spread of cancer because it helps regulate actin assembly and activity. As described later, the spread of cancer depends on the ability of cells to migrate through tissues. This "crawling" motility in turn depends on a musclelike mechanism based on actin and myosin (see Fig. 1.4). Although the basic on-off activity of Ras and Rho are the same as that shown in Fig. 2.6, Rho is connected to actin, whereas active Ras activates the elements of the MAP kinase pathway.

The MAP Kinase Pathway Leads to the Expression of Cyclins and Other Stimulators of the Cell Cycle

GTP-bound Ras causes the sequential activation of a series of protein kinases, called Raf, Mek, and Erk. Raf phosphorylates and activates Mek, which in turn phosphorylates and activates Erk, as shown in Fig. 2.5. This trio of kinases is called *the mitogen-activated protein kinase,* or MAP kinase, pathway (a *mitogen* is a stimulator of mitosis, e.g., a growth factor). If any of these three protein kinases should experience a gain-of-function mutation irreversibly activating the protein kinase, a stimulatory signal is sent down the remainder of the pathway. Thus, as with *ras,* these three kinase genes act as oncogenes.

One important example of a gain-of-function mutation among the three MAP kinases involves the first of these MAP kinases, Raf. A single–amino-acid mutation in the kinase domain of Raf (a substitution of glutamate for normal valine at amino acid 600) causes permanent activation of Raf in approximately 50% of human melanomas, a very deadly cancer, and is also common in thyroid cancers. As described for mutations in Ras, activation of Raf sends an unregulated stimulatory signal downstream to the other MAP kinases, leading to unregulated proliferation of the cancer cells. Recent clinical progress involving melanoma illustrates the importance of the Raf, Mek, Erk pathway and of understanding which particular mutations are involved in a given patient's cancer. Two recently developed drugs, vemurafenib and dabrafenib, target the mutant Raf and significantly prolongs the life span of those melanoma patients harboring this *raf* mutation but has no effect in cases of melanoma with normal Raf/*raf.* However, melanoma patients typically develop resistance to the two Raf-directed drugs and so these drugs are now used in combination with drugs that specifically inhibit Mek. This combination therapy strengthens the blockade of the Raf, Mek, Erk pathway and leads to better outcomes.

Raf, Mek, and Erk are a specific example of yet another conserved but diverse general module of information transduction. There are MAP kinase trios other than Raf, Mek, and Erk. Although it is not worthwhile to give specific names to all the various pathways, it should be noted that these trios have a systematic set of names for their elements. Raf is a MAP kinase, kinase, kinase (a MAPKKK). Mek is a MAP kinase, kinase (MAPKK), and Erk protein is the MAP kinase (MAPK) itself. This jargon is awkward, but it is widely used and logical, as Fig. 2.5 suggests.

When activated, Erk activates one or more transcription factors that control the transcription and translation of a key regulator of the cyclin-CDK engine. One of these transcription factors, Myc ("mick"), is encoded by another important oncogene/proto-oncogene. As with *ras,* the *myc* gene is mutated in a high frequency of human tumors, giving rise to an oncogenic form able to activate

Vet-ebooks.com

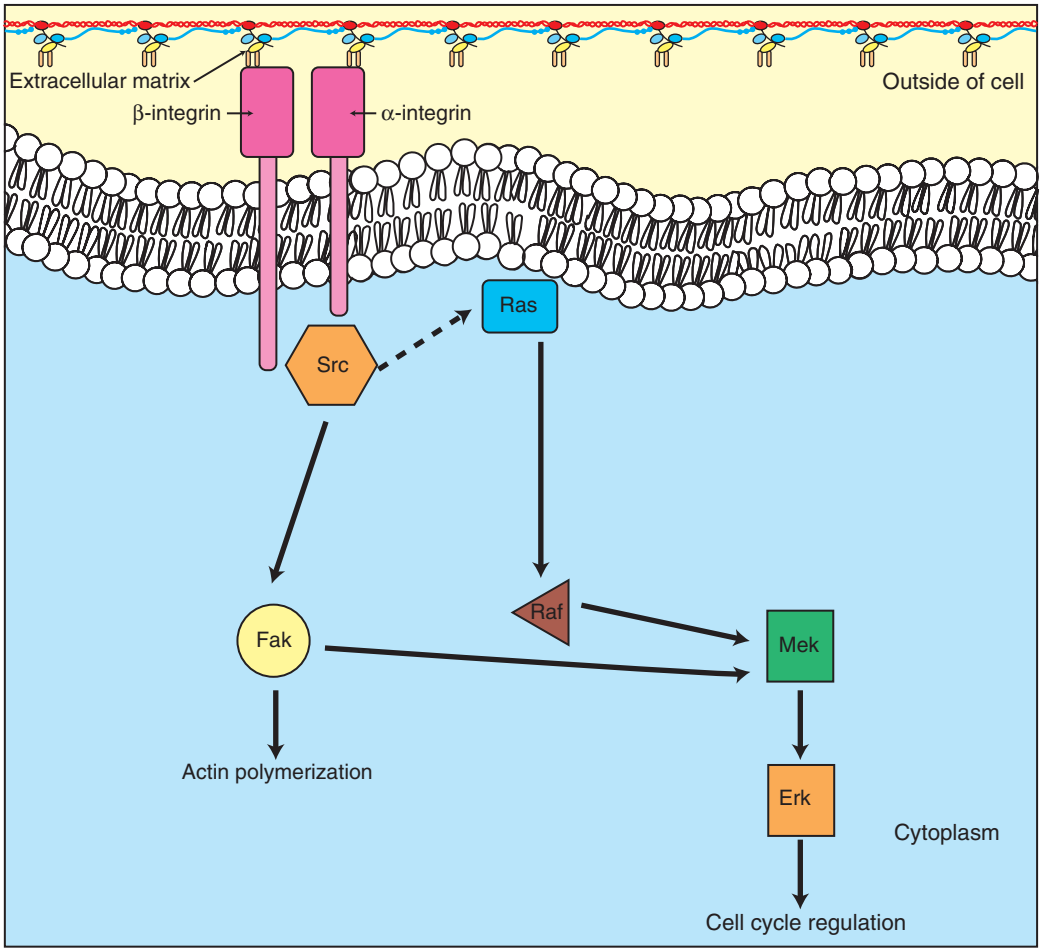
the cell cycle. As shown in Fig. 2.5, Myc protein is involved in the transcription of a variety of cyclins and of the CDK2 catalytic subunit and plays a significant role in allowing the cell to pass from G1 to S phase. Myc is also involved in many other transcription events related to cell growth, differentiation, and cancer.

This completes the growth stimulatory pathway beginning with a growth factor binding to its RTK receptor that, through Ras, a MAP kinase cascade, and a transcription factor, eventually leads to a direct "throttling up" of a cyclin-CDK engine. This same pathway is used similarly to transduce the information about the other major stimulator of cell division, cell attachment.

The MAP Kinase Pathway Also Mediates the Stimulation of the Cell Cycle by Cell Adhesion

As noted earlier, the other major throttle mechanism to regulate the cyclin-CDK engines of the cell cycle is cell adhesion. Cell adhesion, as with growth factor stimulation, ultimately stimulates cyclin-CDK pairs through the MAP kinase pathway. Two types of cell contact are involved in normal growth and proliferation. The most obvious is cell-cell adhesion; most cells are tightly attached to their neighboring cells. The second type is cell adhesion to an extracellular matrix (ECM) of fibrous proteins. Eighty percent of human and mouse cancers arise from epithelial cells (carcinomas), and all epithelial layers are attached to an ECM. The adhesion proteins that bind to other cells or to the ECM are *adhesion receptors.* Adhesion receptors are responsible for the mechanical aspect of attachment but also act similar to other receptors in transducing information across the plasma membrane. In this case adhesion receptors communicate the information that the cell is anchored and can divide.

Both cell-cell and cell-ECM adhesion activate the MAP kinase pathway, similar to growth factors, but the Ras intermediate is less important here. Fig. 2.7 shows the activation of the MAP kinase pathway as a result of cell-ECM adhesion. The adhesion receptors that bind to ECM are called *integrins* and these activate the MAP kinase pathway via two important intermediates that are oncogenes. One is Src ("sark"), a protein tyrosine kinase and the first oncogene (*src*) to be discovered. Unlike the RTKs previously described, Src is not a receptor. However, Src is located on the inside face of the plasma membrane, where it can interact with adhesion receptors. Another important intermediate is also a protein tyrosine kinase, called Fak (focal adhesion kinase). As before, activation of Src and Fak activate the MAP kinase pathway, leading to increased cell



• **Fig. 2.7** Cell adhesion functions through the MAP kinase pathway to stimulate cell division. In addition to the growth factor stimulation of proliferation shown in Fig. 2.5, normal epithelial cells also require stimulation of the MAP kinase pathway through adhesion to the extracellular matrix. The adhesion receptors are integral membrane proteins called integrins, which are activated by binding proteins of the extracellular matrix. Activation of integrins leads to activation of two protein kinases, Src and focal adhesion kinase (Fak), which in turn activate the MAP kinase pathway.

Vet-ebooks.com

division. Again, mutation or overexpression of *src* and *fak* sends inappropriate stimulation to the cell cycle machinery, which facilitates cancer. As mutant oncogenes, *fak* is associated with aggressive melanomas in humans. The *src* oncogene was named because of its ability to cause sarcomas in chickens.

Several other growth stimulatory pathways work in much the same manner as the growth factor and adhesion pathways. Most stimulatory pathways involve protein kinases and G proteins controlling the transcription of genes encoding proteins that are part of or close to the workings of the cyclin-CDK engines.

Having introduced the fundamentals of stimulatory pathways in the cell cycle, we now change our focus to consider the equally Rube Goldberg–like pathways that provide the brakes to the cell cycle.

Tumor Suppressors: Inhibitors of Cell Cycle

Checkpoints in the Cell Cycle Are Manned by Tumor Suppressors

The cell cycle machinery also has crucial "brake" mechanisms that function as checkpoints, as noted earlier. The existence of brake and checkpoint mechanisms operating in the cell cycle were discovered by fusing a normal cell with a cancer cell of the same type to form a hybrid cell with two nuclei. The resulting hybrid cell invariably showed normal regulation of growth. Apparently, a normal copy of some gene or genes present in the normal cell was able to suppress the altered activity of a mutant gene in the cancer cell. Thus these genes and their encoded proteins were called *tumor suppressors.*

Tumor suppressors play several different functional roles in braking and checking, and they can be divided into two broad types: gatekeepers and caretakers. *Gatekeepers* are genes and proteins that are involved in the actual checkpoint machinery connecting cell damage with a halt in the cell cycle. Thus *P53* (protein of 53-kilodalton mass) is a gatekeeper importantly involved in the pathway that detects DNA damage; it causes a halt in the cell cycle and, if the damage cannot be repaired, signals the cell to undergo programmed death. It is thought that about 50% of human cancers have a mutation in P53. *Caretakers* are usually proteins involved in the repair of damage or the normal maintenance of proteins crucial in the cell cycle. A human example of a caretaker gene and protein is Brca1 (breast cancer 1). This protein is normally involved in the repair of nucleotide mismatches (e.g., G paired with T rather than with C in the complementary DNA strand), and its mutant allele has been found to underlie familial (hereditary) breast cancer in some families.

With these normal functions, one can see how these genes and proteins would suppress tumor activity and cell proliferation. If they are working, DNA is repaired before the cell attempts to divide; this would tend to prevent mutation or other types of genetic instability. However, *loss-of-function mutation* in these genes means the cell now has lost the ability to detect or repair DNA damage. For example, when P53 is nonfunctional, even a badly damaged cell may not receive an adequate signal to commit suicide, and this already-mutant cell can continue to divide. Thus tumor suppressor genes are associated with loss-of-function mutations in cancer, not gain-of-function mutations as for oncogenes. Returning to the automobile analogy of brakes, mutant tumor suppressor genes resemble dysfunctional braking systems or no brakes at all.

We focus on two gatekeeper-type tumor suppressors because their role and importance in cancer are clear. The role of caretakers such as Brca1 is both more complex and more uncertain (see suggested reading on *brca* in the Bibliography).

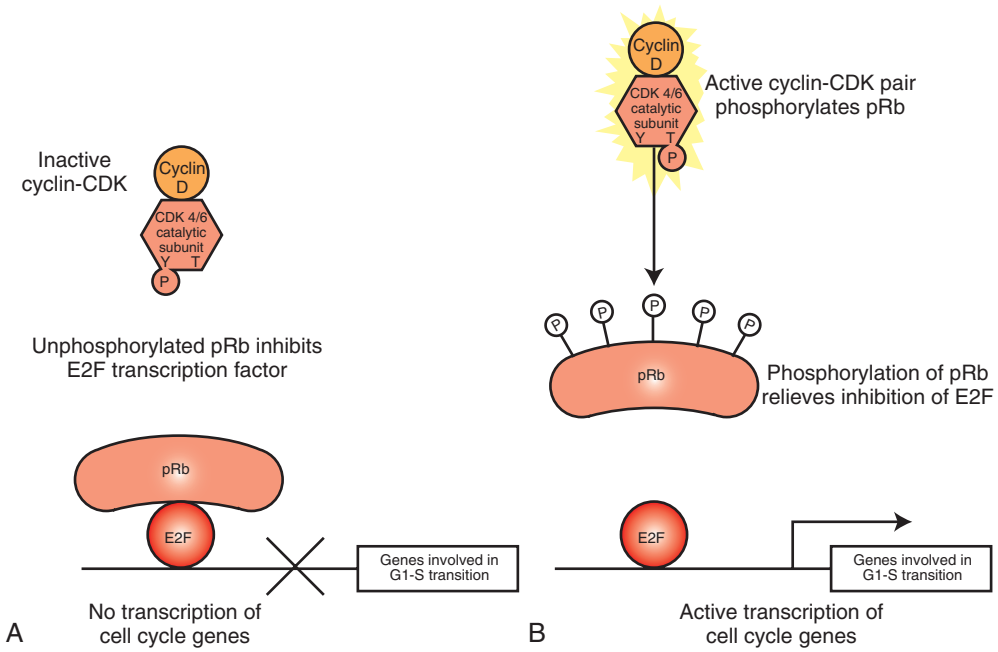
The Retinoblastoma and P53 Proteins Are the Main Gatekeepers for the Cell Cycle

Retinoblastoma is a rare, hereditary, childhood cancer of the retina of the eye. Despite its rarity and the fact that it cannot be induced in mice, retinoblastoma has played an important role in the study of cancer. A statistical study of the disease in the early 1970s provided the best evidence then available that human cancer is a genetic disease. Alfred Knudsen showed that children with retinoblastoma typically inherit one mutant copy from a parent (a *germ line mutation*), but then require a second *somatic mutation* in cells giving rise to the retina. Knudsen's *two-hit hypothesis* was a forerunner to the idea that cancer develops by the accumulation of mutations in a cell lineage. (Retinoblastoma tumors do require the accumulation of additional mutations beyond the two retinoblastoma genes being mutant.) Subsequently, the retinoblastoma gene, *rb,* was the first tumor suppressor gene to be cloned. Study of the encoded protein, pRb, showed that it played a central role in controlling the transition from G1 to S phase of the cell cycle.

The *retinoblastoma protein* is a repressor of a transcription factor whose activity is required for the cell to enter S phase from G1 (Fig. 2.8). The transcription factor is E2F, which controls the expression of a wide variety of genes/proteins required for DNA synthesis, including cyclin A, CDK1 (see Fig. 2.4), and subunits of DNA polymerase. The retinoblastoma protein is a potent inhibitor of E2F only when it is bound to E2F directly, which requires pRb to be in an unphosphorylated state. The repression of E2F is released by phosphorylation of pRb by cyclin-CDK pairs operating early in G1 in the cell cycle. As discussed, growth factor stimulation of the MAP kinase pathway leads to expression of cyclin D (see Fig. 2.5), which in turn makes a pair with either CDK4 or CDK6 to make an active CDK. One of the substrates for cyclin D/CDK4, 6 is the retinoblastoma protein. When pRb is phosphorylated by CDK4, 6, it releases from E2F, allowing this transcription factor to promote RNA polymerase activity on genes with E2F promoter regions (see Fig. 2.8). It is this release of inhibition by CDK-mediated phosphorylation of pRb that constitutes the molecular mechanism underlying the R-point "decision" to divide late in G1 mentioned earlier and shown in Fig. 2.2. If both copies of *rb* are mutant, as in retinoblastoma, there will be no active repressor molecules to bind to E2F, and the decision will always be to divide, regardless of other conditions. E2F then promotes uncontrolled expression of S-phase genes whether or not CDK4, 6 has been activated (in part) by growth factors and adhesion, thus making a contribution to unregulated growth and to cancer. Conversely, in its normal, nonmutant form, pRb tends to suppress tumor formation by acting as a gatekeeper, only allowing the cell "to cross the border" between G1 and into S phase if normal growth factor and adhesion signals are received. Thus pRb plays a crucial gatekeeper role in healthy, normal cell cycle control.

The other crucial gatekeeper between G1 and S phase is P53. Unlike pRb, P53 does not participate in healthy cell cycles; P53 is only active in response to cell damage, usually DNA damage, or stress, such as low O_2 concentration or oncogene activation (Fig. 2.9). The role of P53 is to ensure that stressed/damaged cells are either repaired or, if not, commit suicide before being allowed to replicate their DNA. As a gatekeeper, P53's mechanism is also more direct than pRb; P53 is a transcription factor, and

Vet-ebooks.com



• **Fig. 2.8** Retinoblastoma protein and the G1-to-S transition. (A) In quiescent cells or cells early in G1, retinoblastoma protein (pRb) exists in a nonphosphorylated state that is a direct inhibitor of the E2F transcription factor. The principal CDK pair of G1, cyclin D with CDK4 or CDK6, phosphorylates pRb, releasing its inhibition of E2F. (B) Activated E2F then participates in the expression of a variety of genes required for S phase, including the cyclins and CDKs of S phase and subunits of DNA polymerase.

P53 activation stimulates the expression of a protein that is a powerful general inhibitor of all the cyclin/CDK engines. As a transcription factor, P53 also mediates the expression of genes that encode stimulators of cell death, as discussed shortly. Whether the cell responds to P53 by cell cycle arrest to allow repair, or by committing suicide, depends on multiple factors, but presence of an oncogene is among the most important. Normally, the cell cycle arrest activity of P53 is dominant to its death-inducing activity. However, in the presence of oncogenes, including *myc,* suicide is favored. This illustrates clearly the normal tumor suppressor activity of P53; although a cell expressing an oncogene will tend toward increased proliferation, the same oncogene, acting through P53, activates a death pathway to prevent expansion of the mutant cell population.

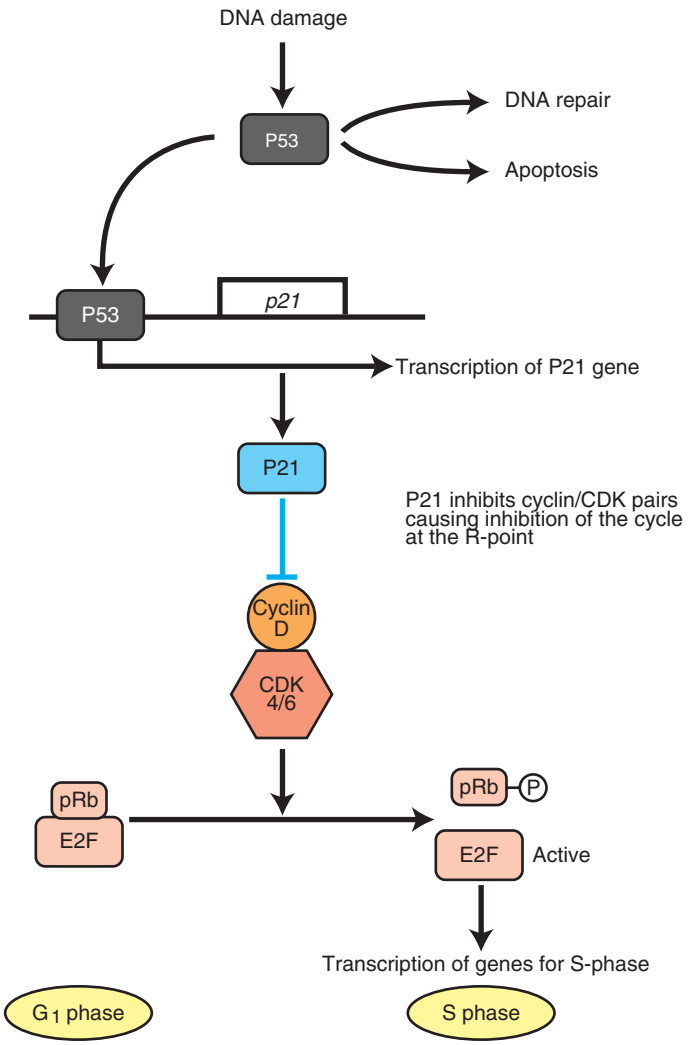
The activation of P53 occurs in part through mechanisms familiar from previous examples of protein control, including phosphorylation and binding with other proteins. In addition, P53 activity is also regulated simply by an increase in its concentration within the cell. That is, P53 is normally synthesized at a steady but slow rate throughout the cell cycle and is normally degraded at a similar rate. In healthy cells, the half-life for a P53 molecule is about 30 minutes, but this increases threefold to sevenfold in response to DNA damage. Even one double-strand break in DNA has been shown to increase P53 concentration rapidly in some cells. Again, it is clear how P53 serves as both a gatekeeper and a tumor suppressor. Activated P53 prevents a cell with DNA damage from crossing the G1-S boundary (its gatekeeper function), which in turn prevents mutant cells from being allowed to accumulate additional mutations (its tumor suppressor function).

However, if the *p53* gene suffers a loss-of-function mutation and the protein cannot act as a transcription factor, a damaged cell will be able to divide, increasing the probability of accumulating further damage and leading to possible cancer. Thus *p53*/P53 is one of the most important single genes and proteins involved in human cancers; in 1993, the journal *Science* even named it "Molecule of the Year." About 50% of human tumors have a mutation in *p53,* with most of these eliminating DNA binding, disabling its transcription factor activity. When the *p53* gene was "knocked out" in mice, 74% of the animals developed cancers by 6 months of age (young adult). Among experimental mice that had one or two normal copies of the gene, only 1% of animals developed a tumor by 9 months.

In addition to a checkpoint for S phase in which DNA damage provides an important regulatory signal, the other major checkpoint occurs during mitosis. This checkpoint responds to mitotic spindle abnormalities or damage and to abnormalities in the array of chromosomes within the spindle. Again, one can easily see how mutations that disrupted such "safety interlocks" could lead to further damage, by segregating both replicated chromosomes into one daughter cell, for example, with no copy of that chromosome in the other daughter cell. This would lead directly to aneuploidy. Among human cancers, colon cancer is frequently found to have mutations in mitotic checkpoint genes.

However, we leave the topic of mitotic checkpoints at this somewhat intuitive level and do not address the molecular mechanisms. Such an effort would require a lengthy background discussion of the structure, functions, and control of the microtubule-based mitotic spindle, more suitable for a course in cell biology than animal physiology. Instead, we now discuss the controls on cell growth other than proliferation and briefly summarize what is known about programmed cell death and the control of cell life span.

Vet-ebooks.com



• **Fig. 2.9** P53 and the response to DNA damage. Normally, P53 is maintained at low levels in the cell by continuous synthesis and breakdown. DNA damage inhibits breakdown, allowing P53 to build up to functional levels. P53 is itself a transcription factor, and its targets include *p21,* whose protein is a potent inhibitor of all cyclin-CDK pairs. Thus upregulation of P53 brings the cell cycle to a halt, typically by inhibiting phosphorylation of pRb, as shown here. Subsequently, if the DNA is repaired, P53 returns to low concentration. If the DNA remains damaged, P53 leads to an apoptotic response by mediating expression of proapoptotic proteins, as described in the text. *CDK,* Cyclin-dependent kinase.

Mechanisms Regulating Cell Suicide and Cell Life Span

Apoptosis Is the Process of Cell Suicide

The process of cell death by external damage, involving cellular swelling, bursting, and engagement of the inflammatory response, has been well described for more than 100 years. This form of cell death is called *necrosis* and is familiar from experiences as common as a cut or abrasion. A rather different process of cell death was described in the 1970s in which cells shrink, the DNA fragments in a systematic way, the plasma membrane bubbles and churns, and the cell breaks up into small pieces that are rapidly engulfed by neighboring cells (Fig. 2.10). This neater and cleaner form of cell death was named *apoptosis* (a-pah-toe-sis; Greek, "falling off"). Apoptosis was largely ignored for the next 20 years, until studies of nematode development discovered genes whose only role was to control apoptosis. Further studies revealed the highly conserved mechanisms of apoptosis and its importance in normal development, immune function, and disease. Resistance to apoptosis is clearly a major contributor to cancer. (Conversely, too much apoptosis plays an important role in neurodegenerative diseases and stroke.) Particularly relevant to clinical practice, most cancer drugs and radiation therapy kill the target cells (and unfortunately many bystander cells) by stimulating apoptosis. Also relevant to clinical practice, recently discovered cancer therapies augment the immune system's ability to trigger apoptosis in cancer cells, as discussed shortly.

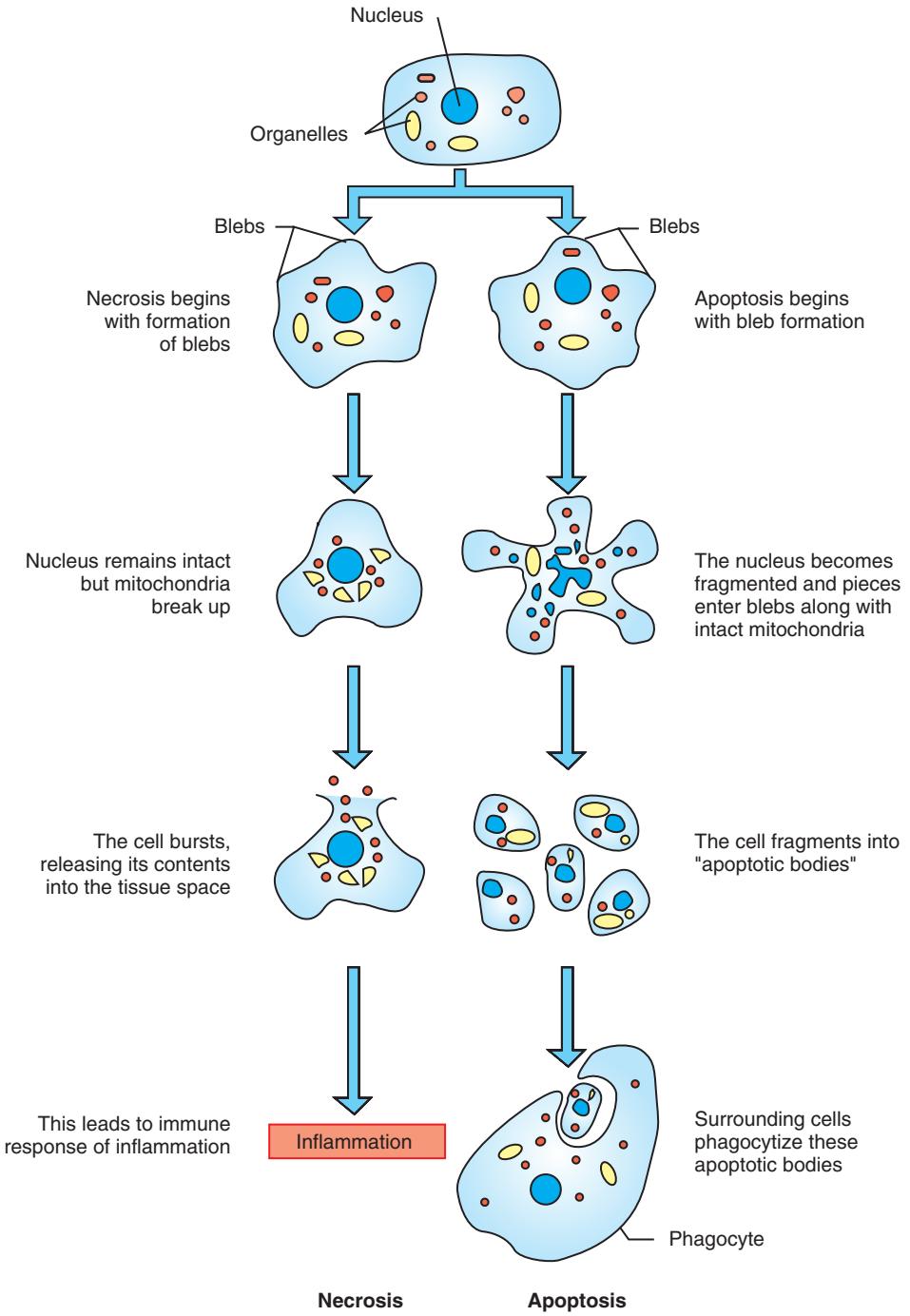
There are two broad pathways that lead to apoptosis. The *intrinsic pathway* of apoptosis responds to internal damage or stress from within the cell. The *extrinsic pathway* begins with a signal molecule binding to a "death receptor" on the cell surface (Fig. 2.11). However, both pathways converge on the same "executioners." *Caspases* are a family of proteolytic enzymes that have a cysteine amino acid at their active site (the "c" in caspase) and that cleave the substrate proteins at an aspartate amino acid (the "asp" in caspase). Similar to many other proteases, including digestive enzymes and blood-clotting factors, caspases are themselves activated by proteolytic cleavage. That is, as initially translated, the protease contains an inhibitory peptide that must be cleaved away to allow active proteolysis by the enzyme. In the case of the caspases, the activating protease is itself another caspase. Thus caspases are divided into *activating caspases,* which respond directly to one or another element in the intrinsic or extrinsic pathway, and downstream *executioner caspases,* which lead to specific cleavage of cellular structures. Among other tasks, executioner caspases cleave cytoskeletal proteins, leading to cell shrinkage, and activate the DNA-degrading enzymes involved in the systematic fragmentation of DNA.

The basic extrinsic pathway of apoptosis, also called the *death receptor pathway,* is unusually short and straightforward considering the extreme and irreversible outcome. An extracellular signal, which can be either soluble or attached to the surface of another cell, binds to and activates a death receptor on the cell destined to commit suicide. The cytoplasmic domain of the death receptor recruits one or two adapter proteins that directly activate an activating caspase, which in turn activates one or more executioner caspases (see Fig. 2.11). The activating caspase of the extrinsic pathway can also engage in "cross-talk" with the intrinsic pathway, described shortly, to increase the extent of caspase activation. The extrinsic pathway plays a crucial role in regulating the immune system, where the vast majority of immune cells initially generated are eliminated. The extrinsic pathway in cancer is also important in that many cancer cells are able to inhibit the immune system's capacity to kill cancer cells via the extrinsic pathway. Restoration of immune attack is among the most promising new clinical weapons against certain cancers, as discussed shortly.

Resistance to Apoptosis via the Intrinsic Pathway Is a Hallmark of Cancer

Internal cellular damage or stress, including DNA damage, absence of cell anchorage, too little or too much oxygen metabolism, oncogene activation, and radiation damage, can stimulate the intrinsic pathway of apoptosis in normal cells. Most, and perhaps all, cancer cells are more resistant than normal cells to apoptosis

Vet-ebooks.com

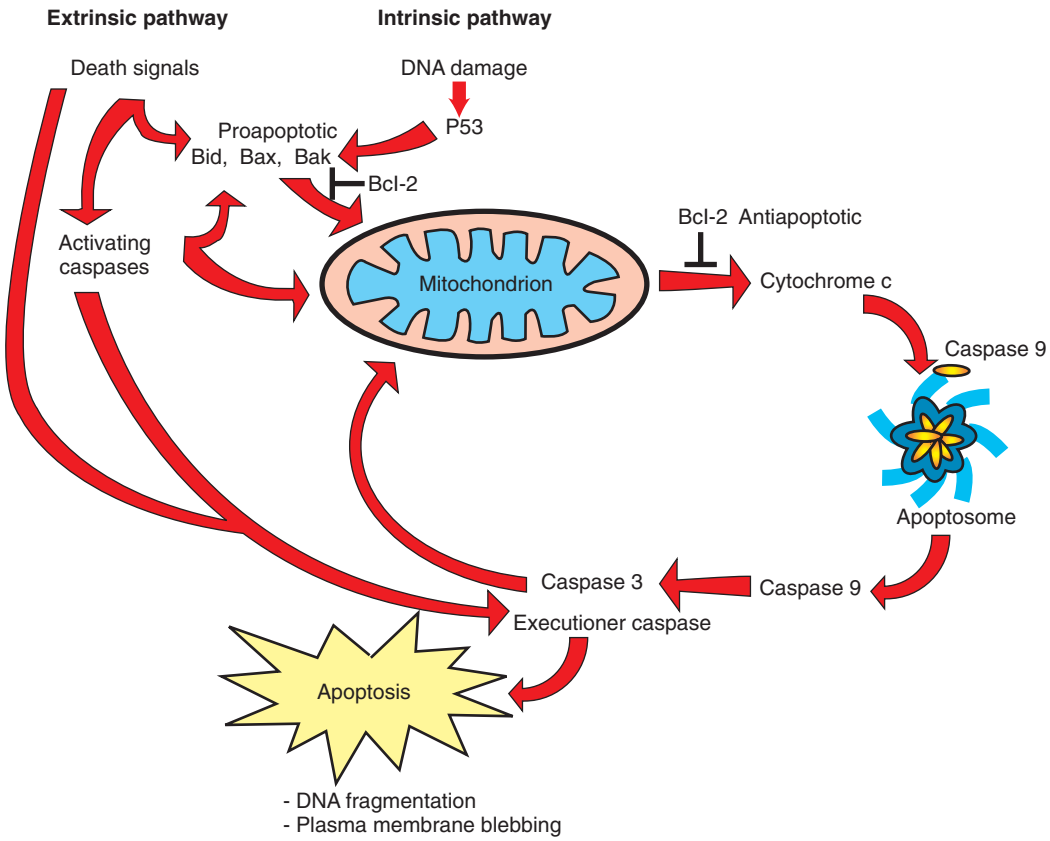


• **Fig. 2.10** Necrosis versus apoptosis. Necrosis is cell death as a result of external damage to the cell that leads to bursting of the cell and release of cell contents, leading to inflammation. Apoptosis is cell death as a result of intrinsic mechanisms in which the cell is broken down into cell fragments that then undergo phagocytosis by neighboring cells. This produces no inflammatory reaction and is so "tidy" that apoptosis is difficult to observe.

through this pathway. Resistance to apoptosis not only increases the probability that the cell will be able to accumulate further genetic damage, but also reduces the likelihood that cancer cells can be eliminated. This is because the antitumor activity of the immune system, as well as most chemotherapy and radiation treatments, depends on apoptosis. Thus resistance to apoptosis often means resistance to treatment.

The intrinsic pathway is considerably more complex than the extrinsic pathway, and this discussion focuses on three major elements of the pathway involved in activating caspases: P53, the mitochondrion, and the *Bcl family* of proteins (see Fig. 2.11). This family of proteins was originally discovered in a cancer ("Bcl" is from B-cell lymphoma, a type of leukemia in which the first such protein was discovered) and includes both proapoptotic and

Vet-ebooks.com



• **Fig. 2.11** Extrinsic and intrinsic pathways for apoptosis. See text for details.

antiapoptotic members. The balance between pro- and antiapoptotic members determines whether the cell lives or dies. The resistance of cancer cells to apoptosis arises not only from mutations, such as those already described for *p53,* but also from underexpression of proapoptotic mediators and overexpression of antiapoptotic proteins.

We begin with the *mitochondrion,* familiar as the "powerhouse" of the cell responsible for generating ATP, but also the central control point for the intrinsic pathway of apoptosis. Recall that the mitochondrion has both an inner membrane, responsible for electron transport, and an outer membrane, responsible for compartmentation of this organelle. Proapoptotic signals cause the outer membrane of the mitochondria to become leaky, releasing several proapoptotic proteins not normally found in the cytoplasm. Among the most important is cytochrome c, an electron transport protein that is only loosely attached to the inner membrane. In the cytoplasm, cytochrome c stimulates the assembly of a multiprotein complex (the apoptosome) that directly stimulates the activity of an activating caspase (caspase-9), ultimately leading to the activation of executioner caspases. What then determines the extent of permeability (leakiness) of the mitochondrial outer membrane?

The Bcl family members are major regulators of mitochondrial outer membrane permeability. The proapoptotic members of this family, such as Bax, lead to permeabilization by assembling to form channels in the outer membrane through which cytochrome c can pass. Proapoptotic members of the family can also cause the channel through which ATP normally passes into the cytoplasm to open wider than usual. The antiapoptotic members of the family, such as Bcl-2, seem to function by binding to proapoptotic members, inhibiting their activity. In a healthy cell, antiapoptotic Bcl members are at high enough concentration to neutralize proapoptotic activity. Damage increases the amount of proapoptotic Bcl molecules and leads to membrane permeabilization. Thus the balance between pro- and antiapoptotic members of the family controls the permeability state of mitochondria and the survival of the cell.

With about 20 different members of the Bcl family, the balance between pro- and antiapoptotic Bcl molecules has multiple controls, but P53 activity is certainly a major player. Recall that, when activated (e.g., by DNA damage), P53 acts as a transcription factor, and at least three different proapoptotic Bcl genes are transcriptionally activated by P53. These include Bax, and also the particularly powerful proapoptotic protein, PUMA. Downstream, P53 also activates the transcription of the activating caspase-9 gene, and the gene of a major cytoplasmic component of the apoptosome. In addition to acting as an activating transcription factor, P53 serves as an inhibitory transcription factor for some genes, including that of the antiapoptotic Bcl-2 protein. Finally and independent of transcription, activated P53 can directly activate Bax, which is required for its ability to assemble into channel structures. With these multiple effects on apoptotic genes and proteins, P53 is regarded as a central apoptotic control point, in addition to its role in cell cycle regulation.

As noted earlier in the discussion of P53, the importance of apoptosis to tumorigenesis is that, with normal apoptosis, almost all damaged cells are eliminated. Without apoptosis, damaged cells live to accumulate additional damage, which illustrates why multiple mutations and dysfunctions are required for tumors to reach a clinically significant stage. The resistance of cancer cells to apoptosis arises from many types of mutations and disruptions of normal gene expression. In some cases, mutation of the *p53* gene eliminates its DNA binding and thus transcriptional activity. Related to P53 activity is a protein that regulates P53's normal proteolytic

Vet-ebooks.com

breakdown (see previous discussion). Overexpression of this protein (MDM2) in various cancers of soft tissues inhibits the accumulation of P53 to active levels and therefore inhibits both cell cycle arrest and apoptosis. The antiapoptotic Bcl-2 protein is overexpressed in a variety of human cancers, including 60% of human follicular lymphomas, but also some lung cancers, melanoma, and prostate cancer. Another common apoptotic lesion seen in cancer cells is overexpression of proteins that bind to and directly inactivate caspases, as well as mutation or loss of expression of the caspases themselves.

Many Types of Cancer Cells Suppress Immune Attack and so Avoid Apoptosis via the Extrinsic Pathway

In contrast to the intrinsic pathway of apoptosis, cancer cells themselves are rarely resistant to the extrinsic pathway of apoptosis. Rather, many types of cancer cells avoid apoptosis via the extrinsic pathway by "jamming" the initiating death signal that would otherwise come from the immune system. The immune system is the "defense establishment" of the body that functions, in large part, by recognizing "self" from foreign, "nonself" cells. These foreign cells are then destroyed, often by apoptosis. The immune system is exceedingly complex; a primer on the immune system can be found in Chapters 54 and 55 of this text. Despite the topic being somewhat out of order, recent highly encouraging advances in cancer therapy mediated through the immune system make it imperative to say just a few words about these new therapies. Indeed, the Nobel Prize in Medicine or Physiology for 2018 was awarded for work, including the discovery of the PD-1 receptor discussed below, that led to this new type of therapy. It is likely that you will hear more about this in your immunology course.

Suffice it to say that cancer cells, although derived from the body's own cells, are sufficiently damaged that the immune system should, and often does, attack them as foreign, nonself cells. This is the same basis on which the immune system attacks virally infected cells (see Fig. 55.3). And the available evidence suggests that spontaneous remission of cancer is the result of the patient's immune system successfully destroying the cancer. One form of immune attack involves immune cells signaling the cancer cell to commit apoptosis via the extrinsic pathway. As with a military defense establishment, it is crucial that only nonself cells (bacteria, fungi, infected and damaged cells) are attacked and not the normal "self" cells, i.e., the latter must avoid "friendly fire" in current military jargon. Consequently, the attacking immune cells have molecular receptors for being activated, or "armed," as well as for being inhibited, or "standing down."

Among the latter receptors is a protein called PD-1 ("programmed death 1"); as the name implies, this receptor was discovered as a "death receptor" whose activation led to apoptosis via the extrinsic pathway. As with the limited accuracy of names of growth factors mentioned earlier, subsequent study revealed that PD-1 only rarely causes apoptosis in some normal immune cells. Rather, this receptor primarily serves as a downregulator of immune cell activation. For example, melanoma cells express high levels of a ligand that activates PD-1 and thus downregulates immune cells that would otherwise attack the melanoma. The therapeutic advance has been the development of drugs (antibodies in this case) that bind to and inactivate the PD-1 receptor, i.e., inhibiting the inhibitor. This enables the immune cells to attack the melanoma and cause remission. This therapy was used on former US President Jimmy Carter to successfully treat his advanced-stage melanoma. The same therapy is used to treat small-cell lung carcinoma and is showing promise for a variety of other solid tissue tumors including breast and renal tumors. More information about this new strategy of immune-directed cancer therapy is provided in a paper by Messerschmidt et al. in the Bibliography for this chapter.

Cellular Life Span Is Determined by DNA Sequences at the Ends of Chromosomes

The final major dysfunction of growth control found within cancer cells is the most recently discovered but also seems to be the most common single molecular lesion in cancers: the expression of a reverse transcriptase called *telomerase.* (A reverse transcriptase is any enzyme that synthesizes DNA from an RNA template.) Telomerase is responsible for elongating *telomeres,* the specialized, noncoding regions of DNA found at the end of chromosomes. However, telomerase is normally expressed only in embryonic cells and in adult *stem cells.* (Stem cells are specialized normal cells that do have limitless replicative potential, such as gamete-generating cells and the blood-forming cells of the bone marrow, as discussed later.) The vast majority of normal somatic cells do not express telomerase, but it is expressed in 85% to 90% of all cancers and is the major determinant of the "immortality" of cancer cells.

Telomeres are segments of highly repetitive DNA, representing hundreds of repeats of the simple nucleotide sequence TTAGGG (in vertebrates), found at the ends of chromosomes. Telomeres serve as caps at chromosomal ends, protecting them against end-to-end joining of chromosomes. Telomeres also prevent the ends of chromosomes from being recognized as sites of DNA damage (double-strand DNA breaks). Crucially, telomeres also protect against the loss of coding DNA from each chromosomal end with every round of DNA replication; this is needed because normal DNA polymerases have a serious limitation; they cannot fully replicate the end of a double-strand DNA molecule. As a result, the ends of chromosomes become shorter with each round of DNA replication. (Bacteria solve this problem by having circular DNA chromosomes.)

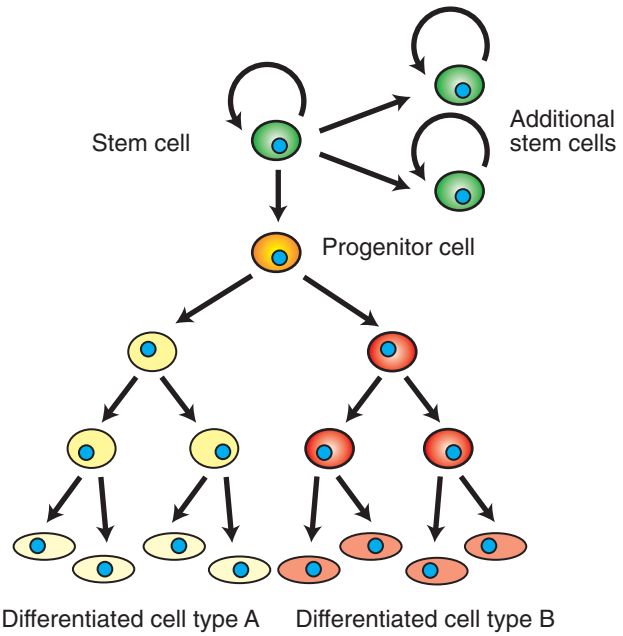
Telomeres are expendable DNA, at the ends of chromosomes, whose progressive shortening does not compromise the coding function of the genome. Although no coding sequence is lost, the shortening of telomeres nevertheless plays an important role in the cell. The shortening of telomeres serves as a kind of clock, measuring the number of times a cell has replicated its DNA, and the length of the telomere reflects the age of the cell. Through poorly understood mechanisms, cells can detect the length of their telomeres, and when they reach a critically short length, the cell ceases to divide and is said to undergo *senescence* (Latin; "growing old"). As noted earlier, normal cells have a finite life span, such that a cell taken from a middle-aged human will divide 20 to 40 times in culture before senescence. When placed into culture, the number of subsequent cell divisions before senescence reflects the original length of the telomeres. Further, various degenerative diseases, including cirrhosis of the liver, have been shown to accelerate telomere shortening. In principle, senescence is a powerful block to cancer because the original damaged cell (see Fig. 2.1) would be unable to divide for a sufficient number of generations to accumulate the necessary multiple mutations required to produce a tumor. Telomerase expression (and other, less common means of elongating telomeres) effectively eliminates this block to cancer development by causing the cells to become immortal.

Telomerase has both protein and RNA components. The protein provides the catalytic reverse transcriptase, allowing the enzyme

Vet-ebooks.com

to elongate the telomere sequence based on the RNA template it carries. That is, the RNA component of telomerase is complementary to the telomere DNA sequence and is used as the template for telomere DNA elongation. Telomerase is not expressed in normal adult somatic cells except for stem cells, as mentioned earlier. However, immortal tissue culture cells do express telomerase, as do cancer cells. Experimental expression of telomerase in human cells dramatically increases the replicative life span of the cells. Thus the observed expression of telomerase in the vast majority of human cancers permits these cells to divide indefinitely, providing yet another selective advantage for these cells to accumulate additional damage over time.

In the last sections of this chapter, we turn our attention to the cancer cell in the context of a *tumor,* which is a population of cancer cells interacting with one another and with surrounding normal tissue. We end our discussion of the intrinsic growth controls of normal and cancer cells with an experimental result that seems to confirm the importance of the types of damage discussed thus far. This experiment showed that four genetic changes were sufficient to transform normal human kidney cells into cancer cells able to form tumors when transplanted into a mouse host (with no immune system). The four genetic changes were to "engineer" into the cells an activating mutation for the *ras* oncogene, inactivation of both the retinoblastoma and P53 proteins, and expression of the catalytic subunit of telomerase. Thus damage to the genes or expression of these molecules, emphasized here, reflects the minimum requirements for a normal cell to grow as a cancer.



• **Fig. 2.12** Stem cells. Stem cells are self-renewing cells of high, sometimes unlimited, replicative potential. (The word "potential" is crucial; typically stem cells replicate only rarely. It is the progenitor cells that are the rapidly dividing population.) The proliferation forms both additional stem cells and progenitor cells. These progenitor cells divide, often rapidly as noted, and eventually differentiate to become one or more types of differentiated somatic cells specialized for certain tasks (e.g., erythrocytes and monocytes of blood).

Tumor Origin and the Spread of Cancer

Cancer Cells May Be Related to Stem Cells

As noted in the previous section, some normal adult cells do have unlimited replicative potential. These are *stem cells,* a cell type that has been in the news for decades and remains so (see K. Servick, *The Stem Cell Skeptic* in the chapter Bibliography). A *stem cell* is a self-renewing cell of high proliferative potential that can also give rise to differentiated cells. Typically, stem cell division produces one cell that remains a stem cell whereas the other daughter cell differentiates into a specialized cell with the usual limited life span (Fig. 2.12). The cell that continues being a stem cell does not lose any developmental capacity and can divide indefinitely, continuing to produce additional stem cells and additional differentiated cells.

Much of the recent attention in the news centers on *embryonic* stem cells. These are embryonic cells that can either continue to form stem cells or differentiate, in principle, to any and every cell type within the body. Even in the adult, however, the maintenance of many normal tissues is critically dependent on stem cells. *Adult* stem cells, however, can only differentiate into a limited array of different cell types, not every cell type in the body. Best understood is that all the various cells of the blood arise from the division of hematopoietic stem cells in the bone marrow; one daughter cell remains a stem cell in the bone marrow, whereas the other differentiates to become one of the several types of blood cells (but the blood stem cell can only form blood cells, not nonblood cells). The cells lining the gut and those giving rise to hair also arise from a stable population of adult stem cells, some of whose descendants differentiate into specialized gut or hair-producing cells. For this reason, chemotherapy that is intended to cause apoptosis in cancer cells typically also affects these same populations of normal stem cells; common side effects of chemotherapy include anemia, hair loss, and digestive dysfunction.

Cancer cells resemble stem cells in their immortality, but the relationship of cancer cells to stem cells may go further. Based on the presentation thus far, you may have the mental image of a tumor composed of a uniform population of badly damaged cells, any of which would be capable of forming a new tumor if transplanted. In fact, real tumors are not a homogeneous population of cells but rather are composed of a variety of cells that differ significantly in their phenotype, despite all being clonal descendants of a single somatic cell, as shown in Fig. 2.1. (Keep in mind that all somatic cells of the body are clonal descendants of the fertilized egg, so phenotypic differences arising within clonal lines is not surprising by itself.) Further, experiments with a variety of cancers show that only 1% or less of tumor cells are capable of forming another tumor, even in the same patient (or mouse). Thus tumors may contain a small subpopulation of *cancer stem cells* that are responsible for producing the heterogeneous cells in the tumor and are uniquely able to continue cancer growth. This would also give tumors the capacity to adapt to their surroundings; because stem cells can differentiate in various ways, differentiated cells that allowed continued growth and survival would be selected.

This hypothesis has been persuasively supported only in leukemias, but it may apply to other cancers as well. For leukemias, the cancer stem cells express some marker proteins characteristic of normal hematopoietic stem cells. Further, only those leukemia stem cells expressing certain normal markers are capable of forming new cancers when transplanted. Finally, a possible relationship between cancer and stem cells is that perhaps the genetic changes summarized in this chapter must occur in a normal adult stem cell to produce cancer cells. Here again, the best evidence in favor of such a mechanism comes from leukemias. But the blood is unusual in ways other than just being a fluid rather than a solid

Vet-ebooks.com

tissue, and it is not at all clear that other types of cancers will prove similar.

Indeed, recent results on melanoma, once again, challenge some of the concepts of the cancer stem cell hypothesis. For example, a relatively large fraction of melanoma cells, perhaps as much as 25% of the tumor cells, can produce tumors after transplantation. This high frequency is not consistent with scientists' conception of stem cells. More troubling still, some evidence suggests that the ability of melanoma cells to form new tumors is transient. It is as if the "stemness" of melanoma cells is unstable and comes and goes. So the stem cell hypothesis for cancer is a controversial one. On the one hand, those tumors that show stem cell properties raise the possibility that cancer therapy should, perhaps, be directed primarily at cancer stem cells and not the majority of cells in a tumor. And, it might be possible to use stem cell markers for drug targeting, thus sparing the vast majority of cells in the body from side effects of the treatment. On the other hand, we do not know how many cancers or individual tumors will or will not be shown to adhere to the stem cell model. Additional readings about the current state of the cancer stem cell hypothesis are provided in the Bibliography at the end of this chapter (by Kreso and Dick, and by van Niekerk et al.).

Death by Cancer Is Usually the Result of Its Spread, Not the Original Tumor

Death from cancer is often the result of the spread of the cancer from the initial tumor, the *primary tumor,* to various distant sites. This process of cancer cells colonizing other tissues is called *metastasis.* For some cancer types, including leukemias and those of the brain, the primary tumor itself can be fatal. In contrast, the primary tumor for melanoma is little more than a molelike dark patch on the skin that does not become life-threatening until these cancer cells spread. Although metastasis is the deadliest aspect of cancer, much less is known about it than about the dysfunctions of cell growth leading to the primary tumor.

The best understood aspect of metastasis is that it occurs by a multistep process called the *metastatic cascade.* In this step-by-step process, cells escape from the primary tumor, breaking through tissue barriers to gain access to the circulatory system. The cells are carried until they escape the circulatory system to invade a new tissue (Fig. 2.13). The steps of the metastatic cascade suggest that dysfunction of three broad types of cellular function is particularly important: cellular adhesion, cellular motility, and secretion of proteases. How these dysfunctions arise from the genetic damage of growth in the primary tumor is, again, unknown, but mutation resulting from the genetic instability of the primary tumor is typically suggested as a link.

The first step of the metastatic cascade is the loss of cell adhesion by the cancer cell, both to neighboring cells and to the ECM. Accordingly, many types of cancer cells show greatly reduced expression of a cell-cell adhesion receptor, E-cadherin, important for epithelial adhesion. Similarly, primary tumor cells show a wide variety of abnormalities in the number and type of cell-ECM adhesion receptors, integrins, they express. In addition to loosening the bonds to the primary tumor, allowing cells to escape, one hypothesis is that these changes in cell adhesion molecules underlie the curious tendency of various cancers to metastasize preferentially to certain other tissues. Melanoma, for example, has a strong tendency to metastasize to the brain and to bone. Melanoma's particular array of abnormal (for skin) adhesion molecules may represent a "postal code" favoring delivery to a particular distant site.

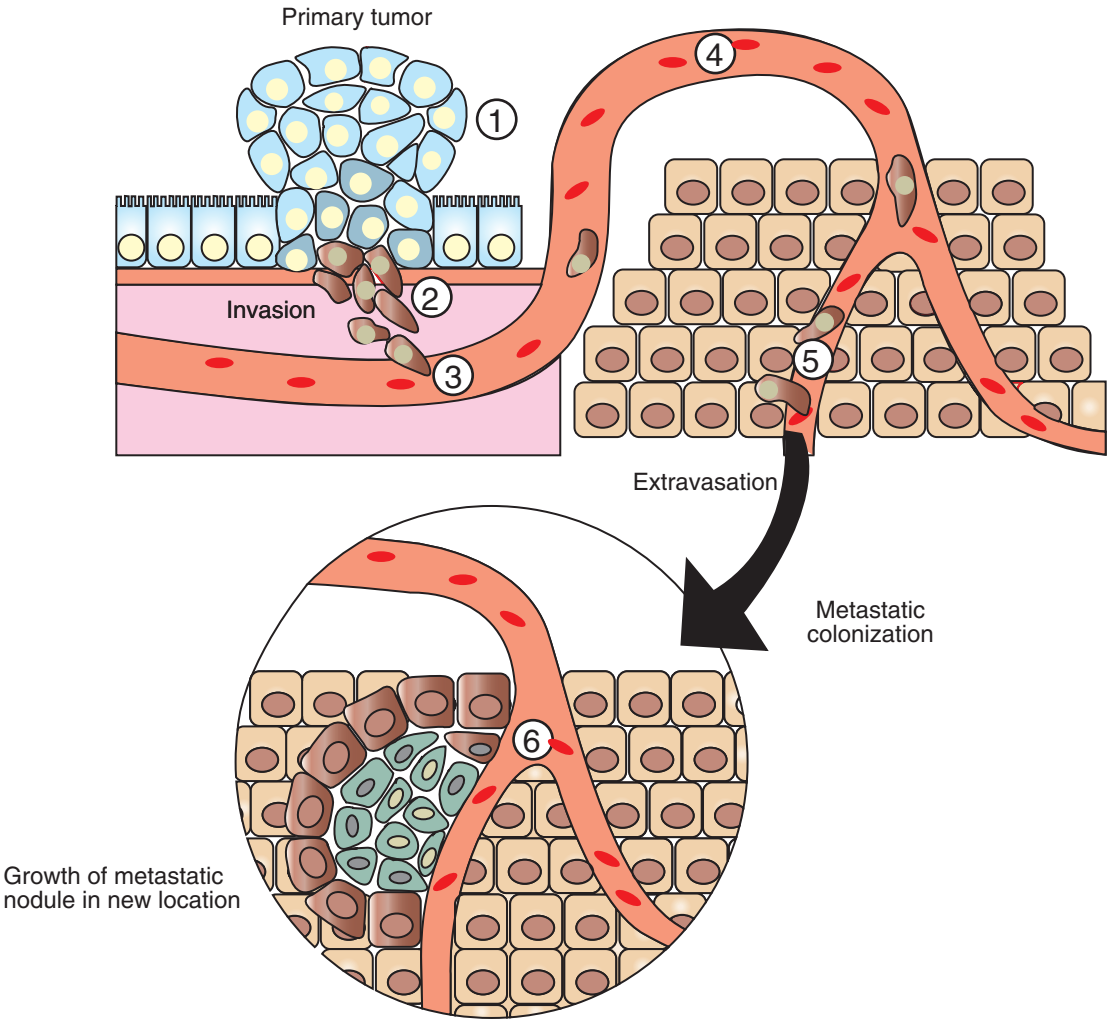
Having altered its adhesion, enabling escape from the primary tumor, the metastatic cell must make its way toward the circulatory system, and enter the circulation (called *intravasation*) to "hitch a ride" around the body. Although "circulation" typically refers to the bloodstream, cancer cells can also be disseminated by traveling within the lymphatic system, which collects extracellular tissue fluid for return to the blood. Indeed, invasion of lymph nodes, which are major collection sites for extracellular fluid and debris, is a common test for initial metastases. For either route, however, the cell's ability to achieve intravasation depends on activation of cell motility and expressing proteases. Most animal cell types are capable of "crawling" locomotion (but do not actually do so) using actin and myosin mechanisms similar to muscle contraction (see Fig. 1.4). This crawling locomotion is roughly similar to the motility of amebae. Migrating breast cancer cells have been imaged directly and show solitary cells with ameboid morphology. The entire actin and myosin system of most cancer cells is dysregulated, causing changes in cell shape and the ability and tendency to locomote. For example, normal skin cells are generally quite stationary, but melanoma cells are highly motile. The dysregulation of the actomyosin system results in part from mutations of the Rho family of small, Ras-like G proteins, mentioned briefly earlier. Mutations in *rho* are common among highly metastatic melanoma cells, but such mutations are rare among weakly metastatic melanoma cell lines.

Because cells in solid tissues are crowded together, increased motility appears to be helped by secretion of proteases that digest some of the cell matrix "obstacles" in the cancer cell's path. Epithelial cells give rise to approximately 80% of human cancers. As noted earlier, all epithelial cells are attached to an ECM, which is characterized by a particular type of collagen. Proteases specific for this type of collagen are generally overexpressed by metastatic cells. The number of different proteases and the net amount of protease secreted tend to increase with increasing metastatic potential. In addition, cancer cells appear to have the capacity to cause surrounding normal cells to increase their secretion of proteases. Proteases not only aid the metastatic cell in intravasation, but also stimulate cell survival and proliferation (e.g., by releasing growth factors bound to the ECM). However, not all ECM represents an obstacle to movement that must be proteolytically degraded. Some types of ECM appear to provide stimulatory pathways for the migrating cells to follow toward the circulation. Migration toward the circulation is also aided by chemoattraction; epidermal growth factor in blood vessels seems to be an attractant for breast cancer cells.

After intravasation, the metastatic cell rides around the circulation until it can attach to a capillary vessel wall. Then, reversing the process of intravasation, the cell "climbs out" of the circulation, which is called *extravasation.* As one might expect, this also depends on changes in adhesion, motility, and protease secretion. After extravasation, the metastatic cell must be able to survive and proliferate in its new environment.

Fewer than 1 in 10,000 cells escaping the primary tumor colonize a new location successfully. It was once assumed that this high rate of failure reflected an "exceedingly rough ride" around the circulatory system. More recent evidence suggests that the limiting factor of metastasis is the survival of the cell in its new location. This represents another example of the natural selection, the "microevolution," occurring in cancer. The foreign environment exerts a strong negative selective pressure on cancer cell arrivals, and the vast majority do not survive. Ironically, our current thinking about metastasis is similar to the "seed and soil hypothesis" first proposed in 1889. Metastasis requires a cancer cell from the primary tumor (the seed) capable of carrying out (selected for) all the steps

Vet-ebooks.com



• **Fig. 2.13** Metastatic cascade, the path from primary tumor to metastatic tumor. Cells of the primary tumor alter their cell adhesion and motility properties to migrate away from the primary tumor site (*1*). These cells secrete proteases to digest their way through the surrounding tissue (*2*). They then crawl into the vasculature (*3*), a process called *intravasation,* where they are then carried passively around the circulation (*4*). At some point, they adhere to the sides of the blood vessel and crawl out of the vasculature (*5*), a process called *extravasation.* Some metastatic cells are able to colonize the new location to form a new, deadly metastatic tumor (*6*).

of the metastatic cascade, and the metastatic cell must colonize a region (the soil) appropriate for its subsequent growth. Normal cells cannot survive in a new location within the same body. Different tissues have different chemical milieus (e.g., different mixes of growth factors), and these are specialized for the survival and growth of the particular cell types found in the region. The process of metastasis necessarily selects for cells capable of growing in a foreign environment. It is thought that most metastatic cells become dormant in their new location and that additional genetic changes are required and must be selected to enable uncontrolled growth. Genes whose products play a role in adapting cells to particular environments have been called *landscaping genes,* another allusion to the current view of metastasis resembling gardening. Mutations in these landscaping genes are postulated to allow growth in the foreign environment, less than 1% of micrometastases grow to a clinically relevant size. A key aspect of this selection process for uncontrolled growth in the foreign location is the secretion by the cancer cells and by surrounding cells of a variety of mediators to stimulate growth of new blood vessels to supply the tumor. Another key feature of selection is remodeling of the local ECM, which normally is required for proliferation. These phenomena of inducing blood supply and remodeling of the ECM are other aspects of metastasis that are poorly understood.

Indeed, it is poorly understood how metastasis actually leads to death except that it often involves a profound and progressive wasting of skeletal muscle and fat in the body called, *cachexia.* This wasting process affects up to 80% of human cancer patients and is responsible for perhaps a third of deaths. Recent experimental results, again with mice, indicate that inhibiting cachexia, even without inhibition of tumor growth, dramatically lengthens the animals' life span. Cachexia is complex but appears to be due in large part to cancer cells releasing signaling molecules that stimulate abnormal developmental pathways in muscle and fat cells. Activation of these developmental pathways, in turn, leads to abnormally high levels of fat and protein breakdown in the affected cells. Another possible mechanism underlying cachexia involves generalized inflammatory reactions, which attack other aspects of the patient's physiology generally. Possibly the presence of foreign cells,

Vet-ebooks.com

selected for growth in an abnormal location causes the body's defense mechanisms to be fully mobilized and attack itself, a nightmare of biological "friendly fire." Presumably, the foreignness of the metastatic tumor explains the highly inappropriate response, which primary tumors typically do not instigate. It is for this reason, in part, that *complete* removal of the primary tumor before metastasis occurs often leads to total recovery. In other cases, death from metastatic disease, like death from some primary tumors, is the result of cancer cells simply overwhelming a vital organ, leading to organ failure.

Growth of Solid Tumors Depends on Development of New Blood Vessels

Tumors, as with normal tissue, require blood vessels to supply them with oxygen and nutrients and to remove waste. Much attention has been focused on the development of tumor capillaries because it is a rate-limiting step in the tumor's growth and progression. Both primary and metastatic tumors require new vessels; without them, the tumor remains too small to be visible or palpable, about 1 to 2 mm in diameter. Dormant tumors of this size have been found in autopsies of people who did not die of cancer, so not all tumors develop the blood supply needed for growth. Thus the ability of tumors to stimulate new blood vessel development is a distinct and important step in tumor progression. As this suggests, it is also a relatively early step in tumor progression but is covered here after metastasis because most new vasculature arises from existing capillaries invading new regions of tissue, sharing some features with metastasis.

The discussion of Starling's hypothesis in Chapter 1 notes that blood capillaries are composed primarily of a single layer of a specialized epithelial cell type, the endothelial cell. The first capillaries in the embryo are formed by *vasculogenesis,* the differentiation of precursor cells (angioblasts) to form a basic capillary network. However, most new capillaries are formed by *angiogenesis,* the sprouting and branching of existing capillaries to supply new tissue regions. Larger blood vessels, such as arterioles and veins, all develop from the subsequent growth of capillaries. In the adult, only angiogenesis normally occurs and depends on invasive cellular processes similar to those involved in metastasis: proliferation of existing endothelial cells; migration of the cells into the region to be supplied, involving changes in actin function and adhesion to the surrounding cells; and remodeling the surrounding ECM so the extending cells intercalate among the tissue cells, ultimately to form a hollow tube. Although the cancer cells within a tumor are abnormal, the endothelial cells composing the new capillaries are normal. Thus tumor capillaries can arise by vasculogenesis (because of the abnormal environment of the tumor) or, and primarily, by angiogenesis. Similarly, the endothelial cells of tumor capillaries respond to the normal stimulatory and inhibitory signals for angiogenesis. Nevertheless, the pathological features of the tumor stimulate abnormal growth of blood vessels, whose pattern, composition, and function differ from normal capillaries.

In normal adult tissue, except for the female reproductive tract, endothelial cells are among the most slowly proliferating cell type. Only 1 in 10,000 adult endothelial cells is in cell division at any one time, compared with about 10% of gut epithelial cells. Normal angiogenesis is under tight regulation by both stimulatory and inhibitory influences. Stimulatory influences include injury and hypoxia, which in turn lead to the secretion of angiogenic growth factors such as *vascular endothelial growth factor* (VEGF, "vedge-eff"). This growth factor strongly stimulates endothelial cell proliferation and migration, and suppresses apoptosis. VEGF also increases permeability of existing vessels. Inhibitory influences include thrombospondin-1, which is an ECM component that inhibits endothelial cell proliferation and motility. Inhibitory influences also include soluble factors such as angiostatin, which stimulates apoptosis in proliferating endothelial cells, and endostatin, which inhibits the migration of endothelial cells. The growth, stasis, or regression of capillaries depends on the balance between pro- and antiangiogenic stimuli, much as cellular life and death depend on the balance between pro- and antiapoptotic signals discussed earlier.

The relative quiescence of normal capillaries is in sharp contrast to capillaries of tumors, which have been compared with wounds that never heal, in that tumor capillaries undergo continuous growth and remodeling. Tumor endothelial cells divide 20 to 40 times more frequently than normal endothelial cells, and tumors typically have a much higher density of vessels than normal tissue. As a result, tumor vasculature is abnormal in structure and function. Tumor vessels can exhibit strange combinations of capillary, venous, and arteriole structures and often incorporate cancer cells as part of the vessel wall. These vessels tend to be convoluted and dilated, follow tortuous paths, and even form dead ends. As a result, blood flow is equally abnormal, with the vessels leakier than normal vessels, and in some cases the blood flows back and forth rather than circulates.

Perhaps the most important factor in this vascular pathology is the high concentration of VEGF in and around tumors. Most human tumors secrete large amounts of VEGF and also cause surrounding tissue to secrete VEGF. Much evidence from experiments on mice supports the crucial role of VEGF in tumor angiogenesis and growth. Antibodies against VEGF suppress growth of existing tumors; cancer cells engineered to be incapable of expressing VEGF were unable to form tumors; and inhibition of the VEGF receptor inhibited the growth of a variety of tumors. In part, the secretion of VEGF by tumor cells seems to be the result of the initial hypoxic conditions of the avascular tumor. Hypoxia is normally a strong inducer of VEGF production, and the centers of many solid tumors show necrotic cells indicative of death from lack of oxygen. In addition, the genetic damage to cells in their progression to a cancer cell seems to contribute to VEGF overexpression. Mutations of *ras* and overexpression of Bcl-2, the antiapoptotic factor, have been shown to play important roles in this regard.

Tumor vessels are also substantially more permeable than normal vessels, to the point of being almost hemorrhagic, which is also thought to be caused by overexpression of VEGF (which has an alternative name of *vascular permeability factor*). The leakiness of tumor vessels has several consequences with respect to tumor physiology, spread, and treatment. The high vascular permeability of tumors is believed to aid metastasis in that metastasis requires intravasation of tumor cells into the circulation, and leakier vessels make this more likely. Leakier vessels also disrupt capillary fluid transport, as discussed in Chapter 1. Recall that capillary filtration and reabsorption depend on the balance between hydrostatic and oncotic forces across the capillary wall. The increased fluid leaking from tumor vessels distends the interstitial space, increasing its hydrostatic pressure and thus reducing the pressure gradient across the capillary wall. The oncotic pressure gradient is also reduced because the leak of proteins into the interstitial space means that the oncotic pressure of the interstitial space approaches that of the blood. The result is uncommonly high net interstitial fluid pressure. This can cause collapse of some vessels, leading to hypoxia of the

Vet-ebooks.com

surrounding tissue and further upregulation of VEGF expression. High interstitial fluid pressure also causes poor fluid transport out of the blood into the tumor. This poor fluid exchange seems to inhibit the delivery of chemotherapeutic agents from the blood to the tumor. Studies on chemotherapy of breast cancer and melanoma show that tumors with high interstitial fluid pressure tend not to respond as well to the therapy.

As with the other insights into tumor biology, the possibility of controlling tumor angiogenesis for therapy is being actively pursued. At this writing, there are more than a dozen antiangiogenesis compounds being tested. One, bevacizumab (Avastin, an antibody to VEGF), is approved as a first-line therapy for metastatic colon cancer, although this same drug has recently been shown to be ineffective for breast cancer. Unlike most cancer therapy that targets the abnormal cancer cell, antiangiogenic therapy would be targeting normal endothelial cells. These cells are not genetically unstable, and therefore development of drug resistance may be less likely (see the following discussion). Also, because normal endothelial cells are unusually quiescent, inhibiting angiogenesis should produce fewer toxic side effects than standard chemotherapies. As with other cell-targeted therapies, however, antiangiogenesis inhibitors that showed dramatic results in preclinical studies have been much less successful in treating patients.

Prospective Cancer Therapy

Cancer Therapy Has a Hopeful but Challenging Future

Most current cancer therapy makes little or no use of the advances in our understanding of the molecular basis of cancer. However, the US Congress has funded a program, "Cancer Moonshot," intended to accelerate cancer therapies by making better use of new knowledge about cancer. Currently, declines in (human) cancer mortality in the industrialized world are primarily the result of better screening for breast and colon cancer and preventive measures (e.g., discouraging smoking). Both chemotherapy and radiotherapy are typically nonselective (at the cell level) cytotoxic treatments intended to shrink the overall size of tumors, with serious side effects from the general cytotoxicity. Clinical trials to test new cancer drugs nearly always enroll large number of patients with no thought of investigating the particular mutations underlying the patient's tumor. This situation is changing slowly, but perhaps at an accelerated pace, to one of *targeted therapy,* in which the genotype of the tumor is taken into account and, if available, drugs targeting the mutations are used preferentially. An example of targeted therapy, mentioned earlier, is the use of vemurafenib to target the Raf mutations occurring in some, but not all, melanomas. A few additional examples of targeted therapy are provided here, but the development of such therapies and of practical molecular diagnosis remains challenging, and often disappointing, with three common themes accounting for treatment failure, reflecting the fundamental properties of cancer. Indeed, a major goal of the Cancer Moonshot is to overcome the difficulties in cancer treatment described later. More information about current work on cancer therapies is provided in a paper on "The Evolving War on Cancer," in the Bibliography for this chapter. Also, the journal *Science,* volume 335, issue 6330, has a special section on emerging cancer therapies.

First, despite the success with Raf and melanoma, the accumulation of mutations, along with the differences in this process from individual to individual, means that single molecular markers have not proved very useful in refining diagnosis. For example, assessing the different mutations occurring in such important genes as *ras* or the *p53* gene in breast cancer has had conflicting results in predicting disease outcomes. Presumably this is because these mutations have differing effects, depending on the other mutations involved in the cancer and their interaction with the normal alleles of the individual patient. As a result, it appears that multiprotein/multigene molecular "signatures" will be needed. If such signatures can be developed as body fluid or other relatively noninvasive tests, it could lead to major improvements in treatment, insofar as diagnosing cancer as early as possible is crucial for a favorable prognosis.

The second common theme is that the multiple types of genetic damage and selective processes required for cancer also function to cause resistance to treatment. That is, the unstable and abnormal genetic status of cancer cells that produce the growth abnormalities also lead to abnormal responses to drugs and other interventions. Ironically, a vivid illustration of this is one of the notable successes for targeted therapy in treating cancer. Chronic myeloid leukemia (CML) is known to begin with a specialized mutation (a particular chromosomal translocation) that disrupts the gene for a specific tyrosine kinase, Abl, so that it becomes an activated oncogene. A fairly specific inhibitor of this tyrosine kinase was developed, imatinib (Gleevec), that blocks binding of ATP, disabling kinase activity. This has had marked benefits for patients in the early, chronic stage of CML, which is debilitating but not fatal. In many patients, this drug causes complete remission of CML and has thus far prevented progression to the fatal, acute stage. However, some patients have developed resistance; in most of these cases the *abl* oncogene has mutated yet again such that ATP binding is restored despite Gleevec. More ominously, careful analysis of the blood of CML patients actually in remission indicates a remaining pool of leukemic cells (cancer stem cells apparently), which may subsequently lead to development of resistance in later years. Nevertheless, there are currently more than 20 specific protein kinase inhibitors in clinical trials, and practitioners would welcome additional drugs with the effectiveness of Gleevec, despite its limitations.

In addition to outright mutations leading to cancer, we have mentioned several examples in which changes in gene expression of normal proteins stimulate cancer development. Such a situation underlies another early success of targeted therapy based on a single genetic lesion. Non–small-cell lung cancer is the leading cause of death from cancer in the United States. Some 40% to 80% of these cancers overexpress the epidermal growth factor receptor (EGFR), which is an RTK as described earlier. Gefitinib (Iressa) blocks the ATP-binding site of EGFR, inhibiting kinase activity in a fashion similar to Gleevec. As with vemurafenib and *raf* mutations, gefitinib has been shown to be effective among those patients whose non–small-cell lung cancer depends on mutation of the receptor but not among patients with normal EGFR.

However, changes of gene expression also present a challenge to cancer therapy, particularly with respect to drug resistance, the mechanism of which underlies another, broader example of the obstacle to treatment presented by the genetic status of cancer cells. *Multiple-drug resistance* (MDR) is a phenotype in which cells develop resistance to many current, initially effective, chemotherapeutic agents for a wide variety of cancers. This is the result of the overexpression of a pump protein that causes the efflux of many different drugs from the cell. As with the selection among cancer cells for continued ability to proliferate, administration of the drug selects for those cancer cell variants that have changes in gene

Vet-ebooks.com

expression, such that the efflux pump reduces the effectiveness of the drug. Thus new drug development must contend not only with the genetics of cancer but also the genes and gene expression involved in drug resistance. (An interesting aspect of the drug efflux pump often expressed in cancer cells is that it is also expressed in normal stem cells!)

The third common theme identified as an obstacle to molecular cancer treatments is that, as discussed, cancers reflect physiological dysfunctions at a particularly fundamental level. It is not easy to interfere with these functions without compromising other functions, or interference engages compensatory mechanisms normally serving as "backup" to crucial functions. At the simplest level, interventions that alter these basic mechanisms of cellular life and death often prove to be too disruptive to the physiology of some normal cells to be useful. For example, although vemurafenib, the Raf inhibitor mentioned earlier, and gefitinib, the EGFR inhibitor, have proven effective against some cancers, many other inhibitors in the growth factor/MAP kinase pathway (see Fig. 2.5) that showed promise on cultured cells and in mice proved to be too toxic for therapeutic use.

Other results indicate that effective treatments will need to take into account the normal molecular biology of the cell very closely. Experiments attempting to target *p53* are noteworthy in this respect. Because mutation of one *p53* gene will predispose to cancer (if the other copy were lost, an important checkpoint would be lost), activation of the remaining normal copy might protect against cancer. Such enhanced P53 activity did protect against cancer in mice, but the mice also showed notably shortened life spans and visible signs of early aging. As shown by this unexpected role of P53 in aging, the central roles in the life of the cell played by proto-oncogenes and tumor suppressor genes mean that they often have multiple roles that complicate development of therapies. In experiments in which expression of activated P53 was limited to mammary tissue, mice were again protected against cancer but at the cost of inhibiting lactation and mammary development. The best anticancer results obtained from experimentally manipulating P53 expression has come from experiments in which whole artificial chromosomes with the *p53* gene and all its normal control elements were introduced into mice. These mice showed increased resistance to chemically induced cancers with no apparent effects on aging. Introducing genes with all relevant control elements, however, is a rather high hurdle for practical therapies.

Finally, the importance of these normal genes and proteins to cell function means that they often have redundant mechanisms of control. This seems to apply to that other "usual suspect" in cancer, Ras/*ras*. Evidence that association with the plasma membrane via lipid "tails" was required for Ras activity (similar to the alpha subunit of the heterotrimeric G protein, see Fig. 1.14) led to the development of drugs, *farnesyl transferase inhibitors* (FTIs), that block addition of the lipid tail. Although FTIs have proved clinically useful against some types of cancer in some patients, their effects are highly variable. One idea is that the FTIs only inhibit one pathway for Ras membrane association. Used alone, these drugs showed only modest effects on tumors, but in combination with standard chemotherapeutic drugs, FTIs worked relatively well on some cancers. However, it was puzzling that some cancers importantly involving *ras* mutations, such as lung cancer, were not affected by the inhibitors. Further, some *ras*-independent tumors were just as susceptible to FTIs. It now seems that these drugs may not be acting only through Ras membrane association. The Bibliography includes a paper (Papke and Der) on current attempts to target Ras in cancer therapy.

As noted earlier, standard chemotherapies and radiation therapies are highly toxic by the usual standards of clinical practice. Cancer therapy is a prime medical example of "drowning men grasping at straws." Thus the handful of clear successes using targeted therapy based on advances of our molecular understanding of cancer are widely regarded as being hopeful. But the effectiveness of chemical therapies is colored by the enormous success against infectious diseases with antibiotics and vaccines, and of preventing organ-system disease, e.g., cardiovascular disease with GPCR-targeted drugs. These may prove unrealistic models of success for disease at a deeply cellular, genetic level, such as cancer. For veterinary practitioners, a very welcome development would be the use of a domestic animal as a model of cancer, particularly for the development of therapies. The Bibliography for this chapter includes two papers (by Pang et al. and Schiffman et al.) describing the potential advantages of the dog as a cancer model.

CLINICAL CORRELATIONS

Dog That Collapsed While Running

History

A spayed, 10-year-old female golden retriever collapsed while running outside earlier today. The dog is still very lethargic and does not want to move.

Clinical Examination

The dog has pale mucous membranes with a normal temperature. The capillary refill time is prolonged. Heart rate and respiratory rate are increased. On palpation, there appears to be fluid in the abdomen, and the dog is in pain.

Comment

Based on this history and physical examination, there is concern that this dog has hemorrhaged into the abdomen. Hemangiosarcoma is a common tumor of older dogs and originates from a transformed endothelial cell. Dogs often present after having collapsed when the tumor, which is present in the spleen, causes internal bleeding. The dogs must often have emergency surgery to have a splenectomy (spleen removed). In some cases, dogs may show other, nonspecific clinical signs (inappetence, lethargy), so a diagnosis may be made before the dog collapses from acute bleeding. A diagnosis is often made through a combination of modalities, including radiographs, ultrasound, biopsies, histopathology, and immunohistochemistry, to determine the nature of the tumor. In many cases, by the time the diagnosis has been made, the tumor has already metastasized, usually via hematogenous route, to other organs. The lung and liver are more frequently affected, but other sites include kidney, muscle, brain, mesentery, skin, and lymph nodes. Recently, it has been demonstrated that canine hemangiosarcomas express platelet-derived growth factor beta (PDGF-β). Suppression of this RTK signaling using imatinib (Gleevec) suppressed the canine cell line in a mouse model.

Treatment

Treatment depends on the stage at which the tumor is diagnosed; in this case, the animal presents with shock and hemorrhage. In these cases the patient is stabilized, surgery is performed, and the spleen (in this patient) is removed. The overall prognosis for these cases is poor because the tumor has usually metastasized by the time the initial diagnosis is made. Radiation therapy is palliative for these cases, and it is sometimes used when there is a large, local unresectable mass. Chemotherapy is usually the treatment of

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

choice, although median survival time for these dogs typically is not long. Medications often include the VAC protocol: doxorubicin, cyclophosphamide, and vincristine. Doxorubicin inhibits DNA synthesis, DNA-dependent RNA synthesis, as well as protein synthesis, and it acts throughout the cell cycle. Cyclophosphamide inhibits DNA replication as well as RNA transcription and replication. Vincristine binds to specific microtubular proteins to inhibit cell division. Complications associated with chemotherapy include myelosuppression and sepsis. Experimental treatments are still being tested that target endothelial cells, blocking adhesion factors and inhibiting growth factors associated with endothelial cell growth.

8-Year-Old Male Neutered Boxer with Seizure

History

An 8-year-old boxer male neutered dog presents to the clinic with a 2-day history of not acting normally. The owners also noticed that he did not want to bend his neck or eat. Two hours before presentation, he seizured for about 30 seconds.

Clinical Examination

The dog appeared mentally dull on examination. His temperature, pulse, and respiratory rate were all increased. He was resistant to palpation of his neck, and he now had a head tilt to the right. The dog had not traveled out of the country, nor had any of the other dogs in the household. The dog was current on vaccinations. Based on the findings that the dog was seizuring, mentally dull, and had neck pain, those signs are all consistent with neurological disease. One of the other major causes of seizures can be hypoglycemia, which can be ruled out with testing his blood glucose levels.

Comment

To make a diagnosis, blood work and cerebrospinal fluid (CSF) were collected. CSF showed increased protein, and increased numbers of neutrophils and lymphocytes present. MRI was performed and samples were collected for histopathology which localized the tumor to within the brain. A diagnosis of malignant glioblastoma multiforme (GBM) was made. Radiographs of the chest and abdomen did not reveal evidence of metastasis.

Treatment

Prognosis has tended to be very poor with survival rates of 0.2–4.9 months depending on care from supportive to cytoreductive surgery to combination (radiotherapy and surgery or cryotherapy). Most recently, an irreversible electroporation (IRE) therapy has shown great promise in a small study of seven dogs. IRE involves nonthermal ablation, principally using minimally invasive electrodes within or around the tumor tissue, to use short high voltage electric pulses to the tissue. This destabilizes the cell membranes of the neoplastic cells. MRIs were repeated after therapy. Four of the five dogs saw improvement. Although refinement is needed, this study is promising for the treatment of intracranial malignant glioblastoma multiforme (GBM).

Practice Questions

1. Which of the following is associated with normal stimulation of cellular proliferation?
- **a.** Oncogenes
- **b.** Tumor suppressor genes
- **c.** Telomerase
- **d.** Proto-oncogenes
- **e.** Caspases

2. In the growth factor pathway, the growth factor first binds to _____, which leads to activation of _____, in turn causing activation of a cascade of _____ enzymes leading to alterations in transcription.
- **a.** G-protein–coupled receptors; G proteins; adenylyl cyclase
- **b.** receptor tyrosine kinases; Ras; MAP kinase
- **c.** receptor tyrosine kinases; Bcl-2; caspase
- **d.** cyclin-dependent kinases; pRb; telomerase
- **e.** tumor suppressors; oncogenes; checkpoint

3. Which of the following mediate(s) apoptosis?
- **a.** Telomerase
- **b.** Cytochrome c
- **c.** Receptor tyrosine kinases
- **d.** Cyclin-dependent kinases
- **e.** Cyclins

4. The tumor suppressor pRb is a(n) _____ and participates in regulating the cell cycles of _____ cells, whereas P53 is a(n) _____ and participates in regulating the cell cycle of _____ cells.
- **a.** inhibitor of transcription; healthy; transcription factor; healthy
- **b.** transcription factor; damaged; inhibitor of apoptosis; healthy
- **c.** caspase; damaged; inhibitor of transcription; damaged
- **d.** inhibitor of transcription; healthy; receptor tyrosine kinase; healthy
- **e.** inhibitor of transcription; healthy; transcription factor; damaged

5. Normal stem cells are similar to cancer cells but differ from normal somatic cells in that normal stem cells and cancer cells both:
- **a.** Are missing checkpoint controls on the cell cycle.
- **b.** Have cell cycles that are independent of activation of cyclin-dependent kinases.
- **c.** Have activated telomerase.
- **d.** Are resistant to apoptosis in response to DNA damage.
- **e.** Are able to metastasize to distant, foreign tissues.

Vocabulary

This chapter is unusual in that it contains a large number of vocabulary words that are specialized for cancer and related topics, words that generally will not be used in later chapters. You should be familiar with these vocabulary words; you should be able to define them and state their role in normal cells, and whether and how they differ in cancer cells.

anchorage dependence of growth
aneuploidy
angiogenesis
apoptosis (intrinsic pathway and extrinsic pathway)
Bcl family
cachexia
caspase(s) (activating and executioner)
cell cycle (G1 phase, S phase, G2 phase, M phase)
checkpoint
contact inhibition of growth
cyclin
cyclin-dependent kinase (CDK)
extravasation

Vet-ebooks.com

gain-of-function mutation
germ line
intravasation
loss-of-function mutation
MAP kinase pathway
MDR phenotype
metastasis
metastatic cascade
oncogenes
P53 primary tumor
Ras
retinoblastoma (pRb)
somatic cells
somatic mutation
stem cells
targeted therapy
telomerase
telomeres
tumor suppressor
vasculogenesis
VEGF

Bibliography

Batlle E, Clevers H. Cancer stem cells revisited. *Nature Med*. 2017;23:1124–1134.

Bieging KT, Mello SS, Attardi LD. Unravelling the mechanisms of p53-mediated tumour suppression. *Nature Rev Cancer*. 2014;14:359–370.

Chan SR, Blackburn EH. Telomeres and telomerase. *Philos Trans R Soc Lond B Biol Sci*. 2004;359:109–121.

Goldman JM, Melo JV. Targeting the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med*. 2001;344:1084–1086.

Haber DA, Gray NS, Baselga J. The evolving war on cancer. *Cell*. 2011;145:19–24.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646–674.

Harris SL, Levine AJ. The p53 pathway: positive and negative feedback loops. *Oncogene*. 2005;24:2899–2908.

Hengartner MO. The biochemistry of apoptosis. *Nature*. 2000;407:770–776.

Klopfleisch R, von Euler H, Sarli G, et al. Molecular carcinogenesis of canine mammary tumors: news from an old disease. *Vet Pathol*. 2011;48(1):98–116.

Kreso A, Dick JE. Evolution of the cancer stem cell model. *Cell Stem Cell*. 2014;14:275–291.

Krontiris TG. Oncogenes. *N Engl J Med*. 1995;333:303–306.

Langley RR, Fidler I. The seed and soil hypothesis revisited—the role of tumor-stroma interactions in metastasis to different organs. *Int J Cancer*. 2011;128(11):2527–2535.

Messerschmidt JL, Prendergast GC, Messerschmidt GL. How cancer cells escape immune destruction and mechanisms of action for the new significantly active immune therapies: helping non-immunologists decipher recent advances. *Oncologist*. 2016;21:233–243.

Pang LY, Argyle DJ. Veterinary oncology: biology, big data and precision medicine. *Vet J*. 2016;213:38–45.

Papke B, Der CJ. Drugging Ras: know the enemy. *Science*. 2017;355:1158–1163.

Schiffman JD, Breen M. Comparative oncology: what dogs and other species can teach us about humans with cancer. *Philos Trans R Soc Lond B Biol Sci*. 2015;370:1673–1685.

Servick K. The stem cell skeptic. *Science*. 2017;357:441–443.

Tisdale MJ. Molecular pathways leading to cancer cachexia. *Physiology (Bethesda)*. 2005;20:340–348.

van Niekerk G, Davids LM, Hattingh SM, et al. Cancer stem cells: a product of clonal evolution? *Int J Cancer*. 2017;140:993–999.

Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*. 2002;108:171–182.

Vet-ebooks.com

3 Introduction to the Nervous System

BRADLEY G. KLEIN

KEY POINTS

1. The neuron is the major functional unit of the nervous system.
2. The mammalian nervous system has two major subdivisions: the central nervous system and the peripheral nervous system.
3. The central nervous system can be divided into six anatomical regions.
4. The central nervous system is protected by the meninges and cerebrospinal fluid.
5. The nervous system collects and integrates sensory information, formulates a response plan, and produces a motor output.

The nervous system is the first multicellular system described in this book because it is one of the major coordinating systems of the body and because clarifying many of the concepts that concern the nervous system is important for understanding other systems of the body.

Most clinical signs in veterinary neurology involve abnormal movement (e.g., seizures, paralysis); therefore, the physiology of muscle control and its central nervous system coordination are emphasized in the following chapters. Because veterinary ophthalmology is an extensive subspecialty, the physiology of vision is also covered. Other sensory systems that can produce easily recognizable clinical signs (e.g., vestibular system, hearing) are discussed in Section II as well. Understanding the autonomic nervous system is essential for understanding pharmacology and the involuntary control of many of the body's most critical functions. Similarly, understanding the blood-brain barrier and the cerebrospinal fluid system is essential to understanding the results of the diagnostic cerebrospinal fluid tap and the homeostasis of the cellular microenvironment of the central nervous system. The electroencephalogram and sensory-evoked potentials are described because of their clinical importance in veterinary medicine. Because of space limitations, only the basic physiological concepts essential to understanding the mechanisms of disease and the practice of veterinary medicine are emphasized. For a more expansive study of neurophysiology, the reader should refer to the texts listed in the chapter Bibliographies.

The Neuron Is the Major Functional Unit of the Nervous System

The major functional unit of the nervous system is the *neuron,* a cell type whose shape can vary considerably with its location and with its specific role in the nervous system. Almost all neurons have an information-receiving area of the cell membrane, usually composed of radiating *dendrites;* a cell body, or *soma,* containing the organelles for most cell metabolic activity; an information-carrying extension of the cell membrane, called an *axon;* and a *presynaptic terminal* at the end of the axon to transmit information to other cells. The axon is often covered with a punctuated fatty coating called the *myelin sheath* that enhances the speed of information transfer along the axon's length. Neurons are the principal integrators and rapid communicators of information in the nervous system.

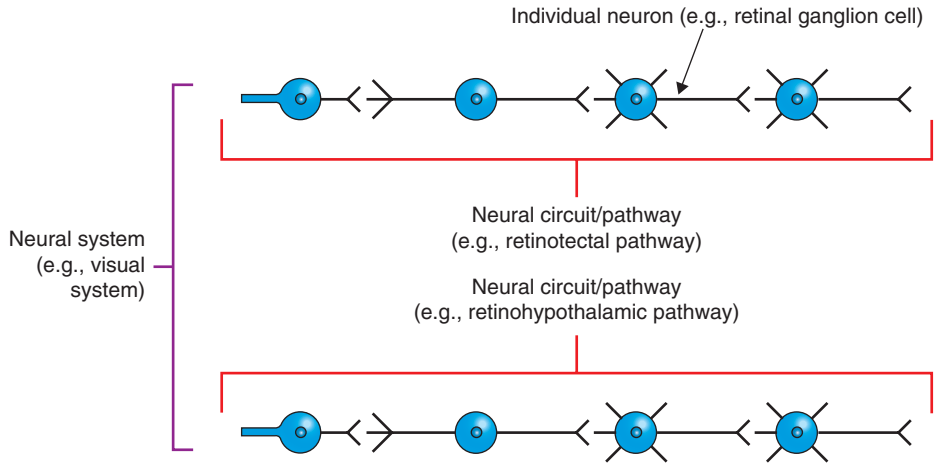
Neurons do not exist in isolation; they are usually interconnected within neural circuits or pathways that serve a specific function (Fig. 3.1). Neural circuits/pathways that are related in function are often collectively referred to as *neural systems.* For example, the retinotectal pathway provides information for reflex orientation of the eyes to the position of a light source, whereas the retinohypothalamic pathway carries information affecting the body's physiological rhythms in response to light-dark cycles. These individual neural pathways are both part of the visual system.

The other cell type in the nervous system is the *glial cell.* These cells are important in maintaining the structural and functional integrity of neurons. Glial cells play important roles in producing the myelin sheaths of axons, modulating the growth of developing or damaged neurons, buffering extracellular concentrations of potassium and neurotransmitters, formation of contacts between neurons (synapses), and they participate in certain immune responses of the nervous system. Glial cells do not produce action potentials, but growing evidence indicates that they can indirectly monitor the electrical activity of neurons and use this information to modulate the effectiveness of neural communication. However, not all glial actions are beneficial to the nervous system. Glial-mediated neuroinflammatory responses have been implicated in some neurodegenerative diseases and in the development of chronic pain conditions.

The Mammalian Nervous System Has Two Major Subdivisions: the Central Nervous System and the Peripheral Nervous System

The *central nervous system* (CNS) is divided into the brain and spinal cord (Box 3.1). A series of protective bones surround the

Vet-ebooks.com



• **Fig. 3.1** Individual neurons are usually interconnected within neural circuits or pathways. Neural circuits/pathways that are related in function are often collectively referred to as neural systems.

• BOX 3.1 Organization of the Nervous System

Central Nervous System (CNS)

Brain
Spinal cord

Peripheral Nervous System (PNS)

Efferent (Motor)

Somatic—to skeletal muscle
Visceral—to cardiac muscle
—to smooth muscle
—to some glands

Afferent (Sensory)

Somatic—from skin
—from retina
—from inner ear membranous labyrinth
Visceral—from thoracic and abdominal organs

entire CNS. The brain is surrounded by the skull, and the spinal cord is surrounded by a series of cervical, thoracic, and lumbar vertebrae and ligaments.

The *peripheral nervous system* (PNS) is composed of the spinal and cranial *nerves* that carry electrical signals, called *action potentials,* away from or toward the CNS. These nerves are bundles of PNS axons. The axons carrying action potentials toward the CNS are called *afferents,* and those carrying such signals away are *efferents.* One way to group the elements of the PNS functionally is into sensory and motor subsystems. The elements of spinal and cranial nerves that serve a motor function are (1) axons of somatic efferent neurons, which carry action potential commands from the CNS toward junctions, called *synapses,* at skeletal muscles, and (2) axons of visceral efferent neurons, which carry action potentials toward synapses with peripheral neurons that control smooth muscle, cardiac muscle, and some glands. PNS components serving a sensory function are axons of afferent neurons that bring action potential messages to the CNS from peripheral sensory receptors. These receptors are directly or indirectly responsible for transducing energy or chemical stimuli from the body's external or internal environment into action potentials that travel to the CNS. The intensity of this stimulation of the receptor is often encoded by changing the frequency of action potentials as the intensity of stimulation changes.

Spinal and cranial nerve sensory components are axons of (1) somatic afferent neurons and (2) visceral afferent neurons. *Somatic* afferent axons carry action potentials resulting from stimulation of receptors such as the photoreceptors of the eye, auditory and vestibular receptors of the inner ear, tactile receptors of the skin, and stretch receptors of skeletal muscle. Action potentials generated by stretch receptors or chemoreceptors (e.g., O_2, CO_2) located within visceral organs of the chest and abdomen are carried to the CNS along *visceral* afferent axons. Most visceral efferent and afferent axons are part of the *autonomic nervous system,* which is made up of the portions of the PNS and CNS responsible for involuntary control of smooth muscle, cardiac muscle, some glands, and many physiological life support functions (e.g., heart rate, blood pressure, digestion).

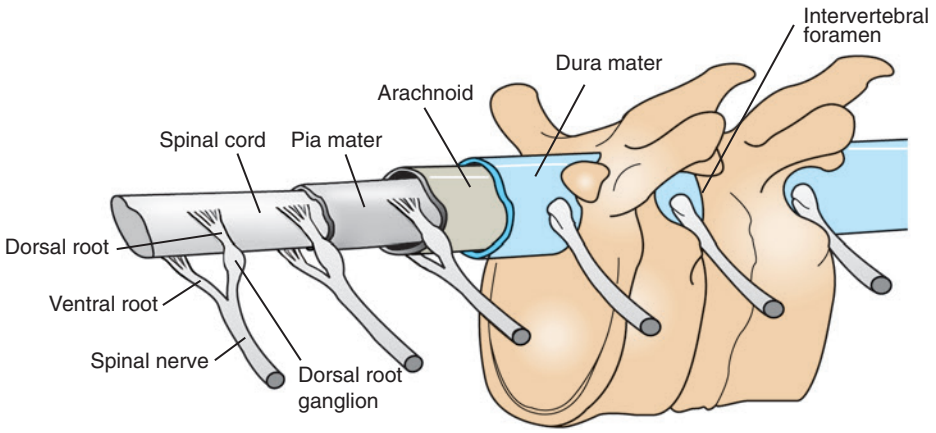
Peripheral nerve axons traveling to or from particular regions of the trunk and limbs converge to form a single spinal nerve at each of the intervertebral foramina. Within the spinal canal, afferent sensory and efferent motor axons are separated; afferent sensory axons enter the spinal cord through the dorsal roots, whereas the efferent motor axons exit the spinal cord through the ventral roots (Fig. 3.2). Peripheral nerve axons running to or from the face and head form cranial nerves of the brain that run through various foramina of the skull.

The PNS and CNS differ in the regenerative ability of their neural axons following physical injury. Peripheral nerve axons can slowly regrow and reconnect to their peripheral targets. Damaged CNS axons do not effectively regenerate due, in large part, to inhibitory features of their local environment. Experimental manipulations of this environment have been shown to improve CNS axonal regrowth. Glial cells are important components of this environment and further understanding of their pro- and antiregenerative responses to neural injury holds promise for therapeutic management of both CNS and PNS axonal renewal after damage.

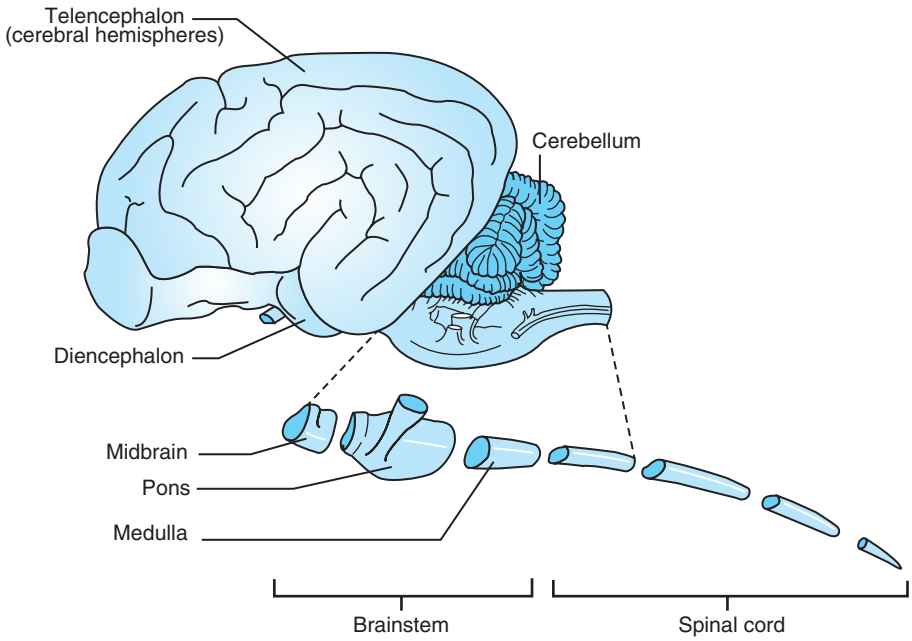
The Central Nervous System Can Be Divided Into Six Anatomical Regions

The CNS has a lengthwise organization characterized by the phylogenetically oldest parts lying more caudal and the newest portions lying rostral. The CNS can be divided into six major regions (Fig. 3.3): the spinal cord and five major brain regions.

Vet-ebooks.com



• **Fig. 3.2** Spinal cord and the three layers of the meninges within the vertebral canal. Action potentials generated on sensory afferents enter the spinal cord along axons in the dorsal roots. Those generated on motor efferents exit the spinal cord along axons in the ventral roots. (Redrawn from Cunningham & Klein: Textbook of Veterinary Physiology; Introduction to the Nervous System Elsevier, 2007.)



• **Fig. 3.3** The central nervous system (CNS) has a lengthwise organization in which the phylogenetically oldest parts are caudal and the newest parts are rostral. The CNS can be divided into six major regions: the spinal cord, medulla, pons, midbrain, diencephalon, and telencephalon (cerebral hemispheres).

From caudal to rostral, these brain regions are the medulla, pons, midbrain, diencephalon, and telencephalon. (The cerebellum, a brain structure that lies dorsal to portions of the pons and medulla, is sometimes named as a seventh major region of the CNS.) The medulla, pons, and midbrain form the *brainstem;* the diencephalon and telencephalon form the *forebrain.*

In general, the spinal cord, brainstem, and forebrain represent a hierarchy of functional organization. The spinal cord receives sensory input from and supplies motor output to the trunk and limbs; the brainstem performs these functions for the face and head. Therefore, simple sensorimotor reflexes can be organized at each of these levels. Sensory information entering the brainstem is passed to the forebrain, where the most sophisticated forms of information processing take place. Sensory information entering the spinal cord can also be relayed to the forebrain by way of the brainstem. The forebrain also formulates the most sophisticated types of motor commands. This output is sent to the brainstem for executing movement of the face and head or for relay to the spinal cord to execute trunk and limb movement. The forebrain is also capable of sending motor commands directly to the spinal cord. Bundles of axons running from one location to another in the CNS are called *tracts.*

Each of the six CNS regions has distinctive anatomical and functional characteristics. Some of these include the following:

1. The *spinal cord* is the most caudal region in the CNS. Sensory dorsal root axons carry action potentials to the cord that were generated by stimulation of sensory receptors in skin, muscles, tendons, joints, and visceral organs. The spinal cord contains the cell bodies and dendrites of motor neurons whose axons exit through the ventral roots either to reach

Vet-ebooks.com

skeletal muscles or to reach out toward smooth muscle. It also contains tracts of axons carrying sensory information to the brain and motor commands from the brain to the motor neurons. The spinal cord is capable of controlling simple reflexes, such as muscle stretch reflexes and limb withdrawal from painful stimuli, even if disconnected from the brain.

2. The *medulla* lies rostral to the spinal cord and resembles it in many ways. By way of cranial nerves, the medulla too receives information from the body's external and internal sensory receptors and sends motor commands out to skeletal and smooth muscle. Large populations of these receptors and muscles lie in the head and neck region. The cell bodies of medullary neurons that receive the sensory input from cranial nerves or that send the motor output are respectively collected in aggregates called sensory or motor *cranial nerve nuclei.* The cranial nerve nuclei of the medulla play a critical role in life support functions of the respiratory and cardiovascular systems and in aspects of feeding (e.g., taste, tongue movement, swallowing, digestion) and vocalization.
3. The *pons* lies rostral to the medulla and contains the cell bodies of large numbers of neurons in a two-neuron chain that relays information from the cerebral cortex to the cerebellum. The *cerebellum* is not a part of the brainstem but is often described along with the pons because of a similar embryological origin. The cerebellum is important for smooth, accurate, coordinated movement and for motor learning. Cranial nerve nuclei of the pons play important roles in receiving sensory information about orofacial touch and in the motor control of chewing.
4. The *midbrain,* or *mesencephalon,* lies rostral to the pons and contains the superior and inferior colliculi, which are important in processing and relaying, respectively, visual and auditory information that has entered the brain at other levels. The midbrain also contains cranial nerve nuclei that directly control eye movement and that induce pupillary constriction. Some midbrain regions coordinate particular eye movement reflexes. A midbrain region called the *periaqueductal gray* plays an important role in endogenous control of analgesia.

   Each region of the brainstem contains axon tracts carrying action potentials to or from the forebrain, as well as tracts that carry action potentials to or from the spinal cord. Discrete aggregates of CNS neuronal cell bodies, that usually give rise to the axons of tracts, are called *nuclei* (not to be confused with the DNA-containing organelle). Each brainstem region also contains a portion of the *reticular formation,* a complex of less tightly aggregated or smaller nuclei mixed with more loosely organized axonal projections, located in the brainstem core. The reticular formation plays important roles in modulating more global functions of the nervous system that include consciousness and arousal, pain perception, and spinal reflexes, as well as movement.
5. The *diencephalon* contains the thalamus and the hypothalamus, both of which are large structures consisting of several subnuclei. The *thalamus* is a relay station for and a modulator of information being passed to the cerebral cortex from sensory systems and other brain regions. The *hypothalamus* regulates the autonomic nervous system, controls hormone secretion of the pituitary gland, and plays a major role in physiological and behavioral aspects of homeostasis (e.g., maintenance of temperature and blood pressure; feeding).
6. The *telencephalon,* also commonly referred to as the *cerebral hemispheres,* is made up of the cerebral cortex and a small number of prominent subcortical structures, such as the basal ganglia and hippocampus. The *cerebral cortex,* the outermost cell layers of the cerebral hemispheres, mediates the most complex forms of sensory *integration* and conscious sensory perception. It also formulates and executes sequences of voluntary movement. The *basal ganglia* are a collection of nuclei that modulate the motor functions of cerebral cortex, and the *hippocampus* plays an important role in memory and spatial learning. Considering the function of the hippocampus, it is fascinating that it is one of the very few regions of the adult mammalian brain where new neurons are born.

The Central Nervous System Is Protected by the Meninges and Cerebrospinal Fluid

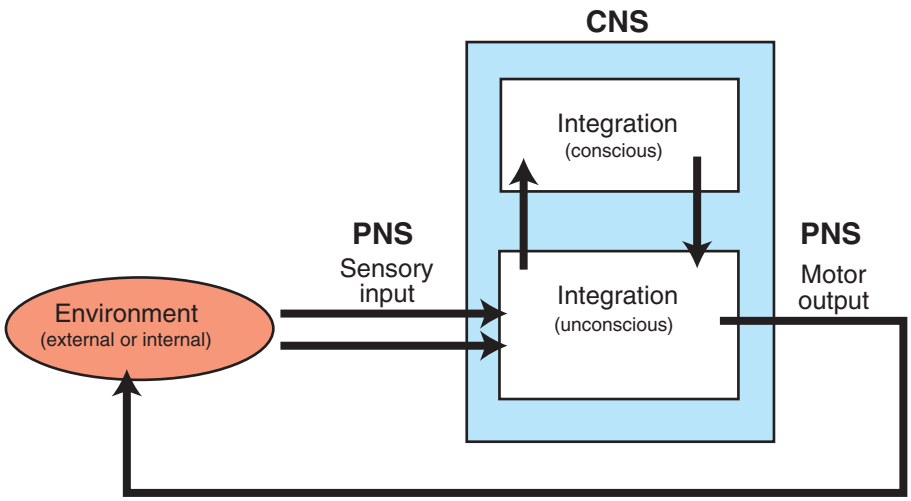
The entire CNS is surrounded by three protective layers called *meninges:* the pia mater, arachnoid, and dura mater (see Fig. 3.2). The innermost layer, lying next to the CNS, is the *pia mater,* which is a single layer of fibroblast cells joined to the outer surface of the brain and spinal cord. The middle layer, the *arachnoid,* so named because of its spiderweb appearance, is a thin layer of fibroblast cells that traps cerebrospinal fluid between it and the pia mater (in the *subarachnoid space*). The outermost meningeal layer, the *dura mater,* is a much thicker layer of fibroblast cells that protects the CNS. Within the brain cavity of the skull, the dura mater is often fused with the inner surface of the bone.

Cerebrospinal fluid (CSF) is a clear, colorless fluid found within the subarachnoid space, the central canal of the spinal cord, and the ventricular system of the brain (see Chapter 15). CSF is produced primarily in the ventricles of the brain, flows down a pressure gradient from the ventricles into the subarachnoid space, where it bathes the surface of the CNS, and from the subarachnoid space eventually passes into the venous system. It is a dynamic fluid, being replaced several times daily. Because CSF can exchange freely with the extracellular fluid of the CNS, it is an important determinant of the neuronal microenvironment, carrying away metabolic waste and potentially harmful macromolecules, and providing certain micronutrients. It can also serve as an important diagnostic tool to indicate CNS infection, inflammation, or tumor activity. The CSF together with the meninges constitutes a fluid-filled, membranous container within which the CNS "floats." Therefore the CSF also serves as a shock absorber for the CNS during abrupt body movement.

The Nervous System Collects and Integrates Sensory Information, Formulates a Response Plan, and Produces a Motor Output

In simplest terms, the nervous system (1) collects sensory information from its external or internal environment, (2) consciously or unconsciously integrates these various inputs to formulate a response plan, and (3) produces a final motor output that can either change the environment (external or internal) or keep it constant (Fig. 3.4). Collecting sensory information and executing the final motor output are the primary responsibilities of the PNS, whereas integration is primarily performed by the CNS. As discussed in Chapter 4, these same functions occur at the level of the individual neuron, which is the principal building block of the nervous system.

Vet-ebooks.com



• **Fig. 3.4** General functional organization of the nervous system. Sensory input and motor output are primarily mediated by the peripheral nervous system (PNS). Integration is a principal role of the central nervous system (CNS).

CLINICAL CORRELATIONS

Neurological Disease in a Horse

History

A client calls and asks you to look at a 4-month-old Arabian filly. The owners have had her since birth, and she has always seemed a little clumsy compared with other foals. They think she is getting worse, however, and say she stumbles in the field. She falls over at times when she is playing with the other foals, and she seems very stiff, almost stabbing at the ground when she is walking.

Clinical Examination

The filly is bright and alert. Her temperature, pulse, and respiration are normal. Abnormalities are limited to your neurological examination. She is weak (paresis) in both the hind and the front limbs (grade II), with the hind limbs being worse (grade III). When you assess her conscious proprioception (ataxia), she is also greatly delayed (grade III hind limbs, grade II front limbs). When she walks, the filly seems to slap at the ground (hypermetria), and she drags her toes forward across the ground. You detect no other neurological deficits.

Comment

This filly has *equine degenerative myeloencephalopathy.* An antemortem diagnosis is difficult. Exclusion of other causes is important. Serum vitamin E levels are often, but not exclusively, low. A definitive diagnosis is made at necropsy.

The pathogenesis of the disease is not clear, but risk factors include diets low in vitamin E, use of insecticides, keeping animals on dirt lots, and exposure to wood preservatives. On histopathology, significant changes occur in the medulla and spinal cord. There is diffuse neuronal degeneration of the white matter. Astrocytosis and lipofuscin-like pigment accumulate in affected areas. Demyelination is marked.

Animals with this disease have loss of functional neurons as well as the myelin sheath that surrounds them. As a result, the ability to conduct impulses is greatly affected. Clinically, this affects the animal's ability to respond to external stimuli as well as initiate conscious responses.

Treatment

Supportive treatment is the only therapy that can be given. Keeping horses on green pasture has been shown to be somewhat protective. Supplementing with vitamin E can improve some horses' condition and slow the progression of disease. There are some familial tendencies in Arabian, Appaloosa, Thoroughbred, and Paso Fino horses.

Neurological Signs in 10-Year-Old Male Neutered Pembroke Welsh Corgi

History

The owners have brought their dog to you because they have noticed he is having a more difficult time walking using his hind end over the last 2 months. They think he is more affected on the right versus the left. They thought it might just be arthritis, but they want you to check their dog to ensure nothing else is wrong. They put him on a nutraceutical (a food component used for medicinal purposes) for arthritis, but it doesn't seem to have helped, as they think the dog is having more issues walking and running and negotiating stairs.

Clinical Examination

The temperature, pulse, and respiratory rate are normal. The dog's mentation seems normal. He has muscle atrophy over both his right and hind gluteal muscles, with the right being more affected. He has decreased patellar responses and has paresis (weakness) of both hind legs, with the right being more severely affected. No evidence of trauma is present. Complete blood count and biochemical profile are normal. CSF is collected and demonstrates increased cells (pleocytosis) and protein. A sample submitted to a research laboratory reveals that phosphorylated neurofilament heavy (pNF-H), which is an abundant structural protein of myelinated motor axons, is increased.

Comment

Degenerative myelopathy, a slowly progressing degeneration of myelin and axons, is suspected. However, there is no definitive antemortem diagnosis for degenerative myelopathy. Diagnosis is made based on clinical signs, ruling out other diseases, and a negative MRI. Some affected dogs have a defect in superoxide dismutase (SOD1), which normally helps control inflammation. The altered SOD1 can be screened for. Most recently, pNF-H has been identified as being increased in some dogs and may be an additional way to obtain a diagnosis antemortem. Affected dogs develop axonal degeneration or loss, as well as demyelination of the white matter in the cervical, thoracic, and lumbar spinal cord. They also have astrocytosis with proliferation of fibrous or reactive (gemistocytic) astrocytes. There appears to be preferential loss of large myelinated fibers. The mechanism underlying the degenerative and progressive changes is still being determined, but there is thought to be a role for oxidative stress.

Treatment

Most dogs are euthanized within a year of them becoming paraplegic. As the exact mechanism is not understood, treatment is supportive.

Vet-ebooks.com

Practice Questions

1. Which part of a neuron is primarily characterized as the information-receiving component?
- **a.** Axon
- **b.** Presynaptic terminal
- **c.** Cell body
- **d.** Dendrite
- **e.** Myelin

2. Which of the following is *not* characteristic of glial cells?
- **a.** Production of action potentials
- **b.** Immune responses of the nervous system
- **c.** Production of the myelin sheaths of axons
- **d.** Modulating the growth of developing or damaged neurons
- **e.** Buffering extracellular concentrations of some ions and neurotransmitters

3. The elements of spinal and cranial nerves that carry action potential commands from the CNS to synapses at skeletal muscles are:
- **a.** Axons of visceral efferent neurons
- **b.** Axons of somatic afferent neurons
- **c.** Axons of somatic efferent neurons
- **d.** The dorsal roots
- **e.** Axons of visceral afferent neurons

4. The thalamus and hypothalamus are components of which major brain division?
- **a.** Medulla
- **b.** Pons
- **c.** Midbrain
- **d.** Diencephalon
- **e.** Telencephalon

Bibliography

Allen NJ, Barres BA. Glia—more than just brain glue. *Nature*. 2009;457(7230):675–677.

Behan M. Organization of the nervous system. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Brosius Lutz A, Barres BA. Contrasting the glial response to axon injury in the central and peripheral nervous systems. *Dev Cell*. 2014;28(1):7–17.

Evans HE, de Lahunta A. *Miller's Anatomy of the Dog*. 4th ed. St. Louis: Saunders; 2013.

Greener M. Don't underestimate glial cells. *Prog Neurol Psychiatry*. 2015;19(1):5–8.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 5th ed. Sunderland, Mass: Sinauer; 2012.

Vallejo R, Tilley DM, Vogel L, et al. The role of glia and the immune system in the development and maintenance of neuropathic pain. *Pain Pract*. 2010;10(3):167–184.

Vet-ebooks.com

4

The Neuron

BRADLEY G. KLEIN

KEY POINTS

1. Neurons have four distinct anatomical regions.
2. Neuronal membranes contain a resting electrical membrane potential.
3. The resting membrane potential is the result of three major determinants.
4. The resting membrane potential can be changed by synaptic signals from a presynaptic cell.
5. Action potentials begin at the axon's initial segment and spread down the entire length of the axon.

There are two major classes of cells in the nervous system: the neuron and the glial cell (see Chapter 3). The neuron is the basic functional unit of the nervous system. The large number of neurons and their interconnections account for the complexity of the nervous system. The number of neurons in the vertebrate nervous system ranges greatly. There are approximately 100 million in a small mammal (e.g., mouse); 100 billion in a human; and more than 200 billion in whales and elephants: far more neurons in a nervous system than people on Earth. It is estimated that the glial cell to neuron ratio ranges from 1:1 to 10:1, and this figure varies across brain region. The structural and functional support provided to neurons by glial cells and their potential to modulate neural communication, development, pathology, and repair make an important contribution to the operational integrity of the nervous system. The numbers of cells in the nervous system are huge, but knowing that they have common elements makes it easier to understand them.

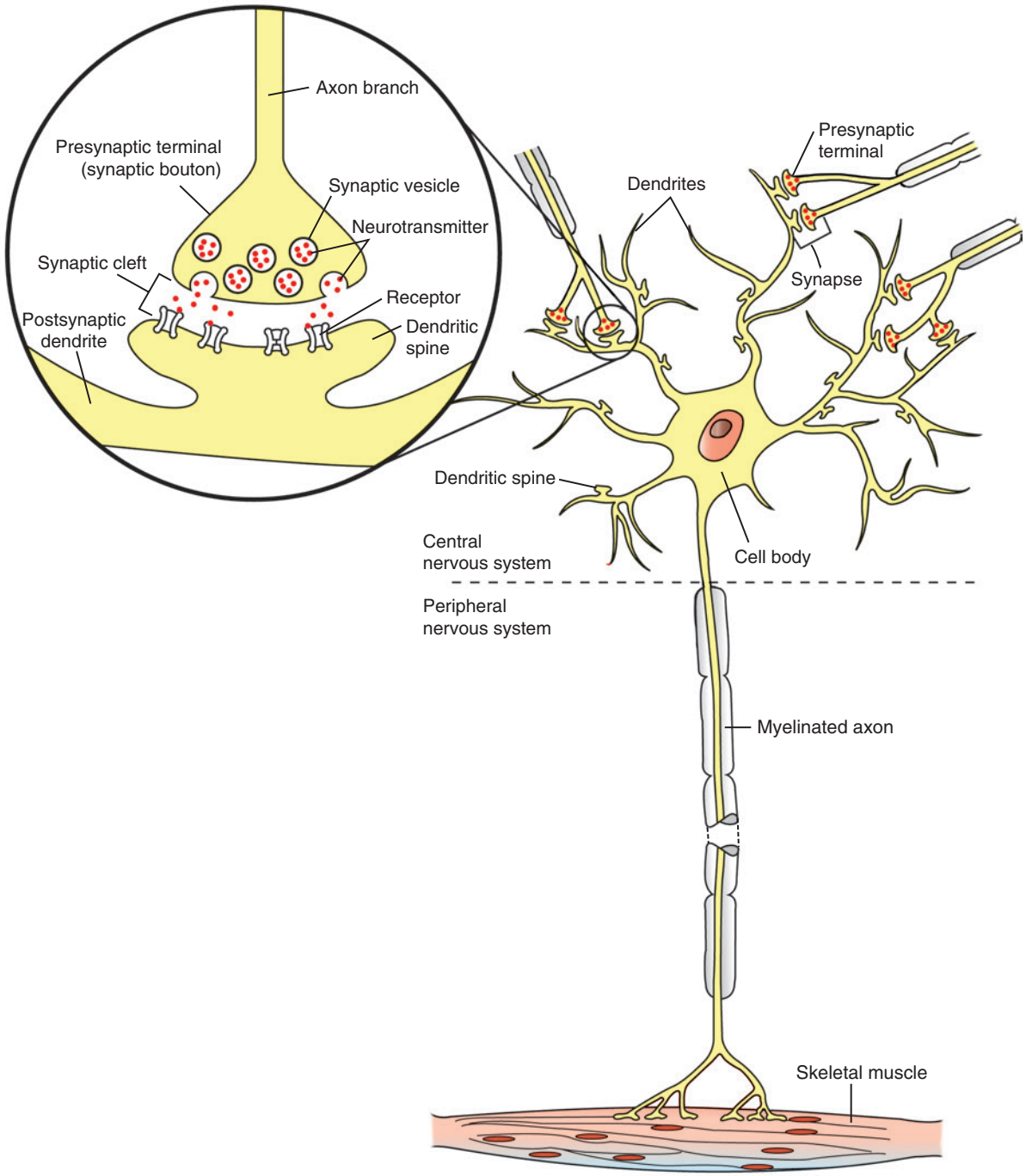
Neurons Have Four Distinct Anatomical Regions

A typical neuron has four morphologically defined regions (Fig. 4.1): the dendrites, the cell body, the axon, and the presynaptic terminals of the axon. These four anatomical regions are important in the major electrical and chemical responsibilities of neurons: receiving signals from the presynaptic terminals of other neurons (on dendrites), integrating these often-opposing signals (on the initial segment of the axon), transmitting action-potential impulses along the axon, and signaling an adjacent cell from the presynaptic terminal. These functions are collectively analogous to the general role of the nervous system: collecting information from the environment, integrating that information, and producing an output that can change the environment.

The *cell body* (also called the *soma* or *perikaryon*) plays a critical role in manufacturing proteins essential for neuronal function. Four organelles are particularly important for this purpose: the nucleus, containing the blueprint for protein assembly; the free ribosomes, which assemble cytosolic proteins; the rough endoplasmic reticulum, in which secretory and membrane proteins are assembled; and the Golgi apparatus, which further processes and sorts secretory and membrane components for transport. The cell body also contains organelles (e.g., lysosome, proteasome) for the degradation of used, damaged, or surplus proteins. The cell body usually gives rise to several branchlike extensions, called *dendrites,* whose surface area and extent greatly exceed those of the cell body. The dendrites serve as the major receptive apparatus of the neuron, receiving signals from other neurons. These signals, usually of a chemical nature, affect specialized receptor proteins or protein complexes (*receptors*) that reside on the dendrites. The cell body also gives rise to the *axon,* a tubular process that is often long (>1 m in some large animals). The axon is the conducting unit of the neuron, rapidly transmitting an electrical impulse (the action potential) from its initial segment at the cell body to its often distant end at the presynaptic terminal. Intact adult axons lack ribosomes and therefore normally cannot synthesize proteins. Instead, macromolecules are synthesized in the cell body and are carried along the axon to distant axonal regions and to the presynaptic terminals by a process called *axoplasmic transport.* Large axons are surrounded by a fatty, insulating coating called *myelin.* In the peripheral nervous system, myelin is formed by *Schwann cells,* specialized glial cells that wrap around the axon much like toilet paper wrapped around a broomstick. A similar function is performed by glial cells called *oligodendrocytes* in the central nervous system. The myelin sheath is interrupted at regular intervals by spaces called *nodes of Ranvier.* The myelin sheath significantly increases the speed of action-potential conduction along the axon.

Axons branch near their ends into several specialized endings called *presynaptic terminals* (or *synaptic boutons*). When the action potential rapidly arrives, these presynaptic terminals transmit a chemical signal to an adjacent cell. The site of contact of the presynaptic terminal with the adjacent cell is called the *synapse,* shown in the inset in Fig. 4.1. It is formed by the presynaptic terminal of one cell (presynaptic cell), the receptive surface of the adjacent cell (postsynaptic cell), and the space between these two cells (the *synaptic cleft*). Presynaptic terminals contain chemical transmitter–filled *synaptic vesicles* that can release their contents into the synaptic cleft. The presynaptic terminals of an axon usually contact the receptive surface of an adjacent neuron or muscle cell, usually on the neuron's dendrites, but sometimes this contact is made on the cell body or, occasionally, on the presynaptic terminals of another cell (e.g., for presynaptic inhibition). On many neurons, presynaptic terminals often synapse on small protrusions of the dendritic membrane called *dendritic spines* (Fig. 4.2 and see Chapter

Vet-ebooks.com

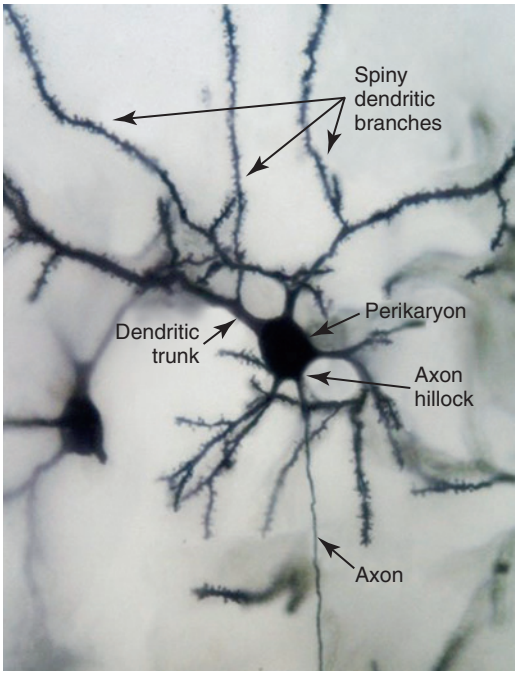


• **Fig. 4.1** A typical neuron has four functionally important regions. The cell body manufactures proteins to maintain the neuron; the dendrites receive signals from neighboring neurons; the axon integrates these signals and rapidly transmits action potentials along its length; and the presynaptic terminal signals adjacent cells. The inset shows an enlargement of the circled synapse.

5). The receptive surface of the postsynaptic cell contains specialized receptors for the chemical transmitter released from the presynaptic terminal. Collectively, the presynaptic terminals of a neuron often make synapses with more than one postsynaptic neuron.

The signaling functions of the morphological components of the neuron can be briefly summarized as follows (Fig. 4.3): receptors, usually dendritic, receive neurochemical signals released from the presynaptic terminals of many other neurons. These neurochemical signals, after being transduced by the receptors into a different form (small voltage changes), are integrated at the initial segment of the axon. Depending on the results of this integration, an action potential (large voltage change) may be generated on the axon. The action potential travels very rapidly to the axon's often distant presynaptic terminals to induce the release of chemical neurotransmitter onto another neuron or muscle cell.

Vet-ebooks.com

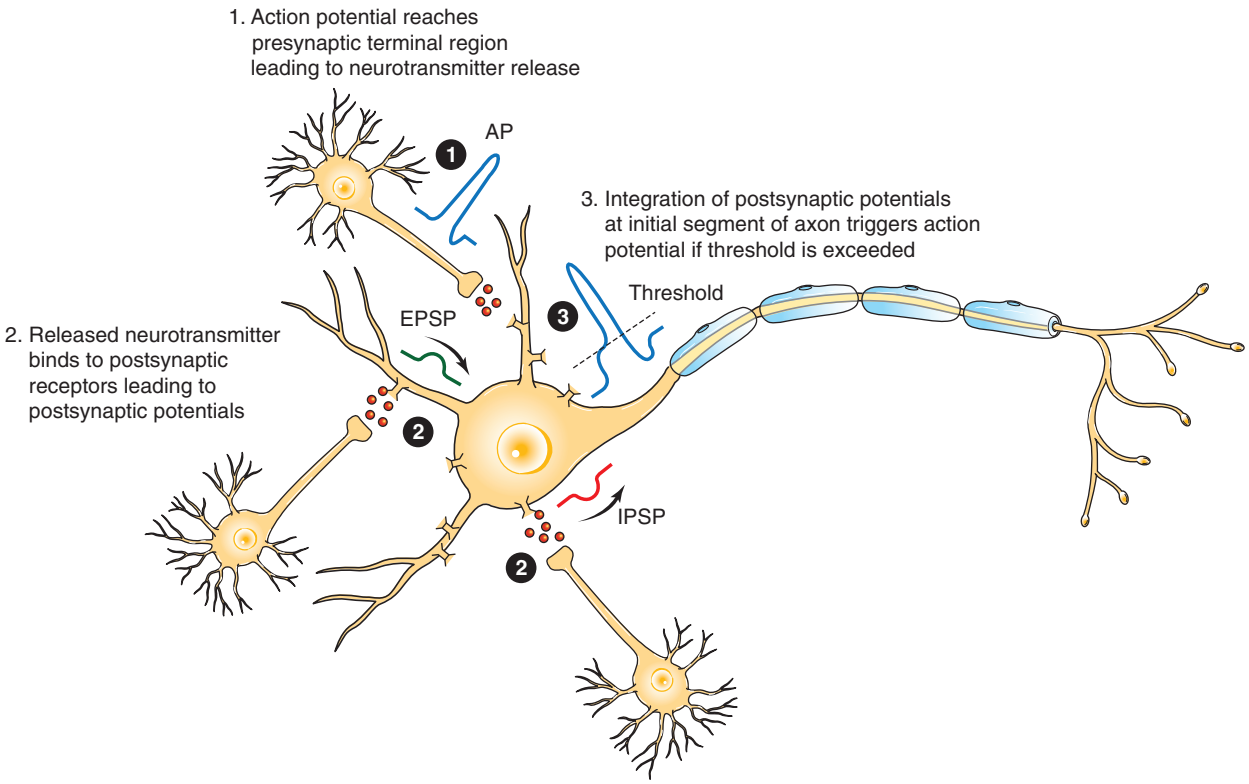


• **Fig. 4.2** Morphology of a neuron in mammalian cerebral cortex revealed with the Golgi staining method. The cell body (perikaryon), dendrites, and proximal portions of the axon are visible. Tiny dendritic spines can be seen along the dendrites. The cell body is approximately 20 µm in diameter. (Image courtesy Dr. Ceylan Isgor.)

Neuronal Membranes Contain a Resting Electrical Membrane Potential

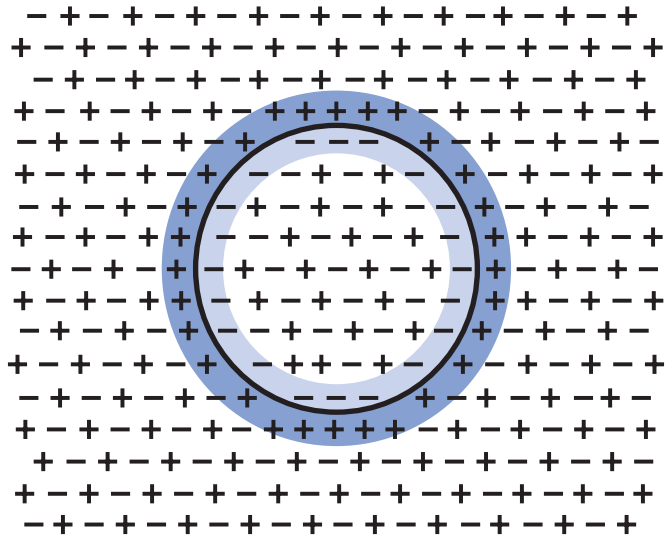
Neurons, like other cells of the body, have an electrical potential, or voltage, that can be measured across their cell membrane (*resting membrane potential*). However, the electrical membrane potential in neurons and muscle cells is distinct from virtually all other cells in that its magnitude and sign can be changed as the result of synaptic signaling from other cells, or it can change within a sensory organ receptor as a response to transduction of some environmental energy. When the change in membrane potential of a neuron or muscle cell reaches a threshold value, a further and dramatic change in the membrane potential, called an action potential, occurs; this action potential moves very rapidly along the entire length of the neuronal axon (see later discussion).

The origins of the resting electrical membrane potential are complicated, particularly in a quantitative way. In qualitative terms, however, the resting membrane potential is the result of the differential separation of charged ions, especially sodium (Na^+) and potassium (K^+), across the membrane and the resting membrane's differential permeability to these ions as they attempt to move back down their concentration and electrical gradients (see Chapter 1). Even though the net concentration of positive and negative charges is similar in both the intracellular and extracellular fluids, an excess of positive charges accumulates immediately outside the cell membrane, and an excess of negative charges accumulates immediately inside the cell membrane (Fig. 4.4). This makes the inside of the cell negatively charged with respect to the outside of



• **Fig. 4.3** Overview of neural communication. *AP,* Action potential; *EPSP,* excitatory postsynaptic potential; *IPSP,* inhibitory postsynaptic potential. (Portions modified from Klein BG. Membrane potentials: the generation and conduction of electrical signals in neurons. In: Reece WO, ed. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.)

Vet-ebooks.com



• **Fig. 4.4** Net concentrations of positive and negative charges are similar in both the intracellular space and the extracellular space. However, positive charges accumulate immediately outside the cell membrane (*blue*), and negative charges accumulate immediately inside the cell membrane (*lighter blue*), producing a potential difference (voltage) across the membrane.

the cell. The magnitude of the resulting electrical difference (or voltage) across the membrane varies from cell to cell, ranging from about 40 to 90 millivolts (mV), and is usually about 70 mV in mammalian neurons. Because the extracellular fluid is arbitrarily considered to be 0 mV the resting membrane potential is –70 mV, more negative on the inside of the membrane than on the outside.

The Resting Membrane Potential Is the Result of Three Major Determinants

Three major factors cause the resting membrane potential:

- *The Na^+, K^+ pump.* Cell membranes have an energy-dependent pump that pumps Na^+ ions out of the cell and draws K^+ ions into the cell against their concentration gradients. This maintains the differential distribution of each of these charged ion species across the membrane that underlies their ability to produce a voltage across the membrane. The pump itself makes a small, direct contribution to the resting membrane potential because it pushes three molecules of Na^+ out for every two molecules of K^+ drawn into the cell, thus concentrating positive charges outside the cell.
- *An ion species will move toward a dynamic equilibrium if it can flow across the membrane.* Using K^+ as an example, the concentration difference across the membrane actively maintained by the Na^+, K^+ pump produces a concentration gradient, or *chemical driving force,* that attempts to push the ion passively across the membrane from high concentration inside the cell toward low concentration outside. If K^+ can flow across ion channels in the membrane, exiting K^+ leaves behind unopposed negative charge (often from negatively charged protein macromolecules trapped inside the cell) that builds an electrical gradient, or *electrical driving force,* pulling K^+ back inside the cell. These opposing gradients eventually produce a dynamic equilibrium, even though there may still be more K^+ inside than outside, as well as a charge imbalance across the membrane. This uneven distribution of charge at dynamic equilibrium produces a voltage across the membrane called the *equilibrium potential* for that ion. When an ion species can flow across a channel in the membrane, it flows toward its equilibrium state and this drives the voltage across the membrane toward that ion's equilibrium potential.
- *Differential permeability of the membrane to diffusion of ions.* The resting membrane is much more permeable to K^+ than to Na^+ ions because there are vastly more K^+ leak channels than Na^+ leak channels in the membrane. This greater membrane permeability to K^+ means that K^+ ions can more closely approach their dynamic equilibrium state, and equilibrium potential, compared with Na^+ ions, which have difficulty moving across the membrane. Therefore, the equilibrium potential for the more permeant K^+ ions (about –90 mV in many mammalian neurons) will have the predominant influence on the value of the resting membrane potential compared with the equilibrium potential of the vastly less permeant Na^+ ions (about +70 mV in many mammalian neurons). Therefore, as noted earlier, the resting membrane potential of many mammalian neurons is about –70 mV, close to the equilibrium potential for K^+.

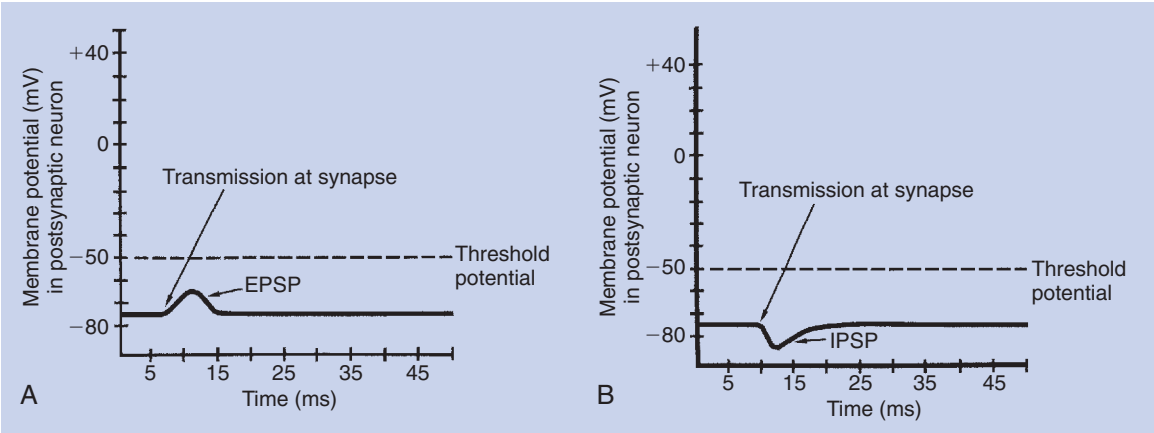
These three determinants—the Na^+, K^+ pump, the movement of a permeant ion toward dynamic equilibrium, and the differentially permeable membrane—are the primary source of the resting membrane potential. The value of this potential can be predicted by the Nernst and Goldman equations; refer to Chapter 1 and the Bibliography for a more quantitative understanding of the resting membrane potential.

This discussion of the resting membrane potential has a number of important clinical implications. The Na^+, K^+ pump requires energy in the form of adenosine triphosphate (ATP), which is derived from the intracellular metabolism of glucose and oxygen. In fact, it has been estimated that 50% to 70% of the brain's ATP-derived energy is expended on the pump. Because the neuron cannot store either glucose or oxygen, anything that deprives the nervous system of either substrate can lead to impairment of the pump and serious clinical neurological deficits. Fortunately, hormones and other factors normally maintain serum glucose and oxygen levels within narrow limits. Because Na^+ and K^+ are important ions involved in establishing the resting membrane potential, it is essential that serum levels of Na^+ and K^+ be regulated carefully. The endocrine system (Chapter 33) and kidney (Chapter 41) maintain serum Na^+ and K^+ levels within narrow limits. Anything altering serum levels of either ion beyond normal limits also leads to potentially severe neurological deficits.

The Resting Membrane Potential Can Be Changed by Synaptic Signals From a Presynaptic Cell

Although most cells of the body have a resting membrane potential, neurons and muscle cells are distinctive in that their membrane potential can be altered by a synaptic signal from another cell. Neurotransmitter released from a presynaptic axon terminal binds with receptors on the postsynaptic membrane, resulting in the opening or closing of ion selective channels and changing the membrane potential of the postsynaptic cell. Even though there are trillions of synapses in the nervous system, a presynaptic signal can alter the postsynaptic membrane potential in basically two ways: by making it more negative or more positive (less negative). The particular change depends on the nature of the receptor activated by the chemical transmitter that is released from the synaptic

Vet-ebooks.com



• **Fig. 4.5** Postsynaptic potentials. (A) Excitatory postsynaptic potential (EPSP) drives the membrane potential toward threshold. (B) Inhibitory postsynaptic potential (IPSP) drives the membrane potential away from threshold.

vesicles of the presynaptic axon terminal. The change in postsynaptic membrane potential is called a *postsynaptic potential.*

If a chemical synaptic transmission leads to a postsynaptic potential that is more positive in comparison with the resting level (e.g., from −75 to −65 mV), this is said to be an *excitatory postsynaptic potential* (EPSP) (Fig. 4.5A). It is called "excitatory" because it increases the chances that the threshold for triggering an action potential will be reached at the initial segment of the postsynaptic cell's axon. When an EPSP changes the postsynaptic membrane potential to a more positive value, the membrane is said to be *depolarized.* Depolarization of the postsynaptic membrane can result if the interaction of the chemical transmitter and its appropriate receptor on the postsynaptic membrane cause (ligand-gated) Na^+ channels to open. This allows Na^+ ions to diffuse into the neuron as they begin to flow toward equilibrium across the membrane, moving the membrane potential toward the more positive sodium equilibrium potential. The ion channels that usually change their conductivity as a result of neurotransmitter binding with a receptor are the *ligand-gated* or *chemically gated* ion channels (see Chapter 1).

Because the chemical transmitter is quickly removed from the synapse, the postsynaptic potential change is transient, lasting only a few milliseconds. Furthermore, because the change in ion flow resulting from receptor activation is limited, the magnitude of a postsynaptic potential is often quite small (e.g., 2–3 mV). However, it is greatest at the synapse. Although the depolarization spreads over the postsynaptic membrane, it decreases with the distance from the originating synapse, much as the waves created by throwing a stone into a lake decrease in size with the distance from where the stone fell.

If instead the presynaptic neurotransmitter's interaction with the postsynaptic receptor results in opening of the membrane's chemically gated K^+ channels, then K^+ ions diffuse out, moving the membrane potential even closer to the equilibrium potential for K^+ (−90 mV). This change from the resting potential to a more negative membrane potential is called *hyperpolarization.* Such hyperpolarization of the postsynaptic membrane is called an *inhibitory postsynaptic potential* (IPSP) (see Fig. 4.5B), because each such transmission makes it less likely that an action potential will result at the axon's initial segment. As with EPSPs, IPSPs spread over the neuron's membrane, and the hyperpolarization decreases with the distance from the originating synapse. It should be noted that only two of the receptor-mediated effects upon chemically gated ion channels, responsible for generating EPSPs or IPSPs, have been discussed here.

Action Potentials Begin at the Axon's Initial Segment and Spread Down the Entire Length of the Axon

Both EPSPs and IPSPs on the postsynaptic membrane are the subsequent result of action potentials that occurred on, and synaptic transmission from, many presynaptic cells. The integration of these postsynaptic potentials on the postsynaptic neuron is important for determining whether neurotransmitter will ultimately be released at the neuron's terminals. However, these postsynaptic potentials decrease in magnitude as they spread along the postsynaptic cell membrane. Because many neurons and muscle cells are long, the cell needs an efficient mechanism for sending a signal from its information-receiving end near the postsynaptic dendritic and soma membranes to the information-transmitting zone at the terminals of the often-lengthy axon to trigger neurotransmitter release. This is accomplished by an explosive event called an *action potential,* a regenerative electrical signal that begins at the axon's initial segment, is triggered by the integration of EPSP and IPSP membrane potential changes, and rapidly spreads down the length of the axon without decreasing in magnitude.

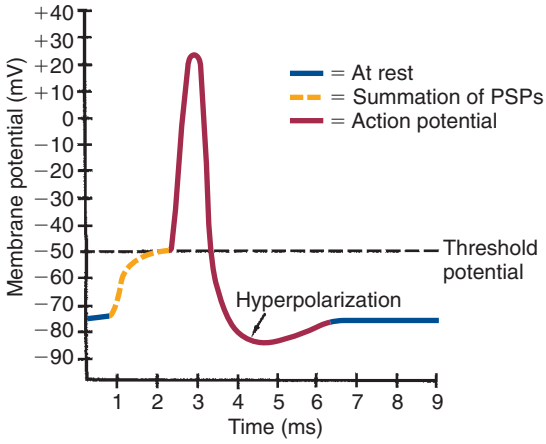
EPSPs and IPSPs can respectively summate on the postsynaptic membrane to produce larger changes in membrane potential than either signal alone. At the axon's initial segment, the arriving EPSPs and IPSPs are integrated. If only a few EPSPs arrive at the axon's initial segment, its membrane potential is not made sufficiently positive to reach its threshold potential (often 10–20 mV more positive than the resting potential) for triggering an action potential. However, if many more EPSPs than IPSPs arrive, the initial segment's membrane potential is made positive enough to reach its *threshold potential,* and an action potential is created on the axon. This action potential is the result of the sequential opening of voltage-gated ion channels in the membrane first to sodium and shortly thereafter to potassium.

The explosive changes in membrane potential that characterize the action potential can be described as follows: first, a dramatic and

Vet-ebooks.com

swift depolarization of the axonal membrane potential occurs, in which the inside of the cell actually becomes more positively charged than the outside, followed by a repolarization of the membrane, in which the membrane potential falls back toward the resting potential. The depolarization phase of the action potential is caused by the immediate and extensive opening of voltage-gated Na^+ channels and the consequent influx of Na^+ ions as they attempt to flow toward their equilibrium. This drives the membrane potential toward the Na^+ equilibrium potential, which is about +70 mV. As the action potential's depolarization phase continues, the voltage-gated Na^+ channels are spontaneously inactivated, and the voltage-gated K^+ channels, which open with a longer delay than the Na^+ channels, begin to allow even more K^+ ions to exit as they move closer to their equilibrium state. This brings depolarization to a halt and allows repolarization to occur as the membrane potential moves toward the K^+ equilibrium potential, which is about −90 mV. As repolarization continues, the membrane potential moves temporarily beyond its resting level to a hyperpolarized state. This hyperpolarization is attributable to the flow of K^+ ions out through the voltage-gated K^+ channels, in addition to the flow out through the K^+ leak channels, bringing the membrane potential even closer to the K^+ equilibrium potential than at rest. The membrane potential eventually returns to its resting state as the K^+ voltage-gated channels gradually close. The whole action potential takes about 2 to 3 ms in many neurons but longer in many muscle cells. Fig. 4.6 illustrates this sequence of events in a neuron.

An analogy may be helpful for understanding these difficult concepts. Imagine the resting neural membrane as a toilet. The toilet has stored potential energy by filling its water tank. (The neuron has done so by generating the resting membrane potential.) If the handle of the toilet is pushed down only briefly, for a short distance, some water runs into the toilet, but the flush cycle is not initiated. (This is similar to an EPSP without the action potential.) However, if the handle is pushed down far enough and held down long enough, a critical threshold is reached, the flush cycle is triggered, and it must run its course, including the refilling of the tank, before another flush cycle can be started. The action potential is analogous to this flush cycle. It is triggered once a critical depolarization threshold is reached. It usually must run its course, including reestablishing the resting membrane potential, before another action potential can be initiated. Because the flush cycle takes a finite amount of time, only a limited number of flush cycles could be completed in an hour, even if the toilet were flushed again each time the tank refilled. Similarly, because the action potential also has a finite duration, there is a limit to the number of action potentials per second that can be generated on an axon. (However, for both toilets and neurons, strategies can be employed to produce a flush or an action potential before the tank is completely refilled or before the membrane completely returns to the resting potential.)



• **Fig. 4.6** Axon's membrane potential changes dramatically during an action potential. After threshold is reached by summating postsynaptic potentials (PSPs), the axonal membrane depolarizes, repolarizes, hyperpolarizes, and then returns to its original resting potential. (Modified from Sherwood L. *Human Physiology: From Cells to Systems*. St. Paul: Wadsworth; 1989.)

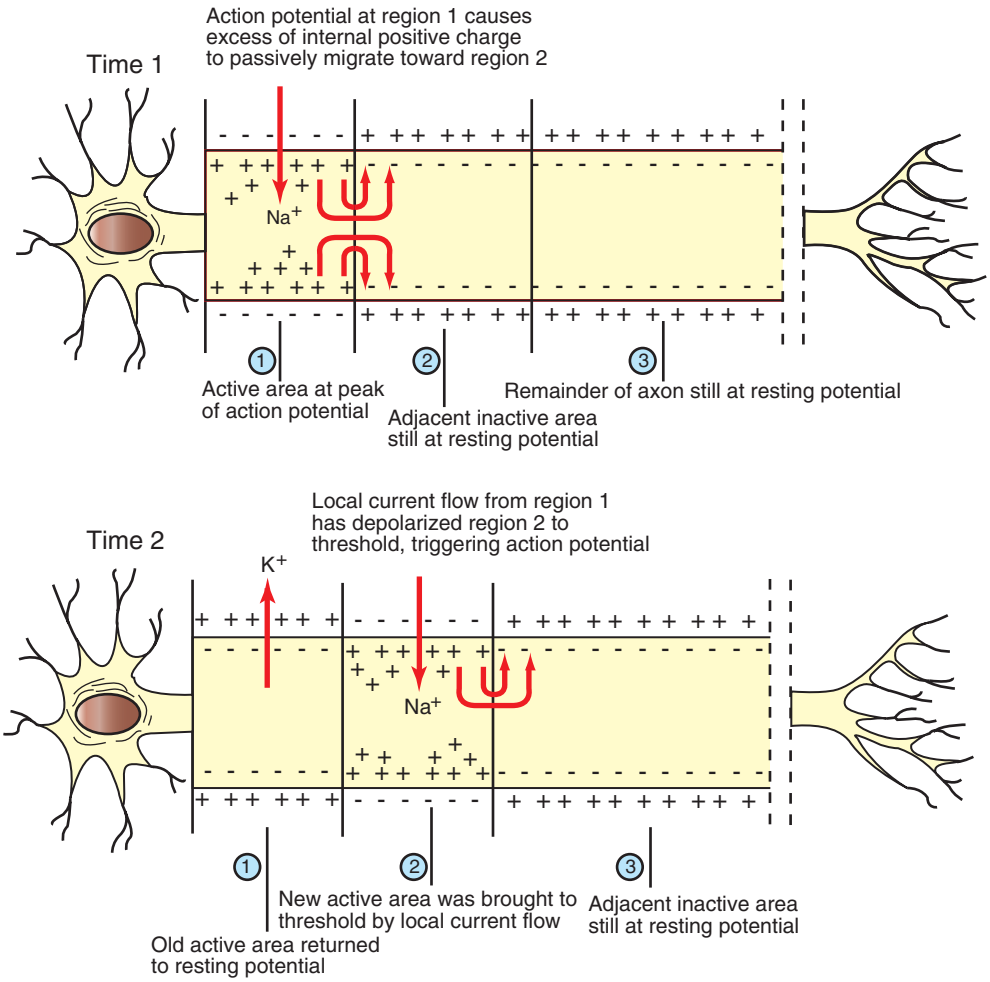
Certain animal toxins, such as tetrodotoxin from the Japanese puffer fish and a variety of peptides derived from spider venoms, can block voltage-gated Na^+ channels and therefore interfere with the generation of action potentials on axons. Many local anesthetics (e.g., lidocaine), which are used in a controlled, clinically efficacious manner, work by a related mechanism. Ongoing studies of animal venom components that can modify actions of specific ion channel components hold promise for even more refined clinical control of ion channel–mediated physiological processes.

The action potential actively propagates from its origin at the initial segment down the axon. The dramatic influx of Na^+ ions that accompanies action-potential depolarization of the initial segment's membrane results in the passive spread of these positive charges toward the adjacent resting segment of the membrane. This migration of positive charge on the inner surface of the membrane, called an *electrotonic current,* depolarizes this adjacent segment to threshold, causing voltage-gated Na^+ channels to open. This causes an action potential to develop, which in turn triggers a similar cycle in its adjacent membrane, and so on down the axon. In this way an action potential spreads from the axon's initial segment down to the presynaptic terminal at the axon's far end (Fig. 4.7).

The speed with which the action potential is conducted down the axon varies. The internal diameter and the degree of myelination of an axon play a critical role in determining this action-potential conduction velocity. In a small-diameter, unmyelinated axon, the conduction velocity is relatively slow (e.g., 0.5 meters/second [m/s]); conduction velocities of greater than 90 m/s (so that a distance as long as a football field is traveled in 1 second), however, are known to occur in large-diameter, heavily myelinated axons. This occurs because the *passive electrotonic current,* responsible for triggering the action potential at the next adjacent patch of axonal membrane, travels faster and farther along wider axons or along myelinated patches of axon. In myelinated axons, exchange of ions across the membrane, and thus generation of the action potential, can only occur at the bare nodes of Ranvier, where a high density of voltage-gated Na^+ channels are found. Given the rapid spread of electrotonic current along the myelinated patches (*internodes*) and the comparatively slower process of ion exchange at the nodes, the action potential seems to functionally jump from node to node (*saltatory conduction*) in myelinated axons (Fig. 4.8).

The normal facilitation of action-potential conduction velocity by myelin can be appreciated by considering diseases that attack myelin, such as acute idiopathic polyradiculoneuritis ("coonhound paralysis"). Slowing of evoked electrical signals along sensory and motor nerves and depressed spinal reflexes are associated with this condition. However, there is emerging evidence that myelin remodeling may occur as a normal, activity-dependent substrate of learning in mature animals.

Vet-ebooks.com



• **Fig. 4.7** Action potential, first generated in the axon's initial segment (time 1, region 1), moves down the unmyelinated axon as positive charges passively migrate to the immediately adjacent membrane to trigger an action potential there (time 2, region 2). (Redrawn from Sherwood L. *Human Physiology: From Cells to Systems*. St. Paul: Wadsworth; 1989.)

CLINICAL CORRELATIONS

Hypoglycemia

History

You examine an 8-year-old male boxer dog whose owner complains that the dog experiences seizures, weakness, and confusion around the time he is fed.

Clinical Examination

The findings of the dog's physical examination, including his neurological examination, were within normal limits. His fasting serum glucose level, however, was 29 mg/dL (normal is 70–110 mg/dL), and the ratio between serum insulin and serum glucose levels was significantly elevated.

Comment

Neurons depend primarily on oxygen and glucose as metabolites for ATP energy production, and neurons cannot store appreciable quantities of glucose. ATP is needed for maintenance of the normal electrical membrane potential. When deprived of glucose and subsequently ATP, the brain does not function properly; associated clinical signs include seizures, weakness, and confusion. In this animal, these signs were more common at the time of feeding because as the dog anticipated eating or actually did begin to eat, insulin was released, causing hypoglycemia.

In this case the ratio of insulin to glucose is elevated, probably because of an insulin-secreting tumor of the pancreas. Because insulin facilitates glucose transport through cell membranes, too much insulin results in the transfer of too much serum glucose to the cytoplasm of other cells of the body, thus depriving the brain's neurons of this essential metabolite.

Treatment

Insulinomas can usually be found and removed from the pancreas surgically. After surgical removal of the tumor, additional medical treatment is warranted to maintain normoglycemia. Medications include glucocorticoids, to stimulate gluconeogenesis; diazoxide, to inhibit insulin secretion; streptozocin, which is toxic to the beta cells; and somatostatin, which increases gluconeogenesis. With this tumor type, there is a high rate of metastasis, which means that other tumor sites may remain, in the liver and elsewhere, to overproduce insulin.

Salt Toxicity in Pot Belly Pig

History

A client calls you and says that they recently got a young pot belly pig through a friend. The pig was doing well for the first week or so, but now

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

seems to be acting "funny." The pig seems to be depressed and not as active, walking into things, uncoordinated, and not as responsive when they call her. She also seems like she is not eating or drinking as well, and she may have some loose feces. They have been feeding her dog food, as they have not had a chance to get to the feed store since they got her. They talked to the owner who they got the pig from, and the owner says that all the other pigs are normal.

Clinical Examination

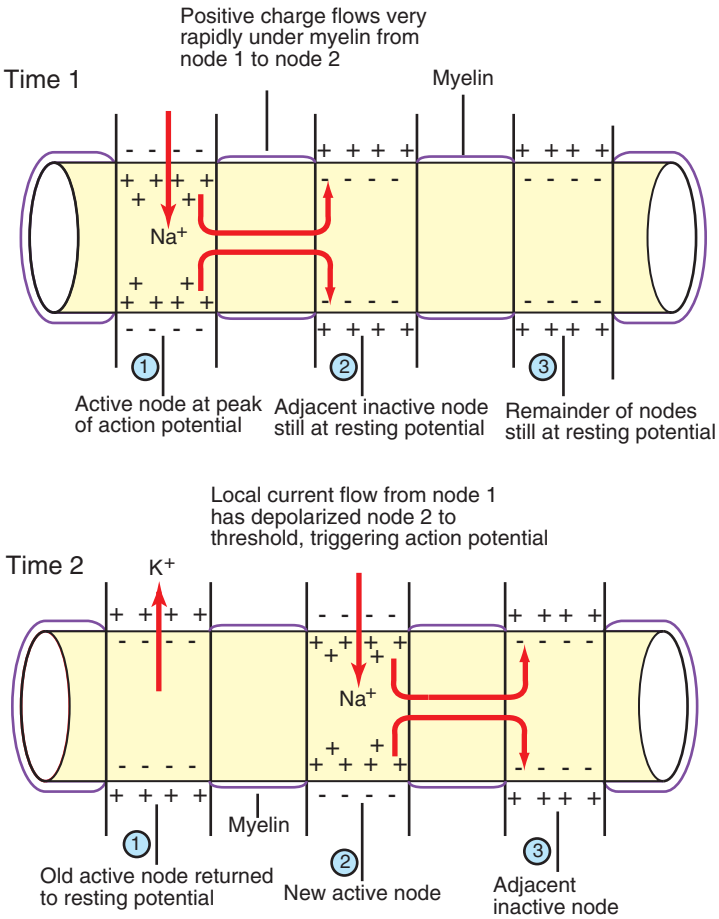
The pig appears depressed and is not responding normally. Her eyes appear sunken from dehydration, and her gastrointestinal sounds are increased. A brief neurological examination demonstrates depression, ataxia (incoordination) with both her front and hind limbs being equally affected, and blindness. You submit blood for complete blood count and biochemical profile. You also discuss with the owner the possibility to perform a cerebrospinal fluid (CSF) tap to collect a sample of the CSF for analysis to identify the cause of the clinical signs, if it cannot be determined based on the blood work.

Comment

The blood work demonstrates markedly increased levels of sodium and chloride (hypernatremia/hyperchloremia) as well as renal disease (increased blood urea nitrogen [BUN] and creatinine). This pot belly pig has salt toxicity due to the excessive amount of sodium contained in the dog food. The high levels of sodium ingested result in increased levels of sodium in the blood. The sodium in the blood passively diffuses into the CSF and brain. The increased sodium in the brain decreases energy-dependent transport mechanisms and anaerobic glycolysis, which normally function to remove the sodium. Increased sodium levels cause passive movement of fluid to equilibrate the electrolyte and fluid levels, thus causing swelling (edema) as well as inflammation.

Treatment

Animals must be treated with sodium-containing fluids, because decreasing the sodium levels too quickly can exacerbate edema in the brain. Prognosis is guarded.



• **Fig. 4.8** Saltatory conduction of action potentials in myelinated axons is faster than action-potential conduction in unmyelinated axons because the passive local current flows very rapidly under the myelin to trigger an action potential at the next node. Thus the action potential seems to jump functionally from node to node. (Modified from Sherwood L. *Human Physiology: From Cells to Systems*. St. Paul: Wadsworth; 1989.)

Vet-ebooks.com

Practice Questions

1. When treating critically ill patients with intravenous fluids, which two ions are most important to the neuronal membrane potential?
- **a.** Na^+ and Cl^-
- **b.** K^+ and Cl^-
- **c.** Ca^{2+} and Cl^-
- **d.** K^+ and Ca^{2+}
- **e.** Na^+ and K^+

2. The energy required by the Na^+, K^+ neural membrane pump is derived from ATP. In the neuron, this energy results from the nearly exclusive metabolism of oxygen and:
- **a.** Amino acids
- **b.** Fatty acids
- **c.** Glucose
- **d.** Glycogen
- **e.** Proteins

3. If the number of IPSPs on the dendritic membrane decreases while the number of EPSPs remains the same, what will happen to the action potentials on that neuron?
- **a.** Probability of triggering action potentials increases.
- **b.** Probability of triggering action potentials decreases.
- **c.** Probability of triggering action potentials remains unchanged.
- **d.** Action potentials would be eliminated.
- **e.** Action potentials would be conducted with increased velocity.

4. During an excitatory postsynaptic potential in a neural membrane, which of the following is the most important ion flow?
- **a.** Sodium ions diffuse out of the cell.
- **b.** Sodium ions diffuse into the cell.
- **c.** Potassium ions diffuse out of the cell.
- **d.** Potassium ions pumped in by the Na^+, K^+ pump.
- **e.** None of the above.

5. Choose the *incorrect* statement below:
- **a.** Conduction velocity of action potentials is slower in myelinated than in unmyelinated axons.
- **b.** Conduction velocity of action potentials is faster in myelinated than in unmyelinated axons.
- **c.** In saltatory conduction of action potentials, the action potential seems to jump functionally from node to node (nodes of Ranvier).
- **d.** Action potentials are of equal magnitude at the beginning and at the end of an axon.

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Fields RD. A new mechanism of nervous system plasticity: activity-dependent myelination. *Nat Rev Neurosci*. 2015;16(12):756–767.

Garrett LD. *Insulinomas: a review and what's new*. *Proceedings ACVIM* 2003.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Herculano-Houzel S. The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution. *Glia*. 2014;62:1377–1391.

Klein BG. Membrane potentials: the generation and conduction of electrical signals in neurons. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Klint JK. Spider-venom peptides that target voltage-gated sodium channels: pharmacological tools and potential therapeutic leads. *Toxicon*. 2012;60:478–491.

Smith MO, George LW. Diseases of the nervous system. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St. Louis: Mosby; 2014.

Utkin YN. Animal venom studies: current benefits and future developments. *World J Biol Chem*. 2015;6(2):28–33.

Vet-ebooks.com

5 The Synapse

BRADLEY G. KLEIN

KEY POINTS

1. The anatomy of the neuromuscular junction is specialized for one-way synaptic communication.
2. An action potential on the presynaptic neuron triggers an action potential on the muscle cell through the release of acetylcholine.
3. There is greater variety in the specifics of neuron-to-neuron synaptic transmission than in transmission at the neuromuscular junction.

Neurons communicate with each other and with other cells of the body, such as muscle and secretory cells. In Chapter 4, the generation of the action potential and its rapid conduction down the axon to the presynaptic terminal was discussed. Using these processes, the neuron can rapidly notify its presynaptic terminals, often located far from its cell body, to initiate the transfer of information to other cells. Such communication occurs rapidly between cells and often focally at specialized junctions called *synapses* (Greek for "junction" or "to bind tightly"). Synaptic transmission between cells can be either electrical or chemical. At *electrical synapses,* ionic current flows directly between cells as the mediator for signal transmission. Electrical synapses in the mammalian nervous system appear to be more widespread than originally thought, and the specific locations, mechanisms, and functional role of these synapses continue to be clarified. Much more is known about synaptic transmission mediated by a *chemical messenger.* Released from the presynaptic terminals on arrival of the action potential, this chemical messenger rapidly diffuses to the postsynaptic cell membrane where it binds with receptors. This binding initiates a postsynaptic change in function, often generating a postsynaptic potential.

The best understood chemical synapse is that between a motor neuron and a skeletal muscle cell (fiber). The specialized area along the muscle cell where this communication takes place is known as the *neuromuscular junction* (Fig. 5.1). Given the emphasis on posture and locomotion in veterinary neurological diagnosis, synaptic communication at the neuromuscular junction is the focus of this chapter. Although the nomenclature for the neuromuscular junction among different authors can be confusing, synaptic communication at the neuromuscular junction is fundamentally similar to that between neurons. However, there is greater variety in the specifics of neuron-to-neuron synaptic transmission, as also discussed.

The Anatomy of the Neuromuscular Junction Is Specialized for One-Way Synaptic Communication

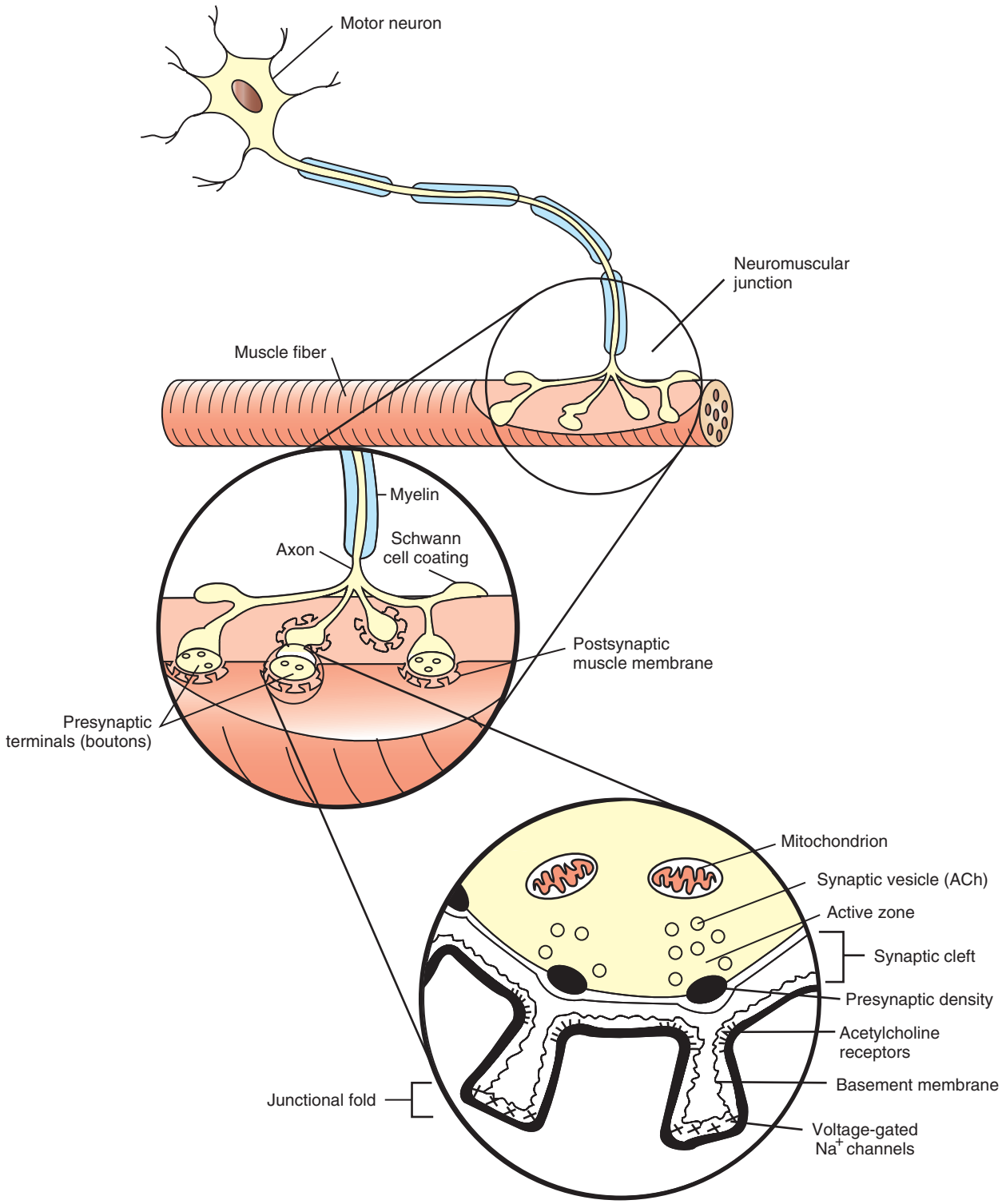
Motor neurons that synapse on skeletal muscles have their cell bodies located within the central nervous system (CNS), either within the spinal cord or the brainstem. The axons of these motor neurons travel within peripheral nerves, out to the muscle, where each motor neuron synapses on several individual fibers (cells) of the muscle. However, an individual skeletal muscle fiber receives synaptic input from, and therefore its contraction is controlled by only one motor neuron.

The neuromuscular junction, like most chemical synapses, has (1) *presynaptic* elements; (2) postsynaptic elements; and (3) a narrow space between these communicating portions of the neuron and muscle fiber, called the *synaptic cleft* (see Fig. 5.1). The presynaptic side of the neuromuscular junction is made up of a spatially focused group of terminals (transmitting portion) of the motor neuron. These presynaptic terminals each have a swelled, button-like appearance and are also called *synaptic boutons.* Each terminal (or synaptic bouton) contains a large number of membranous storage vesicles, called *synaptic vesicles,* which contain the chemical neurotransmitter substance, in this case *acetylcholine.* These synaptic vesicles are lined up in rows along the inner surface of the terminal membrane (Fig. 5.2). The presynaptic membrane region associated with each double row of vesicles is called an *active zone* and is the site where the synaptic vesicles will eventually release acetylcholine into the synaptic cleft. The presynaptic terminal also contains mitochondria, an indication of active metabolism in the cytoplasm. Some mitochondrial products (e.g., acetyl-CoA, adenosine triphosphate [ATP]) play a role in the local synthesis of acetylcholine and in its movement into the synaptic vesicles.

In the region where a presynaptic terminal (neural) of the neuromuscular junction communicates with a locus on the postsynaptic (muscle) cell, the membranes are separated by a narrow space, the synaptic cleft, which is about 50 nm wide (see Figs. 5.1 and 5.2). The cleft contains extracellular fluid and a basal lamina, composed of a matrix of molecules, that is a specialized region of the muscle basement membrane. Some of these matrix molecules mediate synaptic adhesion between neuron and muscle.

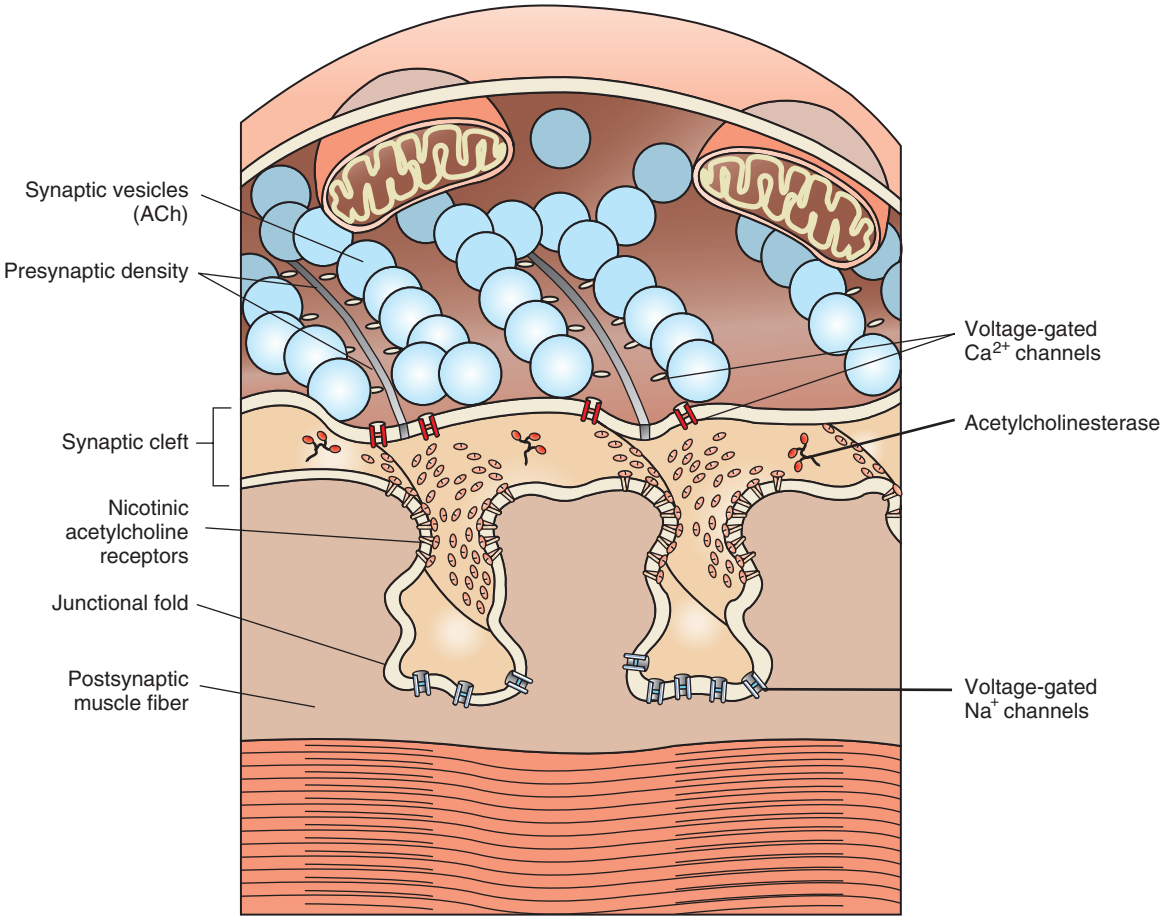
The postsynaptic muscle cell membrane has several specialized features that facilitate synaptic transmission at the neuromuscular junction. Directly opposite the face of a presynaptic terminal, the postsynaptic muscle cell membrane contains receptors for the acetylcholine transmitter (see Figs. 5.1 and 5.2). In this complementary region, the membrane has a series of invaginations, called *junctional folds,* that increase the surface area where acetylcholine receptors can reside. The acetylcholine receptors are most densely

Vet-ebooks.com



• **Fig. 5.1** The region of synaptic communication between a motor neuron and a skeletal muscle cell (fiber) is the neuromuscular junction. The spatially focused group of motor neuron terminals (boutons), which contain acetylcholine (ACh)-filled vesicles, represents the presynaptic elements; the aggregates of junctional folds of the muscle fiber, which contain acetylcholine receptors, represent the postsynaptic elements; the narrow space between the presynaptic motor neuron terminals and the postsynaptic junctional folds of the muscle fiber is the synaptic cleft.

Vet-ebooks.com



• **Fig. 5.2** Presynaptic acetylcholine-filled synaptic vesicles are lined up, near voltage-gated Ca^{2+} channels, at active zones. Released acetylcholine binds with nicotinic acetylcholine receptors at junctional folds on the postsynaptic muscle fiber membrane. (Redrawn and modified from Bear MF, Connors BW, Paradiso MA: *Neuroscience: Exploring the Brain.* 3rd ed. Philadelphia: Lippincott, Williams & Wilkins; 2007.)

packed at the mouth of these junctional folds, and these mouths are closely aligned with the active zones of the presynaptic terminals from which the acetylcholine is released. Thus there is a good match between the specific regions of transmitter release from the motor neuron and the specific location of the receptors on the muscle fiber. Collectively, the regions of the muscle fiber that contain these aggregates of junctional folds and acetylcholine receptors are often referred to as the *motor end plate* of the neuromuscular junction. Because the neurotransmitter is found only on the presynaptic neural side at the neuromuscular junction, transmission can go only from neuron to muscle, not in the reverse direction.

As noted, neurotransmitter signaling across the neuromuscular junction, for purposes of activating muscle fiber contraction, favors the nerve-to-muscle direction. However, there is some evidence that other types of molecules, in the muscle, may play a role during development in the survival, differentiation, and normal functioning of the presynaptic motor neuron terminals.

An Action Potential on the Presynaptic Neuron Triggers an Action Potential on the Muscle Cell Through the Release of Acetylcholine

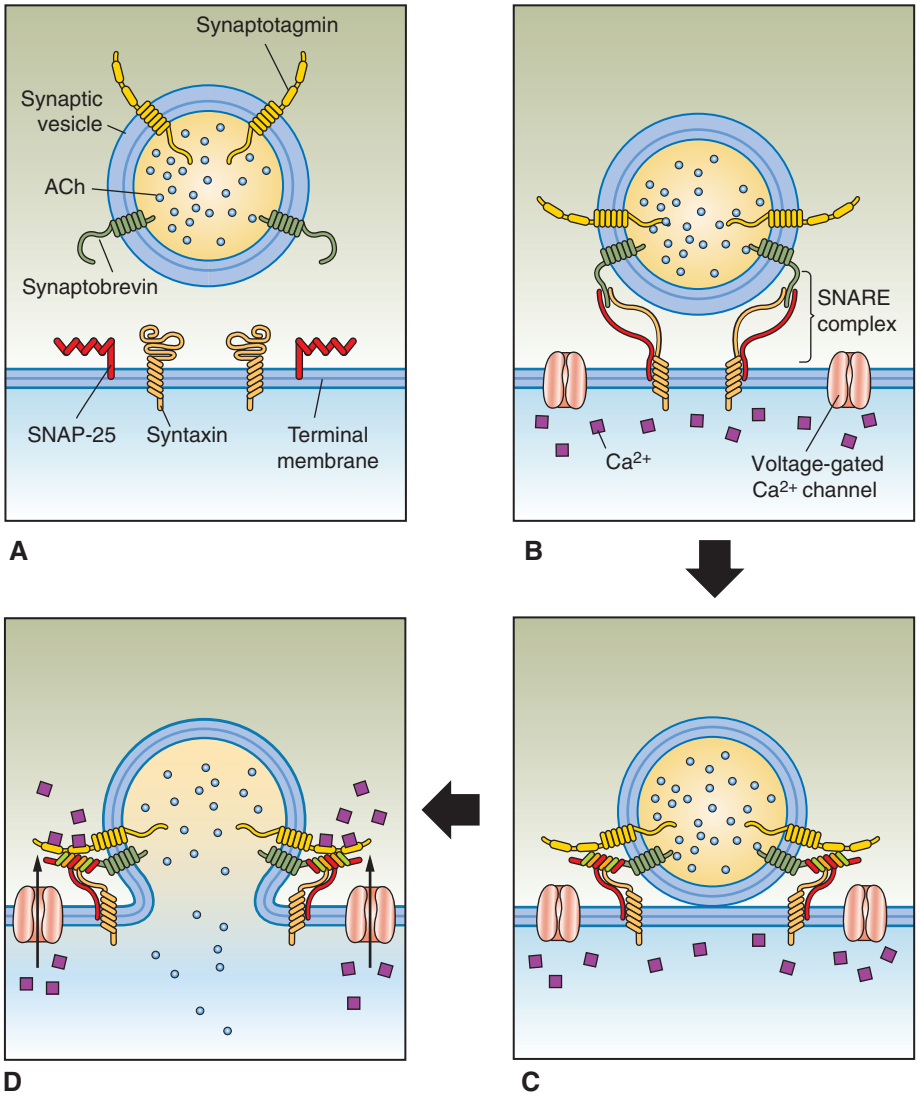
The function of the neuromuscular junction is to transmit a chemical message unidirectionally between a motor neuron and a skeletal muscle cell (fiber) with a frequency established by the CNS. The arrival of an action potential at the motor neuron terminal triggers the release of the acetylcholine transmitter, which then binds with acetylcholine receptors on the postsynaptic muscle fiber membrane. This leads to the genesis of an action potential along the muscle fiber membrane that ultimately leads to contraction of the fiber (see Chapter 6).

An action potential on a motor neuron arises at its initial axon segment and then travels along the entire axon, eventually arriving at the presynaptic terminal (see Chapter 4). As previously noted, the exchange of Na^{+} and K^{+} ions across axonal voltage-gated Na^{+} and K^{+} channels is responsible for the generation of the action potential and its conduction to the terminal. However, as the action potential arrives at the presynaptic membrane, the wave of depolarization opens voltage-gated Ca^{2+} channels located in this region (see Fig. 5.2); as Ca^{2+} flows toward equilibrium across the membrane, the Ca^{2+} enters the presynaptic terminal. This increase in the intracellular Ca^{2+} level is critical for the release of neurotransmitter from the terminal.

Recall that the acetylcholine-containing synaptic vesicles are lined up at the active zones of the presynaptic terminal. They *dock* there through the assembly of a complex of soluble proteins that will bind to both the vesicle and to voltage-gated Ca^{2+} channels of the terminal membrane, recruiting these channels toward the vesicle and the active zone. A portion of the assembled soluble

Vet-ebooks.com

protein complex will also interact with a complement of binding proteins that respectively reside on the vesicle membrane (synaptobrevin) and on the inner surface of the terminal membrane (syntaxin and SNAP-25), setting the stage for the intertwining of these binding proteins (Fig. 5.3). The intertwining, or zipping together, of these complementary binding proteins on the vesicle and terminal membranes is part of a *priming* process that holds the vesicle close to the terminal membrane and the previously recruited voltage-gated Ca^{2+} channels, preparing the vesicle for very rapid Ca^{2+}-induced neurotransmitter release. In the latter stage of this priming process, an additional protein on the synaptic vesicle (synaptotagmin), which is capable of binding with Ca^{2+} and thus serves as a Ca^{2+}sensor, associates with the intertwined binding proteins (the SNARE complex) of the vesicle and terminal membranes. When the action potential induces Ca^{2+} flow into the terminal through the voltage-gated Ca^{2+} channels, the ions bind with the Ca^{2+} sensor molecule (synaptotagmin). This very rapidly triggers fusion of the vesicle with the presynaptic terminal membrane, opening of the vesicle, and release of acetylcholine into the synaptic cleft. After transmitter release, the vesicle membrane is retrieved back into the presynaptic terminal and can be recycled to reform a vesicle that is then refilled with acetylcholine synthesized



• **Fig. 5.3** Release of acetylcholine (ACh) from a synaptic vesicle at the active zone of a motor neuron terminal. (A) Synaptobrevin is a binding protein on the synaptic vesicle. SNAP-25 and syntaxin are binding proteins at the terminal membrane. Synaptotagmin is a vesicle protein that is a calcium (Ca^{2+}) sensor molecule. (B) The binding proteins on the vesicle membrane intertwine with the binding proteins on the inside of the terminal membrane, forming a SNARE complex. This is part of a priming process that holds the vesicle and terminal membranes in close proximity, preparing the vesicle for very rapid Ca^{2+}-induced neurotransmitter release (the prior docking phase is not pictured). (C) As the priming process continues, the vesicle and terminal membranes are drawn closer and tighter together, and the vesicular Ca^{2+} sensor molecule (synaptotagmin) associates with the SNARE complex. (D) Depolarization of the terminal membrane, as a result of the action potential, opens voltage-gated Ca^{2+} channels in the terminal membrane. Ca^{2+} enters the terminal and binds with synaptotagmin. This binding results in extremely rapid fusion of the vesicle and terminal membranes and release of ACh into the synaptic cleft by exocytosis. (Modified from Boron WF, Boulpaep EL. *Medical Physiology.* 2nd ed. Philadelphia: Saunders; 2009.)

Vet-ebooks.com

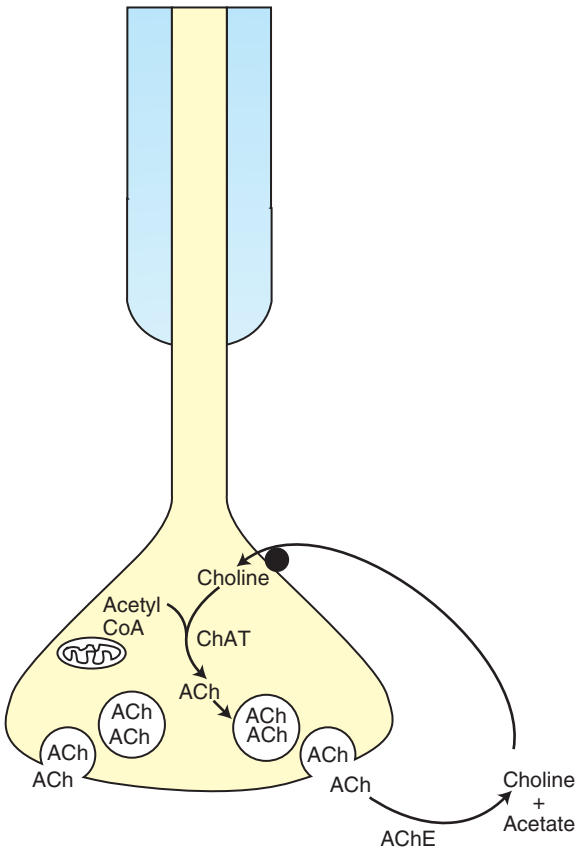
in the cytoplasm. Certain bacterial toxins (e.g., botulinum, tetanus) can destroy the binding proteins of the SNARE complex that are important elements of vesicle priming, ultimately interfering with the ability of the vesicle to release its contents into the synaptic cleft.

After release, acetylcholine then diffuses across the synaptic cleft and binds with transmitter-specific receptors, the *nicotinic acetylcholine receptors,* in the postsynaptic muscle membrane. This specific type of acetylcholine receptor, found at the neuromuscular junction, is so named because it can also bind the alkaloid drug nicotine. There are subtypes of the nicotinic acetylcholine receptor and not all are found on skeletal muscle. Some reside on particular neurons of the central and peripheral nervous systems. The nicotinic acetylcholine receptor is actually a ligand-gated ion channel (see Chapter 1), permeable to small cations, with two binding sites for the acetylcholine molecule. As acetylcholine binds at these two loci, the channel opens and, among other ionic movements, Na^+ ions diffuse into the muscle cell as they attempt to flow toward equilibrium. This contributes to a depolarization of the postsynaptic muscle cell membrane analogous to an excitatory postsynaptic potential (EPSP). However, at the neuromuscular junction, the unitary postsynaptic potential produced, called an *end-plate potential,* is sufficient to open enough voltage-gated Na^+ channels, deep within the junctional folds, to generate an action potential on the muscle cell membrane.

Acetylcholine binds with its receptor only briefly (~1 ms). When free, it is destroyed by the enzyme *acetylcholinesterase.* This enzyme, anchored to the basal lamina of the synaptic cleft, inactivates acetylcholine by breaking it down into acetate and choline molecules (Fig. 5.4). The choline, a precursor of acetylcholine synthesis, can then be transported back into the presynaptic terminal by a high-affinity transporter protein in the terminal membrane and recycled in acetylcholine synthesis. Chemicals that inhibit acetylcholinesterase, such as some organophosphate insecticides (e.g., malathion, chlorpyrifos) and nerve gases (e.g., sarin), can abnormally prolong the presence of acetylcholine at the synapse, often with disastrous physiological consequences. Because the neurotransmitter is normally destroyed soon after its binding with the muscle membrane receptor, and because more transmitter is not available in the synapse to attach to the receptors in sufficient quantities until another motor neuron action potential occurs, there is approximately a 1 : 1 ratio between action potentials on the neuronal and muscle cell membranes.

There Is Greater Variety in the Specifics of Neuron-to-Neuron Synaptic Transmission Than in Transmission at the Neuromuscular Junction

As mentioned earlier, some noteworthy differences exist between synaptic transmission at the neuromuscular junction and neuron-to-neuron synaptic transmission. Although acetylcholine is the neurotransmitter responsible for the primary postsynaptic effect at the neuromuscular junction, a considerable variety of neurotransmitters, in addition to acetylcholine, can be used to produce the principal postsynaptic effect at neuron-to-neuron synapses (Box 5.1). Furthermore, not all of these transmitters are released from morphologically distinct active zones, although their release from the terminal can still depend on Ca^{2+} influx. In such cases, release from the terminal may not always occur directly at the synaptic cleft, resulting in a wider postsynaptic distribution of transmitter. Interestingly, some molecules that are often referred to as *atypical* or *nontraditional neurotransmitters* (e.g., endocannabinoids, nitric



• **Fig. 5.4** Synaptic inactivation of acetylcholine *(ACh).* Released ACh is enzymatically broken down into choline and acetate by synaptic acetylcholinesterase. The choline is actively transported back into the terminal and can be reused in ACh synthesis. *ChAT,* Choline acetyltransferase.

• BOX 5.1 Members of the Major Neurotransmitter Classes

Amino Acids
- Glutamate
- Glycine
- γ-Aminobutyric acid (GABA)

Amines
- Acetylcholine
- Serotonin
- Histamine

Catecholamines
- Dopamine
- Norepinephrine
- Epinephrine

Peptides[a]
- Substance P
- Vasopressin
- Somatostatin

Opioids
- Endorphins
- Enkephalins
- Dynorphins
- Endomorphins

Purines
- Adenosine
- Adenosine triphosphate (ATP)

Atypical (Nontraditional) Gases
- Nitric oxide
- Carbon monoxide
- Hydrogen sulfide

Endogenous Cannabinoids (Endocannabinoids)
- Anandamide
- 2-Arachidonylglycerol

[a]*Only a partial list of peptide neurotransmitters.*

Vet-ebooks.com

oxide) are actually produced in a postsynaptic neuron following traditional synaptic transmission but then diffuse back across the postsynaptic cell's lipid membrane and the synaptic cleft to affect the function of the presynaptic terminal. Therefore neuron-to-neuron communication may not be as specialized for one-way communication as the neuromuscular junction.

The postsynaptic membrane of a neuron-to-neuron synapse can be the soma, dendrites, or even the terminals of the postsynaptic neuron, and junctional folds are not seen at these synapses. However, the dendritic postsynaptic membrane often possesses small protrusions called *dendritic spines* (see Chapter 4). As with the junctional folds on muscle cells, these spines increase the surface area of the postsynaptic membrane and, because of their narrow necks, are thought to provide a means for biochemical isolation between nearby synapses. Spines are also thought to act as collectors of, and reservoirs for, the myriad molecules involved in postsynaptic transduction of neurotransmitter signals. Polyribosomes can sometimes be found at the base of spines suggesting a role in local protein synthesis triggered by synaptic transmission. Furthermore, spines can change size and shape over an animal's lifetime, modulating the functional effectiveness of the synapse and implicating spines in the learning and memory processes. Whereas transmitter release at the neuromuscular junction always produces postsynaptic excitation (membrane depolarization), release at synapses between neurons can produce excitation or inhibition (membrane hyperpolarization). However, synapses on dendritic spines are almost always excitatory.

At the neuromuscular junction, the postsynaptic receptor is almost exclusively the nicotinic acetylcholine receptor, a ligand-gated ion channel. At synapses between neurons, a much greater variety of receptors is available. These may differ from the nicotinic acetylcholine receptor with respect not only to the binding transmitter, but to the receptor mechanism as well (e.g., G-protein coupled; see Chapter 1). Also, several different types of neurotransmitter receptor are often found on a single neuron.

When transmitters other than acetylcholine are employed at neuron-to-neuron synapses, depending on the transmitter, the termination of action of that transmitter may be accomplished by transporter-mediated reuptake of the transmitter itself into the neuron of release or by a less specific, and somewhat slower, form of enzymatic degradation than with acetylcholinesterase. Although simple diffusion of neurotransmitter away from the synapse contributes to the termination of action of most neurotransmitters to some degree, this mode may play a more important role for some neurotransmitters than others. For some atypical gaseous neurotransmitters, rapid spontaneous decay appears to be important for termination of action. Finally, at neuron-to-neuron synapses, a single action potential on a presynaptic neuron rarely results in a full-blown action potential on the postsynaptic neuron. Some form of summation of presynaptic inputs is required to generate a postsynaptic action potential.

As discussed in Chapter 6, action potentials on the muscle cell membrane lead to contraction, or mechanical shortening, of the muscle cell. When this contraction is combined with the shortening of many muscle cells, movement of the body occurs.

CLINICAL CORRELATIONS

Myasthenia Gravis

History

You examine a 5-year-old female German shepherd whose owner states that the dog becomes progressively weaker with exercise. The owner also states that recently, just after eating, the dog has begun to vomit food in formed, cylinder-shaped boluses.

Clinical Examination

All abnormalities found on physical examination were referable to the neuromuscular system. After resting, the dog's neurological examination findings were within normal limits. With even moderate exercise, however, the dog became progressively weaker, particularly in the front legs. Intravenous injection of an acetylcholinesterase inhibitor, edrophonium (Tensilon), eliminated all clinical signs of weakness. Radiographs of the chest revealed an enlarged esophagus and thymus.

Comment

The history of an enlarged esophagus (megaesophagus) and the response to an acetylcholinesterase inhibitor confirm the diagnosis of *myasthenia gravis* (grave muscle weakness). This is caused by a failure of transmission of acetylcholine at the neuromuscular synapse. This transmission failure is caused by antibodies produced by the body against its own acetylcholine receptors. The abnormal antibodies bind with the receptors to form complexes, which prevent acetylcholine from binding to the acetylcholine receptors. As a result, no depolarization occurs on the postsynaptic membrane of the muscle cells. Antibodies also alter the junctional folds and number of acetylcholine receptors available to bind with the transmitter. Acetylcholinesterase inhibitors prevent the metabolism of acetylcholine, allowing acetylcholine to remain longer at the synapse, with additional time for binding to the receptors, and thus facilitating normal transmission.

The large amount of skeletal muscle in the dog's esophagus explains its enlargement from paralysis. These patients often regurgitate formed boluses of food shortly after eating.

Myasthenia gravis can be associated with mediastinal masses, usually of the thymus. The autoantibodies that the body makes are often against antigens from the thymus or acetylcholine receptors. In addition to this cause of myasthenia gravis, idiopathic (of unknown cause) myasthenia gravis is also common.

Treatment

Spontaneous remissions are common, depending on the cause. Until then, oral daily acetylcholinesterase inhibitors are given. Surgical removal of mediastinal masses may also be necessary.

Tetanus

History

A client calls you because their 6-year-old Quarter Horse gelding stepped on a nail a few days ago on the right front foot. They pulled the nail out and soaked the foot, but now the horse seems depressed and does not want to move.

Clinical Examination

The horse has an increased temperature (fever) with increased heart rate (tachycardic), and increased pulses of the digital arteries to the right front foot. The leg is warm, swollen, and painful. The horse appears to be very sensitive to any stimulation. The horse is overdue for his vaccinations.

Comment

Exposure to *Clostridium tetani* can occur through cuts/wounds. The bacterium has a neurotoxin, tetanospasmin, that blocks the synaptic release of glycine and gamma-aminobutyric acid (GABA) by cleaving the synaptic

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

vesicle binding protein synaptobrevin. Blocking the release of these transmitters, which normally have an inhibitory effect on motor neurons that innervate skeletal muscles, results in abnormal excitation of these motor neurons. This results in continued muscle stimulation, manifested as hypertonia and muscle spasms. Affected animals can develop a stiff head and neck as well as stiff gait. When they become recumbent, they can develop a sawhorse posture. Muscle spasms are easily elicited, including jaw and lip rigidity, as well as retraction of the eye with presence of the third eyelid. The respiratory muscles are often affected. Because these muscles do not function properly, animals can develop aspiration pneumonia and hypoxia, which are often the cause of death.

Treatment

Treatment is difficult when horses develop clinical signs. They can be given antitoxin to help bind toxin still present. In some situations, horses are also vaccinated against tetanus to help stimulate an antibody response. Horses are placed on antibiotics to kill the *Clostridium*. Additionally, because they are very sensitive to stimulation, horses are placed in a quiet environment and often require sedation. Prognosis is poor in affected animals with approximately 80% mortality.

Practice Questions

1. At the neuromuscular junction, Ca^{2+} ions are necessary for:
- **a.** Binding the transmitter with the postsynaptic receptor.
- **b.** Facilitating diffusion of the transmitter to the postsynaptic membrane.
- **c.** Splitting the transmitter in the cleft, thus deactivating the transmitter.
- **d.** Fusing the presynaptic vesicle with the presynaptic membrane, thus releasing the transmitter.
- **e.** Metabolizing the transmitter within the presynaptic vesicle.

2. A drug that would prevent the release of acetylcholine at the neuromuscular junction would cause what, if any, clinical signs?
- **a.** Convulsions and excess muscle contractions
- **b.** Paralysis
- **c.** No effect on an animal's movement

3. Which one of the following is *true* with regard to the termination of synaptic action at the neuromuscular junction?
- **a.** The reuptake of intact acetylcholine molecules into the motor neuron terminal is responsible.
- **b.** Diffusion of acetylcholine away from the synapse is solely responsible.
- **c.** Acetylcholinesterase rapidly breaks down acetylcholine into choline and acetate.
- **d.** Dissociation of acetylcholine from the muscarinic receptor, after binding for several seconds, is solely responsible.

4. Several antagonist drugs compete with acetylcholine for the postsynaptic receptor at the neuromuscular junction. If you overdosed your patient with one of these competitive drugs, what would the antidote need to do at the synapse?
- **a.** Decrease the release of acetylcholine.
- **b.** Decrease the effectiveness of acetylcholinesterase.
- **c.** Decrease the influx of Ca^{2+} into the motor neuron terminal.
- **d.** Decrease the action potential frequency on the motor neuron.
- **e.** None of the above.

5. Which of the following statements regarding neuron-to-neuron synapses is *false*?
- **a.** The postsynaptic membrane is always a dendrite.
- **b.** Dendritic spines increase the surface area of the postsynaptic membrane.
- **c.** A single action potential on a presynaptic neuron is usually not sufficient to produce an action potential on a postsynaptic neuron.
- **d.** The neurotransmitter is not always released from a morphologically distinct active zone of the presynaptic terminal.
- **e.** Depending on the presynaptic neurotransmitter released and the postsynaptic receptor activated, the postsynaptic membrane can be either depolarized or hyperpolarized.

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 6th ed. Philadelphia: Lippincott, Williams & Wilkins; 2018.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Brunton L, Hilal-Dandan R, Knollman B, eds. *Goodman and Gilman's the Pharmacological Basis of Therapeutics*. 13th ed. New York: McGraw Hill; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Jones RA, Harrison C, Eaton SL, et al. Cellular and molecular anatomy of the human neuromuscular junction. *Cell Rep*. 2017;21(9):2348–2356.

Li XM, Dong XP, Luo SW. Retrograde regulation of motoneuron differentiation by muscle beta-catenin. *Nat Neurosci*. 2008;11(3):262–268.

MacKay RJ, Van Metre DC. Diseases of the nervous system. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby Elsevier; 2015.

Meyer JS, Quenzer LF. *Psychopharmacology: Drugs, The Brain, and Behavior*. 2nd ed. Sunderland, Mass: Sinauer; 2013.

Nestler EJ, Hyman SE, Holtzman DM, Malenka RC. *Molecular Neuropharmacology: A Foundation for Clinical Neuroscience*. 3rd ed. New York: McGraw-Hill; 2015.

Sudhof TC. Neurotransmitter release: the last millisecond in the life of a synaptic vesicle. *Neuron*. 2014;80:675–690.

Vet-ebooks.com

6

The Physiology of Muscle

BRADLEY G. KLEIN

KEY POINTS

1. Body movement is the result of contraction of skeletal muscle across a movable joint.
2. There are several levels of organization in any skeletal muscle.
3. Action potentials on the sarcolemma spread to the interior of the cell along the transverse tubules.
4. The action potential on the sarcolemma is indirectly coupled to the contraction mechanism through the release of Ca^{2+} from the sarcoplasmic reticulum.
5. The sliding of actin along the myosin molecule results in physical shortening of the sarcomere.
6. Most skeletal muscle fibers can be classified as either fast-contracting or slow-contracting fibers.
7. Muscles change their strength of contraction by varying the number of active motor units or the rate of motor unit activation.
8. The electromyogram is the clinical measurement of the electrical behavior within a skeletal muscle.
9. The structure of cardiac and smooth muscle differs from that of skeletal muscle.
10. The role of Ca^{2+} ions in excitation-contraction coupling in cardiac and smooth muscle is different than in skeletal muscle.

There are three types of muscle in the body: skeletal, cardiac, and smooth muscle. Skeletal muscle makes up about 40% of the body, and smooth muscle and cardiac muscle make up almost 10% more. Because most veterinary patients with disease of the neuromuscular system exhibit abnormalities of movement, it is important to understand how skeletal muscle functions and is controlled by the nervous system. Abnormalities of cardiac muscle and smooth muscle feature prominently in many other clinical disorders (e.g., dilated cardiomyopathy, hypertension, detrusor hypertrophy), and such muscle is often the target of pharmacological clinical intervention (e.g., sympathomimetic drugs, adrenergic receptor antagonists).

This chapter explains the physiology of skeletal muscle and includes brief comparisons with cardiac and smooth muscle. Cardiac muscle is discussed more extensively in Section III chapters. The role of smooth muscle in other body systems is mentioned throughout this book, and its neural control is discussed in Chapter 13.

Body Movement Is the Result of Contraction of Skeletal Muscle Across a Movable Joint

Skeletal muscle consists of a central, fleshy, contractile muscle "belly" and two tendons, one on each end of the muscle. The muscle and its tendons are arranged in the body so that they originate on one bone and insert on a different bone while spanning a joint. As the muscle contracts, shortening the distance between the origin and insertion tendons, the bones move in relation to each other, bending at the joint (Fig. 6.1). When activated by a motor nerve, a skeletal muscle can only shorten. Most joints have one or more muscles on both sides, either to decrease its angle (*flexion*) or to increase its angle (*extension*).

Body movement performed by an animal is the result of contraction of skeletal muscle across a movable joint. It is therefore important to understand the anatomy and physiology of skeletal muscle before discussing how the nervous system choreographs the contraction of groups of muscle cells to perform an impressive array of body movements.

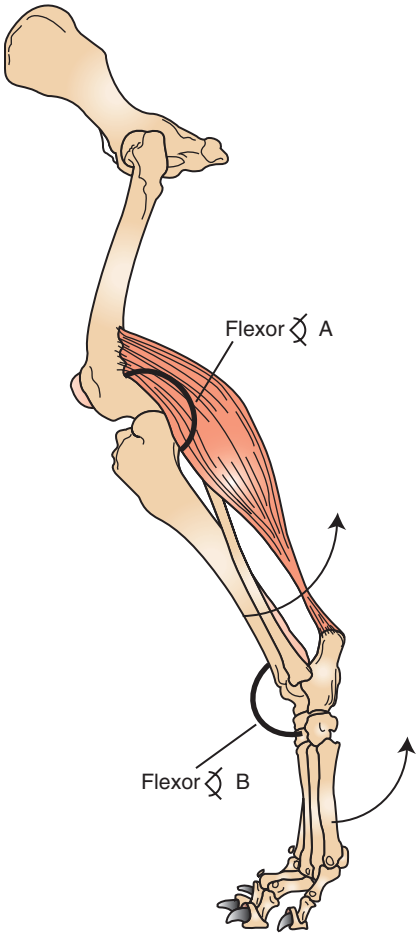
There Are Several Levels of Organization in Any Skeletal Muscle

Fig. 6.2 illustrates the levels of organization in a typical skeletal muscle. Each muscle belly seen during dissection is made up of differing numbers of muscle cells (usually called *muscle fibers*) that span the several inches between the origin and insertion tendons. The fibers range between 5 and 100 μm in diameter and contain multiple nuclei, multiple mitochondria, and other intracellular organelles. The outer limiting membrane of the fiber is called the *sarcolemma.* It consists of a true cell membrane, called the *plasma membrane,* and an outer polysaccharide layer that attaches to the tendons at the cells' extremities. With few exceptions, each muscle fiber is innervated by only one motor neuron, with the neuromuscular junction region located near the middle of the fiber, relative to the ends.

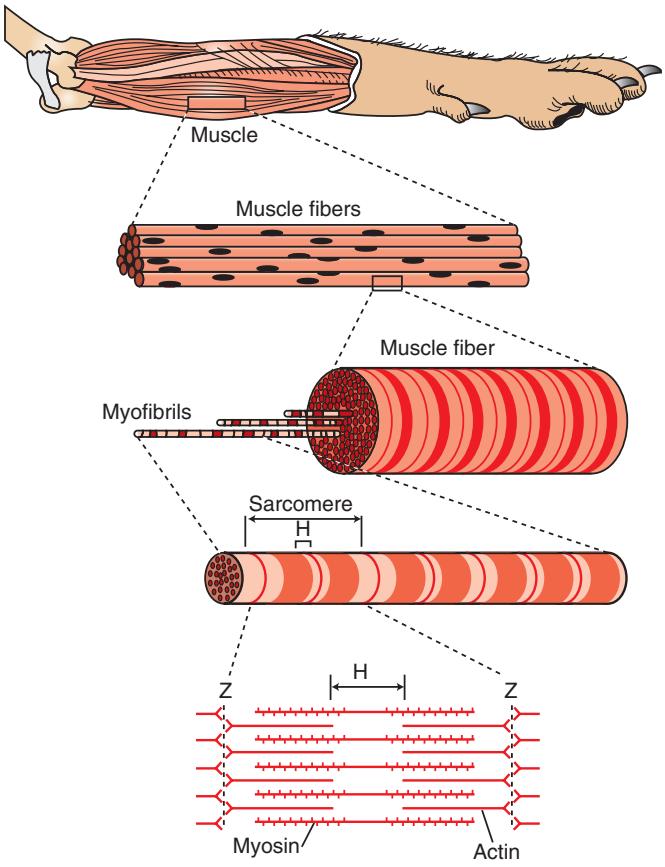
Each muscle fiber is made up of successively smaller subunits (see Fig. 6.2). Each fiber contains several hundred to several thousand *myofibrils* arranged in parallel along its length, like a handful of spaghetti. Each myofibril is made up of a linear series of repeating *sarcomeres,* the basic contractile units of the muscle fiber, which can number in the tens of thousands.

The sarcomere has a disk at each end called the *Z disk* (see Figs. 6.2 and 6.3). The sarcomere contains several types of large protein molecules responsible for muscular contraction, many of which are polymerized. Numerous thin protein filaments, called *actin,* are attached to the Z disks and extend toward the center of the sarcomere, similar to parallel fingers pointing at each other. Each

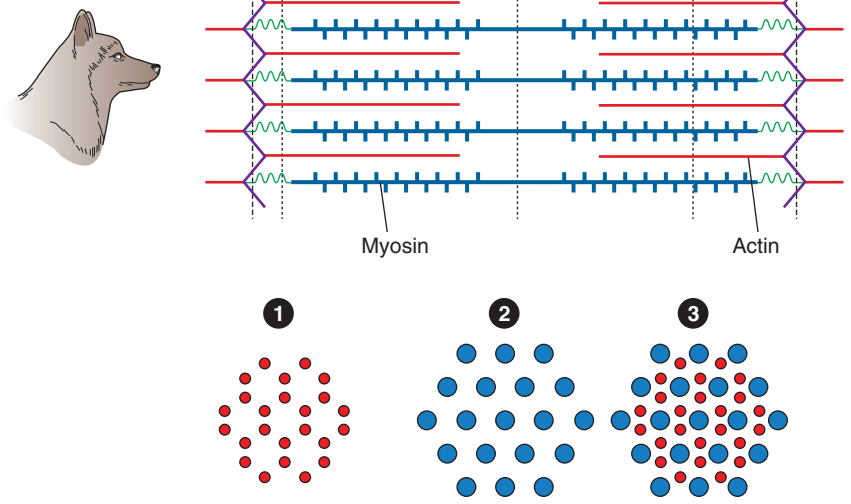
Vet-ebooks.com



• **Fig. 6.1** Body movement is the result of contraction (shortening) of a skeletal muscle attached across a movable joint. Contraction of the muscle will decrease the flexor angle at joint *A* (the stifle joint) and increase the flexor angle at joint *B* (the tarsal joint). This will produce the respective movements about the joints indicated by the *arrows.*



• **Fig. 6.2** A typical skeletal muscle has several levels of organization. *H* and *Z* are letters assigned to stripes seen during microscopical examination of skeletal muscle.



• **Fig. 6.3** Parallel arrangement of actin and myosin filaments in a sarcomere. *Top,* The viewer is looking at one end of a sarcomere. *Bottom,* The view of actin and myosin filament organization respectively seen by the observer at each of the three transverse sectioning points indicated in the top part of the figure. (Modified from Boron WF, Boulpaep EL. *Medical Physiology.* 2nd ed. Philadelphia: Saunders; 2009.)

Vet-ebooks.com

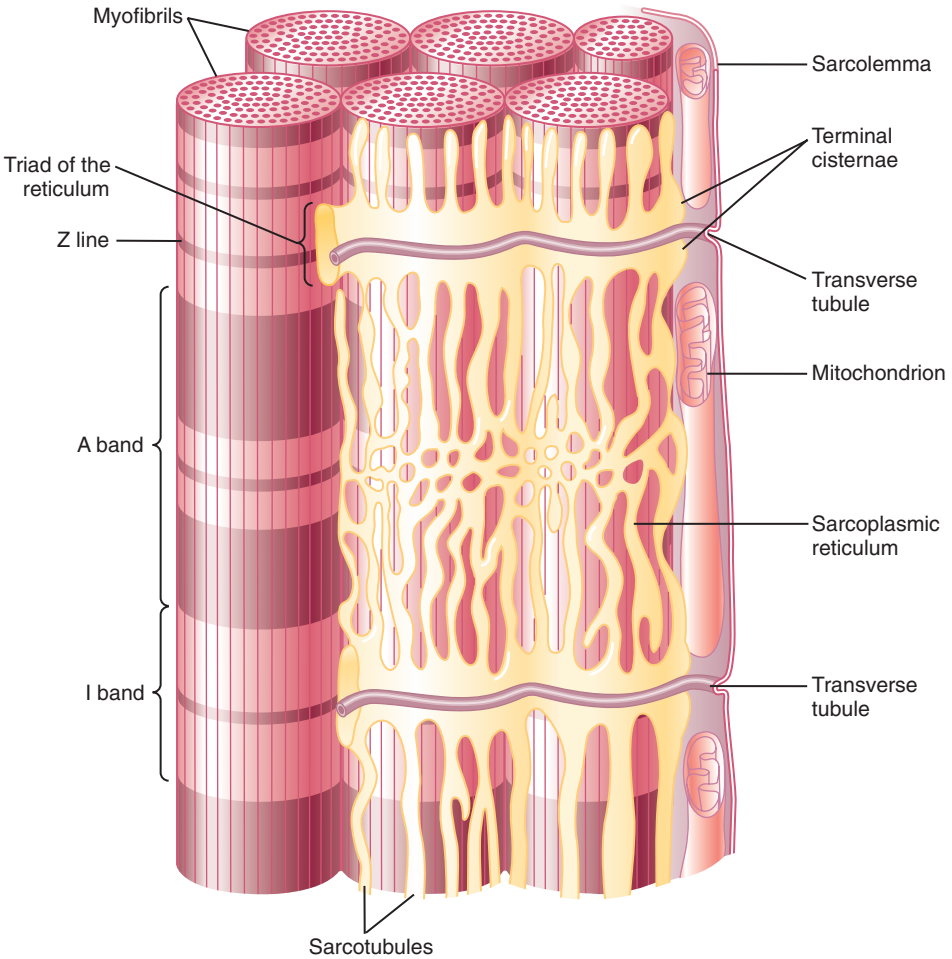
actin filament consists of two intertwined, helical strands of actin protein and two such strands of *tropomyosin* protein, all wound together as a larger helical complex (see Chapter 1 and Fig. 1.5). Also located intermittently along the tropomyosin-actin strand are complex globular protein molecules called *troponin* that can bind tropomyosin and actin, and that have an affinity for calcium (Ca^{2+}) ions. Suspended between and parallel to the actin thin filaments are thicker filaments of *myosin* protein polymers (Fig. 6.3). A myosin molecule contains a tail of intertwined helices and two globular heads that can bind both adenosine triphosphate (ATP) and actin (see Figs. 1.3 and 1.4). Approximately 500 myosin heads of a thick myosin filament form cross-bridges that interact with actin to shorten the sarcomere as the myosin heads flex and relax. The sarcomere also contains a large, springlike protein called *titin.* Titin, attached to both myosin and the Z disk, helps to maintain the side-by-side relationship of actin and myosin, as well as the elasticity and resting length of the sarcomere.

Beneath the plasma membrane of the muscle cell lies the *sarcoplasmic reticulum,* an intracellular storage organelle that forms a reticulated network around the myofibrils (Fig. 6.4). This extensive storage sac sequesters Ca^{2+} ions in relaxed muscle and is analogous to the smooth endoplasmic reticulum in other cells.

Located perpendicular to the long axis of the muscle fiber are tubes of plasma membrane formed by periodic invaginations of the sarcolemma (see Fig. 6.4). These *transverse tubules,* or *T tubules,* traverse the diameter of the muscle fiber, similar to a flexible drinking straw passing perpendicularly through the handful of spaghetti (myofibrils) noted earlier. The T tubules snake between the myofibrils, forming junctions with the network of sarcoplasmic reticulum that surrounds the myofibrils (Fig. 6.5). Because they are formed by invaginations of the sarcolemma (plasma membrane), these tubules are filled with extracellular fluid and are important because they allow the electrically excitable plasma membrane of the muscle fiber to carry the depolarization of the action potential to the interior of the fiber.

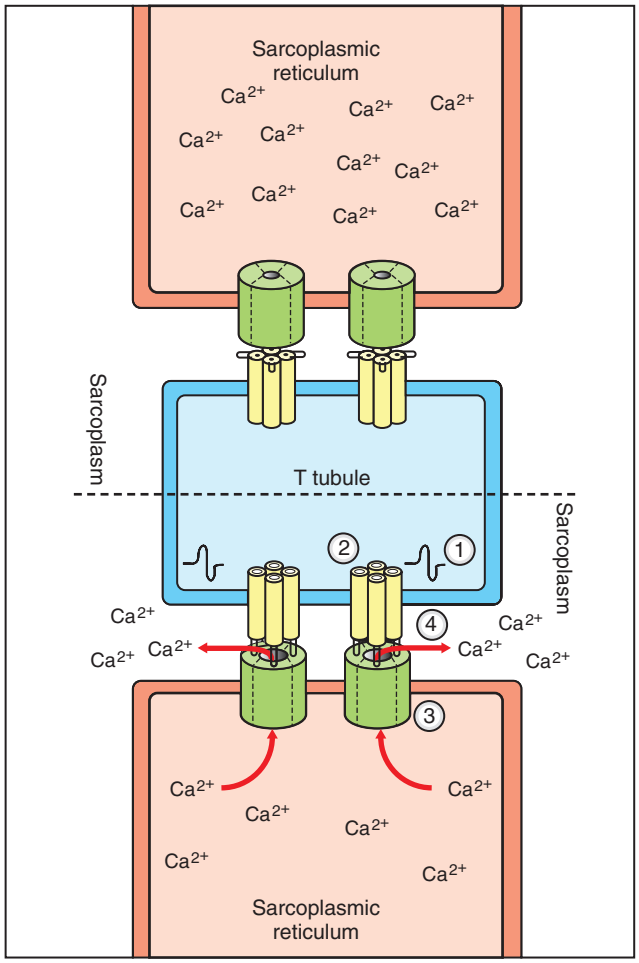
Action Potentials on the Sarcolemma Spread to the Interior of the Cell Along the Transverse Tubules

Skeletal muscle cells have a resting membrane potential, as do neurons, and the muscle cell membrane can be depolarized by synaptic transmission at the neuromuscular junction (see Chapter 5). At this junction, the acetylcholine released by the motor neuron activates nicotinic acetylcholine receptors on the sarcolemma of the muscle cell's motor end-plate region. The resulting depolarization, called an *end-plate potential (EPP),* is analogous to an EPSP (see Chapter 4 and Fig. 4.5) in a neuron. However, the magnitude



• **Fig. 6.4** Diagram of skeletal muscle showing the juxtaposition of myofibrils, transverse (T) tubules, and sarcoplasmic reticula. (Redrawn from Bloom W, Fawcett DW. *A Textbook of Histology.* Philadelphia: WB Saunders; 1986. Modified after Peachey LD. *J Cell Biol* 1965;25:209. Drawn by Sylvia Colard Keene. In Guyton AC, Hall JE. *Textbook of Medical Physiology.* 11th ed. Philadelphia: Saunders; 2006.)

Vet-ebooks.com



• **Fig. 6.5** Relationship between the T tubules (TT) and sarcoplasmic reticulum (SR) during excitation-contraction coupling. *Above dashed line,* TT-sarcoplasmic reticulum junction at rest. *Below dashed line, 1,* Propagation of action potential produces depolarization of the TT membrane. *2,* Depolarization induces a configurational change in voltage-gated Ca^{2+} channel aggregates (yellow) in the TT membrane. *3,* Opening of Ca^{2+} release channels on the SR membrane (green) results from their mechanical coupling with the configurational change in voltage-gated Ca^{2+} channel aggregates on the TT. *4,* Ca^{2+} is released from the SR into the sarcoplasm, where it can bathe the sarcomeres *(not shown)* to induce contraction. (Modified from Boron WF, Boulpaep EL. *Medical Physiology: A Cellular and Molecular Approach.* updated edition. Philadelphia: Saunders; 2005.)

of the individual EPP is sufficient to open enough voltage-gated sodium (Na^+) ion channels, also found at the motor end-plate sarcolemma (see Figs. 5.1 and 5.2), to trigger a muscle fiber action potential. Thus it is at the sarcolemma of the neuromuscular junction that muscle fiber action potentials are generated.

When an action potential is generated at the neuromuscular junction, near the midpoint of the muscle fiber, it spreads in both directions along the length of the fiber by mechanisms similar to action potential spread in unmyelinated neuronal axons. However, action potentials on the sarcolemma are also transmitted to the interior of the muscle fiber along the T tubules (see Fig. 6.5). This allows the action potential to reach the location of the sarcoplasmic reticulum that surrounds myofibrils, even in the innermost regions of the muscle fiber. The consequences of the action potential's arrival at the location of the sarcoplasmic reticulum are critical for the coupling of excitation (the action potential) with contraction (shortening) of the sarcomeres of the myofibrils.

The Action Potential on the Sarcolemma Is Indirectly Coupled to the Contraction Mechanism Through the Release of Ca^{2+} From the Sarcoplasmic Reticulum

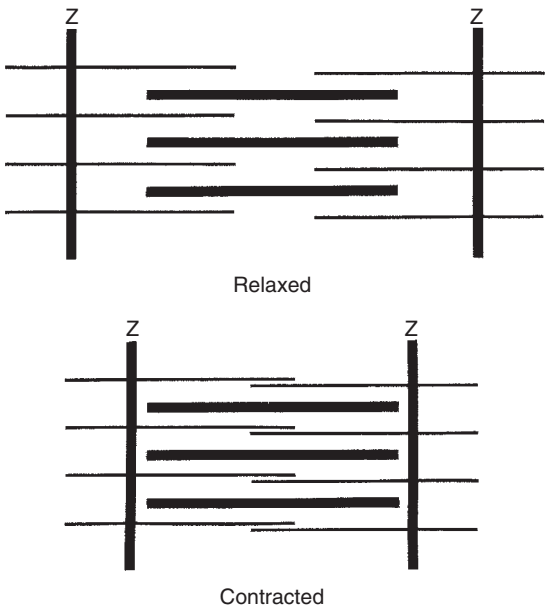
Whereas in the neuron a rise in cytoplasmic Ca^{2+} at the terminal is critical for initiating the process of transmitter release, a rise in Ca^{2+} in the muscle cell sarcoplasm (cytoplasm of a muscle cell) is critical for initiating *contraction.* At rest, Ca^{2+} ions are pumped out of the sarcoplasm and stored in the sarcoplasmic reticulum using an energy-dependent pump in conjunction with Ca^{2+} binding proteins within the sarcoplasmic reticulum. This leaves a concentration of Ca^{2+} in the sarcoplasm that is too low to trigger contraction. However, as an action potential spreads along the muscle fiber surface and into the fiber's core along the T tubules, the depolarization arrives at the junction between the tubules and the sarcoplasmic reticulum (see Fig. 6.5). The arrival of the action potential at this junction leads to the release of stored Ca^{2+} ions from the sarcoplasmic reticulum. These Ca^{2+} ions diffuse down their concentration gradient into the sarcoplasm, bathe the sarcomeres of the myofibrils, then trigger contraction. As the action potential passes, Ca^{2+} is pumped again into the sarcoplasmic reticulum and relaxation results. This cycle of events is known as *excitation-contraction coupling.*

The link between the action potential on the transverse tubule and Ca^{2+} release from the sarcoplasmic reticulum is mediated by voltage-gated Ca^{2+} channel complexes on the T tubule and Ca^{2+} release channels on the sarcoplasmic reticulum (see Fig. 6.5). The voltage-gated Ca^{2+} channels are also called dihydropyridine (DHP) receptors because they can be inhibited by this class of antihypertensive/antiarrhythmic drugs, also known as *Ca^{2+} channel blockers.* The Ca^{2+} release channels are also referred to as ryanodine receptors and, interestingly, can be activated by caffeine. In skeletal muscle, a mechanical coupling of the voltage-gated Ca^{2+} channel complex with subunits of the Ca^{2+} release channel is thought to exist such that action potential opening of the voltage-gated Ca^{2+} channels of the T tubule produces a direct configurational change in the Ca^{2+} release channels on the sarcoplasmic reticulum, allowing stored Ca^{2+} ions to escape from the sarcoplasmic reticulum into the sarcoplasm. In addition to opening due to their mechanical interaction with the voltage-gated Ca^{2+} channels of the T tubule, the Ca^{2+} release channels on the sarcoplasmic reticulum can be opened by Ca^{2+} that moves through the open voltage-gated Ca^{2+} channels. However, this mechanism plays a more important role in excitation-contraction coupling in cardiac muscle than in skeletal muscle (see later discussion).

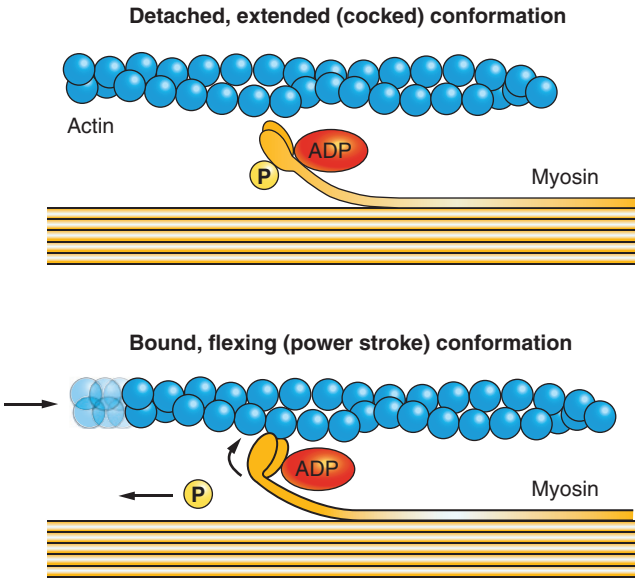
The Sliding of Actin Along the Myosin Molecule Results in Physical Shortening of the Sarcomere

Fig. 6.6 illustrates the sarcomere in the relaxed state and in its shorter, contracted state. The sarcomere is changed from its relaxed state to the shorter, contracted state when Ca^{2+} ions become available to the sarcomere. In the presence of Ca^{2+} ions and a sufficient source of ATP, the actin thin filaments are pulled in parallel along the myosin thick filaments by the repetitive movement of the myosin molecule heads, thus shortening the sarcomere. Because each myofibril is made up of a linear series of repeating and connected sarcomeres, the net result is the physical shortening of the

Vet-ebooks.com



• **Fig. 6.6** Sliding of actin along the myosin molecule results in the physical shortening (contraction) of the sarcomere.



• **Fig. 6.7** Actin-myosin binding, and flexing of the myosin heads, slides actin along the myosin filaments. *Top,* ATP bound to the myosin head has been hydrolyzed to ADP and inorganic phosphate *(P).* In this state the myosin head is extended (cocked) and detached from actin. *Bottom,* The resting myosin head has bound to actin, forming a cross bridge. The dissociation of P from the myosin head induces flexing of the head, pulling the actin filament along the myosin. This has been called the *power-stroke.* (Redrawn from Boron WF, Boulpaep EL. *Medical Physiology.* 2nd ed. Philadelphia: Saunders; 2009.)

distance between the two ends of the muscle. A more detailed molecular explanation of this sliding filament mechanism of sarcomere shortening is provided in Chapter 1 as an example of the binding specificity and allosteric interactions of proteins. However, the events can be briefly summarized as follows.

At several points along the actin thin filament, there are sites that can bind with the head of the myosin molecule (see Figs. 1.4 and 1.5, and accompanying Chapter 1 text). In the absence of Ca^{2+} ions, these sites are covered by the tropomyosin molecules that are normally interwoven within the actin helix. When Ca^{2+} is present and binds with troponin—a regulatory molecular complex attached to tropomyosin—the troponin complex undergoes a configurational change. It is thought that this then causes the tropomyosin molecule to move away from and uncover the myosin-binding site on the actin thin filament, permitting actin-myosin binding. This is coupled with an actin-myosin cross-bridge cycle. Through binding with and hydrolysis of ATP (to adenosine diphosphate [ADP] + phosphate) and subsequent dissociation of phosphate, the myosin heads alternately extend (the "cocked" state) and flex (the power stroke) while coordinately detaching and attaching to the exposed binding sites on the actin thin filament (Fig. 6.7). This results in the actin thin filaments sliding in parallel along the myosin thick filaments to shorten the sarcomere (see Fig. 6.6). In the absence of Ca^{2+}, the myosin-binding sites on actin again become blocked, and sarcomere relaxation results.

Most Skeletal Muscle Fibers Can Be Classified as Either Fast-Contracting or Slow-Contracting Fibers

Skeletal muscle fibers with rapid contraction speeds are sometimes called *fast-twitch fibers.* They tend to be thicker, have extensive sarcoplasmic reticulum for rapid release of Ca^{2+} ions, and have less extensive blood and mitochondrial supplies because aerobic metabolism is less important. Fast-twitch fibers are fairly rapidly fatigued but are well adapted for jumping, sprinting, and other brief, powerful movements.

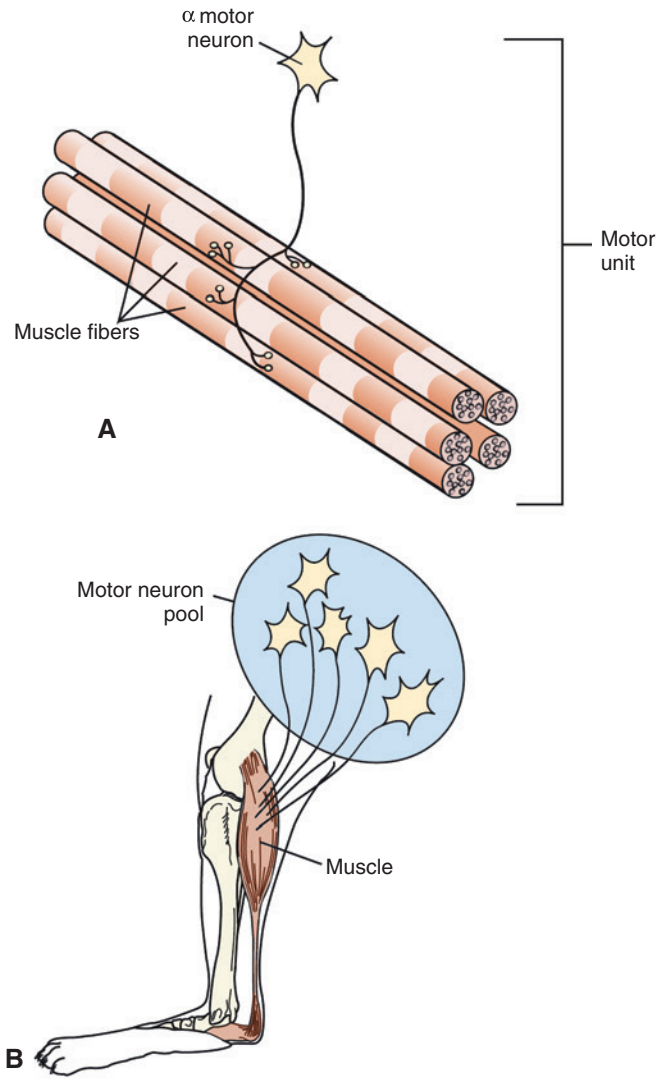
In contrast, *slow-twitch fibers* are thinner muscle fibers, have a rich blood and mitochondrial supply, and have a large amount of myoglobin, an iron-containing and oxygen-storing protein similar to hemoglobin. These fibers rely more heavily on oxidative metabolism and, although they contract at slower speeds, are less amenable to fatigue and better adapted for the continual contraction of antigravity extensor muscles.

Because slow-twitch fibers have more myoglobin, they are sometimes called *red muscle;* fast-twitch fibers are sometimes called *white muscle.* A third type of fiber, a subclass of fast-twitch fibers, has properties between the two types. Usually, a muscle belly is made up of a blend of these three types, the proportions varying in accordance with the muscle's use. This blend can be changed somewhat with exercise, such as in an athlete training for a different type of sports event. In reality, muscle fiber types may be better represented along a continuum, with fast-twitch and slow-twitch at the extremes, as opposed to belonging to exclusive categories.

Muscles Change Their Strength of Contraction by Varying the Number of Active Motor Units or the Rate of Motor Unit Activation

Even though each muscle fiber is innervated by only one neuron, each motor neuron's axon branches as it reaches the muscle and innervates several muscle fibers. A *motor unit* is defined as one alpha (α) motor neuron and all the *extrafusal* (force-generating, striated) muscle fibers that it innervates (Fig. 6.8A). All the muscle fibers of each motor unit are of the same functional type (e.g., fast or slow twitch), and an action potential on the motor neuron causes all the muscle fibers to contract simultaneously. In motor units, a relationship exists among the functional type of muscle fiber innervated, the number of muscle fibers innervated, and

Vet-ebooks.com



• **Fig. 6.8** Innervation of skeletal muscle by α motor neurons of central nervous system (CNS). (A) A motor unit is an α motor neuron and all the skeletal muscle fibers it innervates. (B) Neuronal cell bodies of all the motor units from a single muscle form a cluster within the CNS called the *motor neuron pool* of that muscle. (Redrawn from Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain.* 3rd ed. Philadelphia: Lippincott, Williams & Wilkins; 2007.)

motor neuron size. Small motor units tend to be made up of a motor neuron with a small cell body and a narrow, slower-conducting axon that innervates a small number of slow-twitch fibers. Large motor units have a motor neuron with a large cell body and a faster-conducting, wide axon innervating a large number of fast-twitch fibers. Activation of a small motor unit produces a smaller, slower, less fatigable increment of contractile force in the muscle compared with a larger motor unit. The neuronal cell bodies of all the motor units from a single muscle form a cluster within the central nervous system (CNS) called the *motor neuron pool* of that muscle (see Fig. 6.8B). Within the motor neuron pool for a given muscle, there is a range of motor unit sizes. Muscles with a larger proportion of smaller motor units tend to be amenable to finer control of contractile force.

Although an action potential on a motor neuron produces a simultaneous, brief twitch in all the muscle fibers of the motor unit, the pattern of excitation of the motor units originating from within the CNS produces the smooth, graded contraction of which most muscles are capable. The CNS can instruct a muscle to contract with greater force primarily by increasing the number of motor units that contract at any one time; this is called *multiple-fiber summation* (or *spatial summation*). The force of contraction can also be increased by increasing the frequency of activation of a motor unit, in which a subsequent twitch begins before relaxation of the previous twitch; this is called *frequency summation* (or *temporal summation*). These two processes work together, over much of the range of muscle contraction, to precisely adjust muscle force.

The recruitment of motor units to increase contractile force by multiple-fiber summation occurs in an orderly manner, according to motor unit size, with the smaller units activated first. This results in force being increased gradually in small, more precise, amounts when the muscle force required is low. As the force required increases, faster and larger increases in contractile force are progressively added by orderly activation of the larger motor units. This produces an overall smoothness of contraction, keeping the movement as precise as possible until larger, grosser increments are needed, usually when significant tension has already been generated in the muscle. In some skeletal muscles, the CNS can command a percentage of motor units to be active for extensive periods, with various motor units taking turns, thus keeping the contractile force, and distance between the origin and insertion tendon, relatively constant.

During frequency summation, when stimulation frequency increases to a point where there is minimal or no relaxation between individual twitches, the twitches fuse to produce smooth, continuous contraction of the muscle. When this brings the muscle to maximal contractile force, it is called *complete tetany.* Tetanization of cardiac muscle would be fatal, because heart muscle must relax to allow cardiac filling before it contracts to pump out the blood. Chapter 19 discusses how cardiac muscle prevents tetany.

The Electromyogram Is the Clinical Measurement of the Electrical Behavior Within a Skeletal Muscle

As an action potential spreads along a muscle fiber, a small portion of the electrical current generated spreads away from the fiber, even to the overlying skin. Electrodes placed on the skin or inserted into the muscle belly can record a summated electrical potential when the muscle contracts. Such a measurement, when visually displayed, is called an *electromyogram* (EMG) and is for skeletal muscle what the electrocardiogram (ECG) is for cardiac muscle. The EMG, often used in conjunction with nerve conduction analysis, helps to determine whether weakness or paralysis is caused by disease in the skeletal muscle, neuromuscular junction, motor neuron, or CNS.

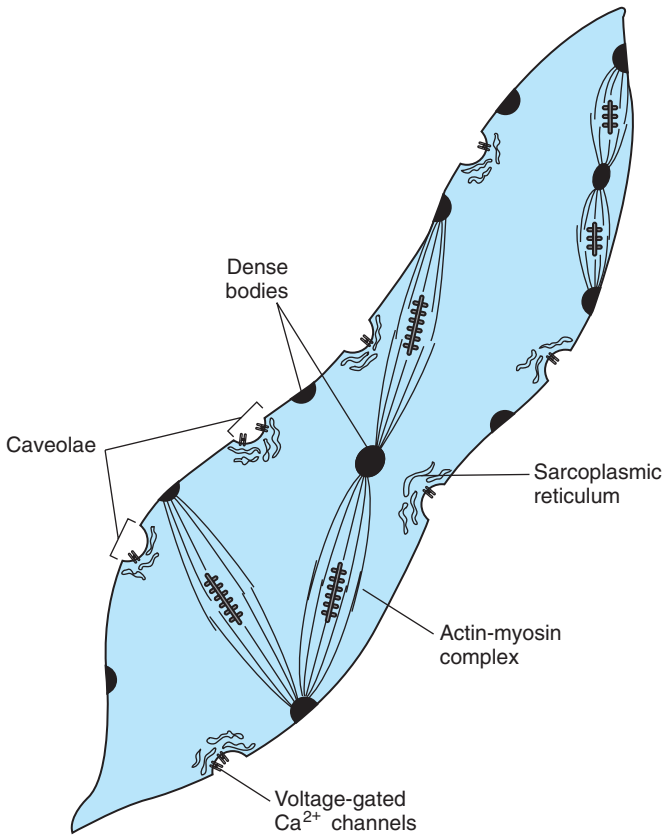
The Structure of Cardiac and Smooth Muscle Differs From That of Skeletal Muscle

As with skeletal muscle, *cardiac muscle* is striated and contains sarcoplasmic reticulum and myofibrils; the fundamental contractile component is formed by actin and myosin subunits (see Fig. 19.1). Cardiac muscle also contains transverse tubules, but cardiac muscle differs from skeletal muscle in some important ways. The shorter cardiac muscle cells are electrically coupled to each other through interconnected gap junctions that bridge *intercalated disks,* the latter being regions of close apposition between the membranes

Vet-ebooks.com

of cardiac cells aligned end to end. Because gap junctions provide continuity between the cytoplasm of adjacent cells, action potentials can spread from one cardiac muscle cell to another, across these intercalated disks, without the need for chemical neurotransmission to each cell. Such cytoplasmic continuity and electrical coupling are not characteristic of the longer skeletal muscle fibers. Also unlike skeletal muscle fibers, the cardiac muscle cells can possess branchlike extensions that form the aforementioned interconnections with some of their parallel neighbors. In fact, as explained in Chapter 19, action potentials arise spontaneously in specialized cardiac muscle cells and then spread throughout a large population of cardiac muscle cells as if they were a functional syncytium. This can result in coordinated contraction of a large region of the heart muscle. The frequency of such action potentials and the force of the resulting contraction are influenced by the autonomic nervous system, but such innervation is not necessary for action potential genesis.

Smooth muscle cells, as with cardiac myocytes, are smaller and shorter than skeletal muscle cells. They do not contain T tubules, and their sarcoplasmic reticulum is poorly developed (Fig. 6.9). These cells rely primarily on the transmembrane diffusion of Ca^{2+} ions from the extracellular fluid to induce the actin-myosin interactions responsible for contraction (see the following discussion). Although overlapping actin and myosin molecules form the contractile units of smooth muscle cells, the arrangement of these units lacks the structural regularity responsible for the striated appearance of skeletal and cardiac muscle cells. Actin filaments are anchored to *dense bodies* (instead of Z disks), which are found within the cytoplasm as well as in the cell membrane. Therefore these cells can appear to wrinkle on contraction.



• **Fig. 6.9** General organization of a smooth muscle cell. T tubules are absent, and the sarcoplasmic reticulum is poorly developed. Transmembrane diffusion of extracellular Ca^{2+}, through voltage-gated Ca^{2+} channels in caveolae, plays an important role in initiating contraction. Actin and myosin are present, with actin anchored to dense bodies. Activating the actin-myosin complex can change the cell's shape. (Modified from Guyton AC, Hall JE. *Textbook of Medical Physiology.* 11th ed. Philadelphia: Saunders; 2006.)

Some smooth muscle cell tissues, usually called *visceral* or *unitary smooth muscle* (as in acting together as a unit), have gap junctions between cells and operate similar to a functional syncytium, with cell-to-cell action potential transmission, and coordinated contraction, much as in cardiac muscle. Visceral smooth muscle is abundant in the gastrointestinal tract and other organs of the thoracic and abdominal cavities. This type of smooth muscle is described more fully in Chapter 28. Another type of smooth muscle cell tissue, usually called *multiunit smooth muscle,* has electrically isolated muscle cells capable of contracting independently of each other. Multiunit smooth muscle can be found, for example, in the iris and ciliary body of the eye, where precise control of muscular contraction is needed.

Smooth muscle tissue is innervated by neurons of the autonomic nervous system. In contrast to neuromuscular junctions at skeletal muscle, either acetylcholine or norepinephrine can be released (by different neurons) at junctions with smooth muscle, and smooth muscle cells can be either excited or inhibited by their presynaptic input. Visceral smooth muscle tends to receive a more diffuse innervation from an autonomic neuron, and the neurotransmitter is released at greater distance from the smooth muscle cell, compared with a more focal skeletal neuromuscular junction. In multilayer visceral smooth muscle tissue, neurotransmitter activation of surface cells can spread to deeper layers through the many gap junctions between cells. In multiunit smooth muscle tissue, it is more common to find synaptic input onto each cell and a synaptic cleft width similar to a skeletal neuromuscular junction. This is concordant with the more independent actions of these cells that have limited intercellular electrical continuity. In addition to control by autonomic neurons, various types of smooth muscle tissue can contract in response to self-induced generation of electrical activity, hormonal action, or stretch.

The Role of Ca^{2+} Ions in Excitation-Contraction Coupling in Cardiac and Smooth Muscle Is Different Than in Skeletal Muscle

Contraction of both cardiac and smooth muscle cells results from the sliding together of actin and myosin protein filaments, just as in skeletal muscle. This sliding of actin over myosin requires ATP and does not occur unless Ca^{2+} ions are present, again as in skeletal muscle. However, the origins of the intracytoplasmic Ca^{2+} ions that permit contraction differ.

In skeletal muscle, Ca^{2+} is sequestered in the sarcoplasmic reticulum. With the arrival of the action potential along the sarcolemma and T tubule, Ca^{2+} is released from the sarcoplasmic reticulum and diffuses out into the cytoplasm where it triggers contraction. Here, the mechanical opening of the Ca^{2+} release channels of the sarcoplasmic reticulum by the opening of the voltage-gated Ca^{2+} channels of the T tubule is the most significant vehicle for the rise in the cytoplasmic Ca^{2+}. With passage of the action potential, Ca^{2+} is pumped back into the sarcoplasmic reticulum, and the muscle relaxes. In skeletal muscle, little if any influx of extracellular Ca^{2+} (the T tubule lumen is an extension of the extracellular space) through the voltage-gated Ca^{2+} channels of the T tubule is needed for contraction.

Vet-ebooks.com

In cardiac muscle, the sarcoplasmic reticulum is not as well developed as in skeletal muscle. Therefore the influx of extracellular Ca^{2+} through the voltage-gated Ca^{2+} channels of the T tubule and the release of Ca^{2+} from the sarcoplasmic reticulum are both important in triggering contraction. In cardiac muscle, the arrival of the action potential along the cell membrane and the T tubules opens voltage-gated Ca^{2+} channels, allowing the influx of extracellular Ca^{2+} ions into the cytoplasm. These Ca^{2+} ions can activate the Ca^{2+} release channels on the sarcoplasmic reticulum, and these combined sources of increased cytoplasmic Ca^{2+} can trigger contraction. This nonmechanically coupled scenario of opening of the Ca^{2+} release channels on the sarcoplasmic reticulum of the cardiac muscle cells is called *Ca^{2+}-induced Ca^{2+} release*. If antihypertensive calcium channel blocker drugs (e.g., dihydropyridines) are used to block the entry of extracellular Ca^{2+} ions, the force of contraction is reduced. When the action potential has passed, muscle relaxation is accomplished primarily by pumping cytoplasmic Ca^{2+} back into the sarcoplasmic reticulum, although some Ca^{2+} is transported across the sarcolemma into the extracellular space.

In many smooth muscle cells, the sarcoplasmic reticulum is poorly developed and extracellular Ca^{2+} influx plays the principal role in initiating the contractile process. Although smooth muscle cells have no T tubules, this Ca^{2+} influx is achieved, upon membrane depolarization, through activation of voltage-gated Ca^{2+} channels located in shallow depressions of the membrane *(caveolae)* (see Fig. 6.9). Calcium channel blockers interfere with this process and can relax smooth muscle in arterial walls, which dilates arteries and lowers blood pressure. Contraction is terminated in many smooth muscle cells primarily by Ca^{2+} transport back into the extracellular space, which is a fairly slow process.

Some receptor-mediated stimulation of smooth muscle cells by neurons or by hormones does not lead to depolarization of the cell membrane but can lead to contraction of the smooth muscle cell nonetheless. For example, activation of some smooth muscle G protein–coupled membrane receptors (see Fig. 1.19), such as those that employ the inositol triphosphate-diacylglycerol (IP_3-DAG) second messenger pathway, can induce the release of Ca^{2+} stores from the smooth muscle cell sarcoplasmic reticulum, leading to contraction.

The mechanism by which Ca^{2+} induces actin-myosin cross-bridge cycling in smooth muscle differs from that in skeletal and cardiac muscle (see Fig. 1.17). In skeletal and cardiac muscle, cross-bridge cycling relies primarily on the Ca^{2+}-troponin-induced removal of the tropomyosin block of the actin-binding site. This is because the ATPase activity of the myosin head, which facilitates the effective interaction between myosin and actin, is normally high in these muscle types. In smooth muscle, where there is no troponin, cycling relies on Ca^{2+} to increase the ATPase activity of the myosin head (a slow process) so that it is able to efficiently interact with actin. This increase in ATPase activity is brought about when increased intracellular Ca^{2+} binds with calmodulin to activate the enzyme *myosin (light chain) kinase,* which, in turn, induces phosphorylation of the myosin head.

CLINICAL CORRELATIONS

Down Cow After Calving

History

A 4-year-old Jersey cow calved earlier this morning; this was her second calf. The producer called you because the cow stood after calving but appeared uncoordinated. Now, a few hours later, the cow is just recumbent and appears dull. She has been offered water and hay but has not accepted either. No other cows are affected. This cow has no history of medical problems.

Clinical Examination

The cow appears dull and does not pay much attention to you or the other activity in the barn. The cow's temperature is slightly low, and her heart rate is slightly increased. Her respirations are normal. She is slightly dehydrated. Her ears are cool to the touch, peripheral pulses are weak, and her rumen contractions are decreased. On examination, she does not appear to have any injuries that would prevent her from standing. Your brief neurological examination is normal, but the cow does have an S-shaped curve to her spine.

Comment

This cow most likely has *hypocalcemia*. Because of the high demands for calcium in the development of the calf in late pregnancy, combined with the production of colostrum and milk, this cow has become hypocalcemic. As reviewed in this chapter, calcium is critical for muscle contractions. Calcium also assists with membrane stabilization (return toward more normal membrane physiology) in peripheral nerve pathology. The release of acetylcholine (ACh) at the neuromuscular junctions is mediated by calcium. Hypocalcemia causes decreased ACh release, which can cause paralysis. Calcium deficits can also cause mild tetany, which is sometimes seen in cows with hypocalcemia. The cause is not clear but may be related to a direct influence of extracellular calcium on channels for other ions. All the clinical signs—hypothermia, increased heart rate, weak pulses, paresis, cool extremities, S-shaped curve of the spine, and decreased rumen contractions—can be attributed to hypocalcemia. A definitive diagnosis can be made by measuring ionized calcium. However, most veterinarians and producers will treat based on clinical signs, with the diagnosis confirmed based on response to treatment.

Treatment

Cows are treated with calcium gluconate, which is slowly given intravenously. Most cows will show improvement in clinical signs during the treatment. Cows often become brighter, their rumen contractility and peripheral circulation improve, and their core body temperature normalizes. Most cows attempt to stand after treatment, which is usually about 1 g per 100 pounds. Some cows will relapse and will need to be retreated.

Malignant Hyperthermia

History

You are called as the attending veterinarian to the veterinary teaching hospital. The large animal clinicians and anesthesiologists have just anesthetized a pig with halothane. The pig is starting to spasm and its temperature is increasing.

Clinical Examination

They have turned off the halothane. The pig is rigid with an increased temperature and increased heart (tachycardic) and respiratory (tachypneic) rates. Laboratory tests demonstrate that the pig has increased muscle enzymes from muscle damage and is acidotic.

Comment

For most species that develop malignant hyperthermia, there is a defect in the ryanodine receptor (another name for the Ca^{2+} release channels on the sarcoplasmic reticulum). This results in increased release of Ca^{2+} into the sarcoplasm, which stimulates the muscles. The uncontrolled muscle activation leads to rigidity, and because of the constant muscle contractions, the pig sustains muscle damage and consumes considerable oxygen. Therefore the muscle enzymes are increased, and the animal becomes acidotic due to energy and oxygen consumption.

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

Treatment

Dantrolene can be given to interfere with the release of Ca^{2+} from the sarcoplasmic reticulum. The exact mechanism is unknown; however, because it interferes with Ca^{2+} release, the Ca^{2+} is not available for muscle contractions. As a result, it can cause relaxation of the muscles. In addition, patients should be well ventilated to normalize the blood gas. Fluids are often given to reduce accumulation of creatinine kinase in the muscle, blood, and kidneys. If patients are rehydrated, their acid-base status may normalize due to correction of the lactic acidosis. However, if necessary, bicarbonate can be given to correct acid-base status.

Practice Questions

1. Troponin and tropomyosin are components of which one of the following structures?
- **a.** Myosin thick filament
- **b.** Sarcolemma
- **c.** T tubule
- **d.** Actin thin filament
- **e.** Sarcoplasmic reticulum

2. Action potentials in skeletal muscle cells trigger the release from the sarcoplasmic reticulum of what ion critical to the muscle's contractile process?
- **a.** Ca^{2+}
- **b.** Na^{+}
- **c.** K^{+}
- **d.** Cl^{-}
- **e.** HCO_3^{-}

3. A gross skeletal muscle belly can be instructed (by the central nervous system) to contract more forcefully by:
- **a.** Causing more of its motor units to contract simultaneously.
- **b.** Increasing the amount of acetylcholine released during each neuromuscular synaptic transmission.
- **c.** Increasing the frequency of action potentials in the α motor neuron's axon.
- **d.** Both a and c.
- **e.** Both b and c.

4. Which one of the following is *not* found in smooth muscle?
- **a.** Actin filaments
- **b.** Myosin filaments
- **c.** T tubules
- **d.** Voltage-gated calcium channels
- **e.** Sarcoplasmic reticulum

5. Choose the *incorrect* statement:
- **a.** The muscle fiber and neuronal cell membranes are similar because they both have a resting membrane potential.
- **b.** A whole muscle, such as the gastrocnemius muscle, can be made to contract more forcefully by increasing the number of motor units contracting.
- **c.** The muscle membrane's transverse tubular system transmits the action potential to the interior of the muscle cell.
- **d.** The muscle cell membrane transmits action potentials by saltatory conduction.
- **e.** The shortening of a skeletal muscle during contraction is caused by the sliding of actin along myosin filaments.

6. Which one of the following is *least likely* to be significantly associated with a muscle that is primarily involved in brief, powerful movements?
- **a.** Large α motor neuron cell body
- **b.** Small motor unit
- **c.** Fast-twitch fibers
- **d.** White muscle
- **e.** Large motor unit

Bibliography

Bailey JG. Muscle physiology. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Goff JP. Acute hypocalcemia (milk fever) in dairy cows (and goats). In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby; 2015.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Jiang D, Chen W, Xiao J, et al. Reduced threshold for luminal Ca^{2+} activation of RyR1 underlies a causal mechanism of porcine malignant hyperthermia. *J Biol Chem*. 2008;283(30):20813–20820.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Koeppin BM, Stanton BA. *Berne and Levy Physiology*. 7th ed. Philadelphia: Elsevier; 2018.

Matthews GG. *Cellular Physiology of Nerve and Muscle*. 4th ed. Malden, Mass: Wiley-Blackwell; 2003.

Myhre JL, Pilgrim D. A titan but not necessarily a ruler: assessing the role of titan during thick filament patterning and assembly. *Anat Rec*. 2014;297:1604–1614.

Rosenberg H, Davis M, James D, et al. Malignant hyperthermia. *Orphanet J Rare Dis*. 2007;2:21–35.

Witherspoon JW, Meilleur KG. Review of RyR1 pathway and associated pathomechanisms. *Acta Neuropathol Commun*. 2016;4:121–140.

Vet-ebooks.com

7

The Concept of a Reflex

BRADLEY G. KLEIN

KEY POINTS

1. A reflex arc contains five fundamental components.
2. Reflex arcs can be segmental, intersegmental, or suprasegmental.
3. Reflex arcs are widespread in the nervous system, and reflexes underlie a major portion of the neurological examination.

The *reflex arc,* the neural substrate of a reflex, is fundamental to the physiology of posture and locomotion, as well as to the clinical examination of the nervous system. A *reflex* can be defined as an involuntary, qualitatively unvarying response of the nervous system to a stimulus. Reflexes are the simplest behavioral example of the general function of the nervous system: collection of sensory input, integration, and motor output. Reflexes are often both critical for survival and components of more complex behaviors. The anatomy and function of a reflex arc are often fully developed at birth.

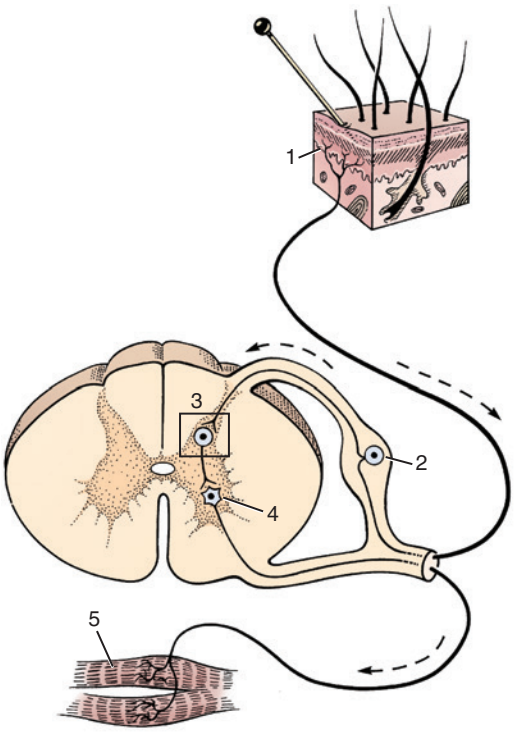
A Reflex Arc Contains Five Fundamental Components

All reflex arcs contain five basic components (Fig. 7.1). If any one of these five components malfunctions, the reflex response is altered.

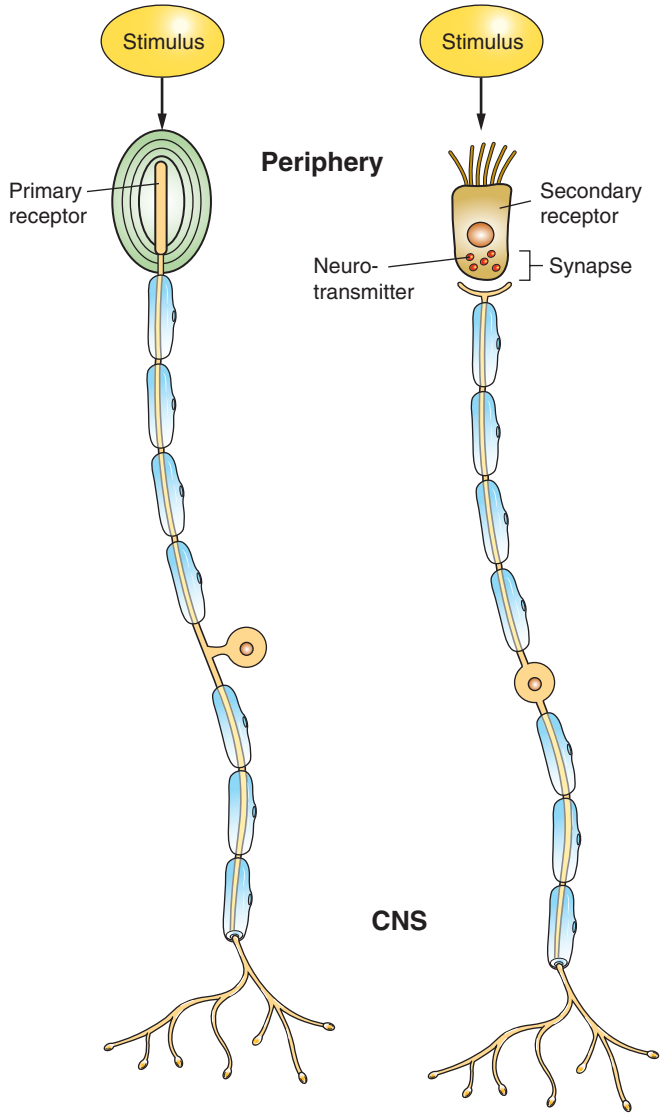
1. All reflex arcs begin with a *sensory receptor* (Fig. 7.1, #1). Sensory receptors vary widely within the body but share a common function: they transduce a range of environmental energy, or the presence of an environmental chemical, into a cellular response that directly or indirectly produces action potentials along a sensory neuron. In other words, these receptors collect environmental signals and turn them into a format that can be understood by the nervous system. For example, receptors of the retina transduce light; those in the skin transduce heat, cold, pressure, and other cutaneous stimuli; muscle spindle receptors transduce muscle stretch; and taste receptors transduce chemical stimuli from ingested material. A *primary sensory receptor* is a neuron with a specialized region for stimulus transduction (Fig. 7.2; also see Chapter 14, retinal photoreceptors). A *secondary sensory receptor* is a nonneural cell specialized for stimulus transduction that, in turn, affects neural activity by releasing neurotransmitter onto a neuron (see Fig. 7.2; also see Chapter 11, vestibular hair cells). Action potentials resulting from stimulus transduction are generated along sensory neurons at a frequency proportional to the intensity of the transduced stimulus. This proportionality between the intensity with which the receptor is stimulated and the frequency of the resulting sensory neuron action potentials is called *frequency coding*; it is one major way the receptor communicates to the central nervous system (CNS) the intensity of light, heat, stretch, and so forth that it has transduced. Stronger stimuli will also activate a larger number of sensory receptors, known as the *population code* of stimulus intensity.
2. The next component in a reflex arc, alluded to earlier, is a *sensory neuron* (CNS afferent; Fig. 7.1, #2). These neurons carry action potentials, resulting from receptor activation, to the CNS. Again, in some cases, the receptor is just a specialized, usually peripheral, region of the sensory neuron (primary receptors). In other cases, the receptor is physically separate from and synapses on the sensory neuron (secondary receptors). Sensory neurons enter the spinal cord by way of the dorsal roots or enter the brain through cranial nerves.
3. The third component of a reflex arc is a *synapse* in the CNS (Fig. 7.1, #3). Actually, for most reflex arcs, more than one synapse occurs in series (*polysynaptic*). However, some reflex arcs that originate from the muscle spindle are monosynaptic (single synapse between sensory neuron input and motor output neuron). In polysynaptic reflexes, where one or more neurons lie between the sensory neuron input to the CNS and the motor neuron output, these interposed neurons are called *interneurons* and can be considered part of this third component of the reflex arc.
4. The fourth component is a *motor neuron* (CNS efferent; Fig. 7.1, #4), which carries action potentials from the CNS toward the synapse with the target (effector) organ. Motor neurons leave the spinal cord through the ventral roots, and motor neurons leave the brain through the cranial nerves.
5. The last component is a *target organ* (effector organ) that causes the reflex response (Fig. 7.1, #5). This is usually a muscle, such as the skeletal muscle fibers of the quadriceps muscle of the leg, in the case of the "knee jerk" muscle stretch reflex, or the smooth muscle of the iris in the pupillary light reflex. The target could also be a gland, such as a salivary gland in the salivary reflex.

In reality, the final reflex response to a stimulus in mammals is rarely, if ever, the product of a monosynaptic reflex arc acting in isolation. Even if a sensory neuron participates in a monosynaptic reflex arc, it will often give off branches in the CNS that participate in polysynaptic reflex circuits. In addition, even the simplest mammalian reflex responses often involve both the excitation of a given muscle or muscles and the inhibition of another (usually antagonistic) muscle or muscles. The knee jerk reflex is a good

Vet-ebooks.com



• **Fig. 7.1** A reflex arc contains five fundamental components: *1,* a receptor; *2,* a sensory neuron; *3,* one or more synapses in the central nervous system; *4,* a motor neuron; and *5,* a target organ, usually a muscle. (From De Lahunta A, ed. *Veterinary Neuroanatomy and Clinical Neurology*. 2nd ed. Philadelphia: Saunders; 1983.)



• **Fig. 7.2** Primary and secondary sensory receptors. A primary sensory receptor (*left*) is a neuron with a peripheral ending specialized for stimulus transduction. In this particular case, the encapsulated peripheral ending of the neuron transduces the stimulus. The secondary receptor (*right*) is a nonneural cell designed for stimulus transduction, which subsequently releases neurotransmitter onto an adjacent neuron. *CNS,* Central nervous system.

illustration of these points (see Chapter 8). With regard to the individual sensory neurons that underlie this reflex, some terminal branches make excitatory monosynaptic connections with motor neurons that activate the quadriceps muscle. Other terminal branches of that same sensory neuron participate in a disynaptic circuit that inhibits motor neurons innervating the antagonistic hamstring muscle.

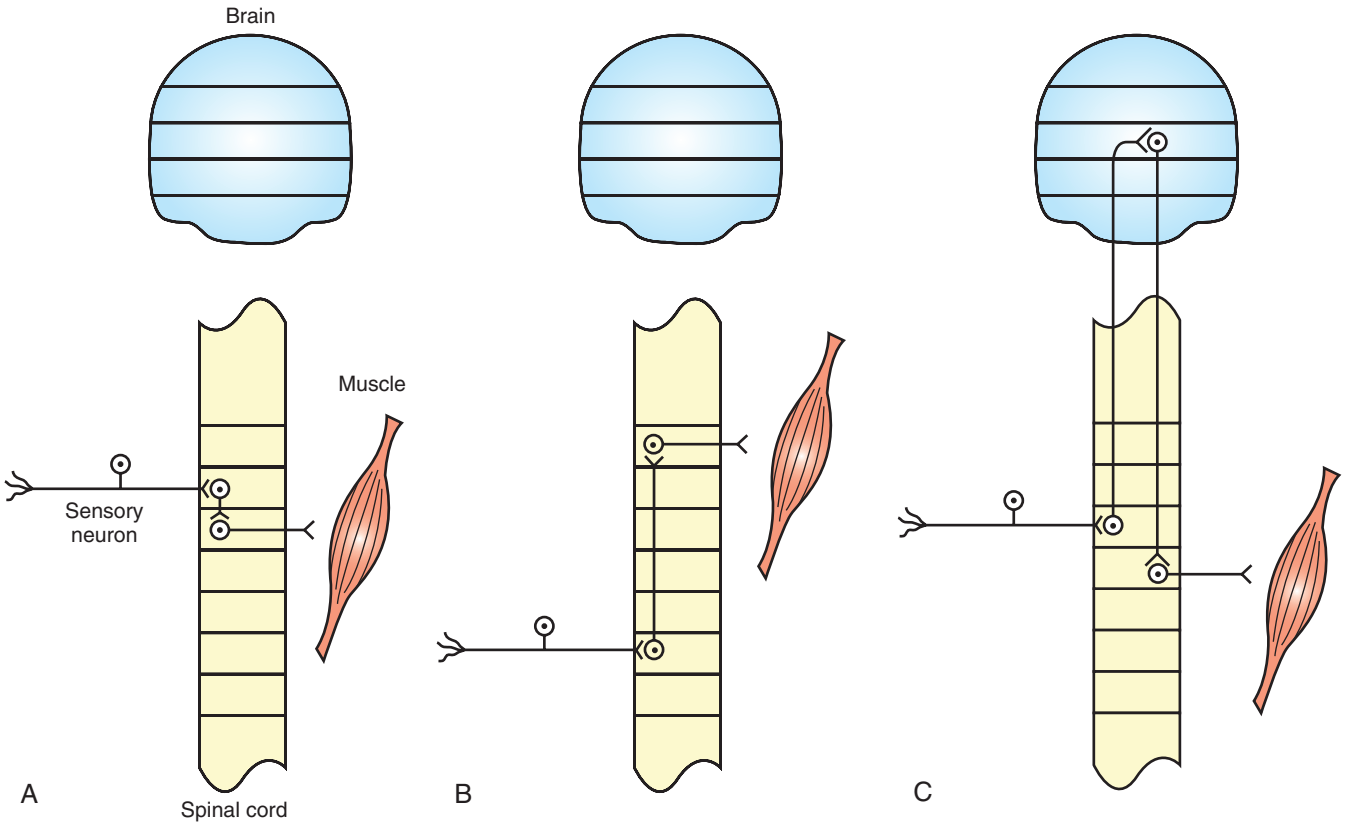
Also, reflexes do not operate in isolation from the rest of the nervous system. Inputs from other regions of the nervous system, onto the CNS components of a reflex arc, can modulate the sensitivity of the reflex response. For example, if input from the brain to the spinal cord is damaged, reflexes in that region of the cord can become exaggerated.

Reflex Arcs Can Be Segmental, Intersegmental, or Suprasegmental

A *segmental reflex* is a reflex in which the reflex arc passes through only a small rostrocaudal portion of the CNS (Fig. 7.3A). In such cases, the sensory neuron entrance to the CNS, the CNS circuitry, and the motor neuron exit region all have a similar rostrocaudal location. The term, segmental reflex, is most often used in reference to spinal cord reflexes, but some reflexes in the brain are of a segmental-like nature because their circuitry is restricted to one or two major brain divisions (e.g., midbrain, diencephalon). The quadriceps stretch reflex (knee jerk reflex) and the pupillary light reflex are examples of such reflexes because they, respectively, use only a small number of spinal cord segments (e.g., L4–L6) or a small rostrocaudal region of the brainstem.

The term *intersegmental reflex* most commonly refers to a reflex arc where the CNS circuitry traverses several segments of the spinal cord rostral or caudal to the site of entry of the sensory neuron signals (Fig. 7.3B). An example is the cutaneous trunci reflex. In this reflex, signals from a skin pinch of the dorsum of the trunk, entering the spinal cord as far caudally as the L5–L6 region, are carried as far rostrally as the T8–C1 region to excite motor neurons leaving the cord there to activate the cutaneous trunci muscle. If an intersegmental reflex arc's travel also includes a portion of the brain, it is often referred to as a *suprasegmental reflex.* The vestibulospinal reflexes that produce postural adjustments in response

Vet-ebooks.com



• **Fig. 7.3** Segmental, intersegmental, and suprasegmental reflexes. (A) In the segmental reflex, the sensory neuron input, CNS circuitry, and motor neuron output traverse only a small number of rostrocaudal segments of the spinal cord. (B) Intersegmental reflex arcs traverse several spinal cord segments from the site of sensory neuron input. (C) Suprasegmental reflexes are intersegmental reflex arcs that include part of the brain. For some suprasegmental reflexes, the sensory input and motor output are located in close rostrocaudal proximity in the cord, with the central nervous system components between having traveled to, and back from, the distant brain. These are called long-loop reflexes. Horizontal lines delimit either spinal cord segments (e.g., L1, L2) or major brain divisions (e.g., medulla, pons).

to acceleration or tilt of the head are examples of supraspinal reflexes. In these vestibulospinal reflexes, the sensory input, originating in the vestibular apparatus of the inner ear, enters the CNS near the pontomedullary border of the brainstem, and the motor neurons that exit the CNS are often distributed over several segments in the spinal cord.

For some suprasegmental reflexes, the sensory neuron entrance to the CNS and the motor neuron exit are at a similar rostrocaudal location, but the CNS circuitry lying between travels to and returns from a distant region of the CNS (see Fig. 7.3C). These are sometimes referred to as *long-loop* reflexes. The proprioceptive positioning (paw replacement) reaction is a clinical assessment often associated with this category. This reaction involves the animal promptly returning its paw to the normal pad-down position after the clinician flexes it to make the dorsal surface touch the floor or tabletop. The sensorimotor circuitry of this reaction courses from the limb's peripheral nerves and associated spinal cord segments, all the way through the brain, and back to the same limb's spinal cord segments and peripheral nerves. Although sometimes referred to as a reflex given its reliable nature in normal animals, the term "reaction" is more commonly used to acknowledge the potential for conscious, cortical influences.

Reflex Arcs Are Widespread in the Nervous System, and Reflexes Underlie a Major Portion of the Neurological Examination

Reflex arcs are ubiquitous in the nervous system and are the basis of much of an animal's subconscious, involuntary response to its environment. Much of a veterinarian's clinical examination of the nervous system involves evoking reflex responses, such as the pupillary light reflex, muscle stretch (knee jerk) reflex, and flexor reflex.

If any of the five components of the reflex arc malfunctions, the expected reflex response is altered or does not occur. It is important to know the general anatomy, physiology, and expected normal clinical response of the common reflexes to perform a neurological examination so that lesions can be localized. For example, loss of pupillary constriction to light in conjunction with normal responses to other visual stimuli, such as avoiding a rapidly oncoming object, suggests that the problem is probably not with the sensory receptor or sensory neuron components of the reflex arc. Several such reflexes are discussed in detail in subsequent chapters.

Vet-ebooks.com

CLINICAL CORRELATIONS

Trauma in a Foal

History

Distraught owners call you about their 4-month-old Tennessee Walking Horse colt. He appeared normal this morning when they let him out to pasture with his mother, but later this afternoon, the mare and the foal did not come in to be fed. The owners went out to the pasture and found the mare with the foal, who would not get up. He was lying on his side and seemed unable to position himself sternal. When the owners tried to reposition him, the foal thrashed, trying to get away. You tell the owners not to move the foal and that you will be there soon.

Clinical Examination

The foal appears to be responsive but in great pain and unable to rise. Temperature, pulse, and respirations are all mildly increased. There appears to be a swelling along the cervical (neck) area in the region of C1 to C3. The swelling is hard (bone) and has some fluid (inflammation) as well. There is some crepitus (crackling sound made by bone fragment friction) in the area of the swelling (possible fracture site). The foal displays no other areas of swelling or trauma. Neurological examination reveals normal cranial nerves. In the front limbs, the biceps and triceps reflexes seem increased on both sides. Deep pain is present, and cutaneous sensation is increased bilaterally. In the hind limbs, the femoral, sciatic, and tibial responses are increased.

Comment

Although it is difficult to localize a fracture definitively, based on history and physical examination, a fracture seems likely. The fracture appears to be in the region of C1 to C3. Radiographs would be ideal to make a definitive diagnosis. On neurological testing of the biceps, triceps, sciatic, femoral, and cranial tibial responses, all assess segmental reflex arcs. Because of a high cervical fracture, the descending motor tracts that supply both the thoracic and the pelvic limbs are affected. From point of initiation, the segmental reflexes tested involve the muscle spindles, which detect stretch, followed by sensory fibers in the peripheral nerves, the dorsal root and its ganglion, and the central branches of the sensory fibers projecting onto the ventral horn cell of the same spinal segment. The efferent pathway involves the ventral horn cell (lower motor neuron) followed by the ventral root, motor fibers in the peripheral nerve, neuromuscular junction, and then the myofibers being tested. The presence, but exaggeration, of the segmental reflexes reflects a change in the modulation of these reflexes by the damaged descending motor tracts that pass through the area of the lesion.

Treatment

The prognosis for this foal is poor. Based on the physical examination and clinical signs, a fracture is likely, and there is little hope for recovery. The complications associated with trying to manage a foal as the fracture heals are enormous. The fracture may not heal, and the foal could have severe residual neurological deficits. In most cases, these foals are euthanized fairly quickly because of the poor prognosis.

Practice Questions

1. Which of the following is *not* always a component of a reflex arc?
- **a.** Receptor
- **b.** Sensory neuron (CNS afferent)
- **c.** Central nervous system (CNS) interneuron
- **d.** Motor neuron (CNS efferent)
- **e.** Target (effector) organ

2. Which of the following regarding sensory receptors is *false*?
- **a.** They transduce environmental signals, directly or indirectly, into neuronal action potentials.
- **b.** A primary receptor can be a specialized peripheral region of a sensory neuron.
- **c.** They directly transduce CNS action potentials into physical activity of a target organ.
- **d.** They are the initial component of a reflex arc.

3. When the intensity with which a receptor is stimulated is increased, what happens to the *frequency* of action potentials along the sensory neuron from that receptor?
- **a.** Increases
- **b.** Decreases
- **c.** No change

4. Which one of the following is an example of a segmental reflex?
- **a.** Quadriceps stretch reflex
- **b.** Cutaneous trunci reflex
- **c.** Vestibulospinal postural reflexes

5. An intersegmental reflex arc is one in which:
- **a.** The course of the arc is restricted to one or a small number of segments of the CNS.
- **b.** No target organ is present.
- **c.** No receptor is present.
- **d.** The course of the arc traverses several segments of the CNS.
- **e.** Both b and c.

Bibliography

Binder MD. Peripheral motor control: spinal reflex actions of muscle, joint and cutaneous receptors. In: Patton HD, Fuchs AF, Hille B, et al, eds. *Textbook of Physiology*. 21th ed. Philadelphia: Saunders; 1989.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Divers TJ. Spinal fractures and luxations and spinal cord trauma. In: Smith BP, ed. *Large Animal Internal Medicine*. 4th ed. St Louis: Mosby; 2009.

Fischer DB, Truog RD. What is a reflex? A guide for understanding disorders of consciousness. *Neurology*. 2015;85:543–548.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Lorenz MD, Coates JR, Kent M. *Handbook of Veterinary Neurology*. 5th ed. Philadelphia: Saunders; 2011.

Schatzberg SJ. Neurologic examination and neuroanatomic diagnosis. In: Ettinger SJ, Feldman EC, Cote E, eds. *Textbook of Veterinary Internal Medicine*. 8th ed. St Louis: Elsevier; 2017.

Vet-ebooks.com

8

Skeletal Muscle Receptor Organs

BRADLEY G. KLEIN

KEY POINTS

1. The muscle spindle stretch receptor is an encapsulated organ of specialized muscle fibers with separate motor and sensory innervations.
2. The muscle spindle conveys information about muscle length to the central nervous system.
3. Muscle stretch and action potentials along spindle sensory neurons lead to reflex contraction of the extrafusal muscle fibers.
4. The central nervous system can control spindle sensitivity directly through the gamma (γ) motor neurons.
5. The Golgi tendon organ lies in series between muscle and tendon, and detects muscle tension.
6. Free (non-organ) sensory receptors in joints and muscles can provide information about joint position, joint movement, and pain-inducing stimuli of joints and muscles.

Movement, characteristic of all animals, is the end product of skeletal muscle contraction. It is orchestrated by the central nervous system (CNS) through its control of the motor unit (see Chapter 6). To control body movement appropriately, the CNS must (1) assess the effect of gravity on the many muscles of the body, (2) determine the initial position of the body parts to be moved, and (3) detect any discrepancy between the intended movement and the movement that actually occurs. When such discrepancies are detected, appropriate adjustments can be made.

Two important receptor systems have evolved in the skeletal muscles of mammals to provide the CNS with the aforementioned information: the muscle spindle and the Golgi tendon organ (Fig. 8.1). The *muscle spindles,* arranged in parallel to the contracting skeletal muscle fibers, provide information about muscle length. The *Golgi tendon organ,* arranged in series with the contracting skeletal muscle fibers, detects muscle tension. This chapter discusses the anatomy and physiology of these two receptor organs; Chapter 10 discusses how the CNS uses the information gathered from these organs to coordinate posture and locomotion. Some of this information is used in reflex arcs of the type described in Chapter 7.

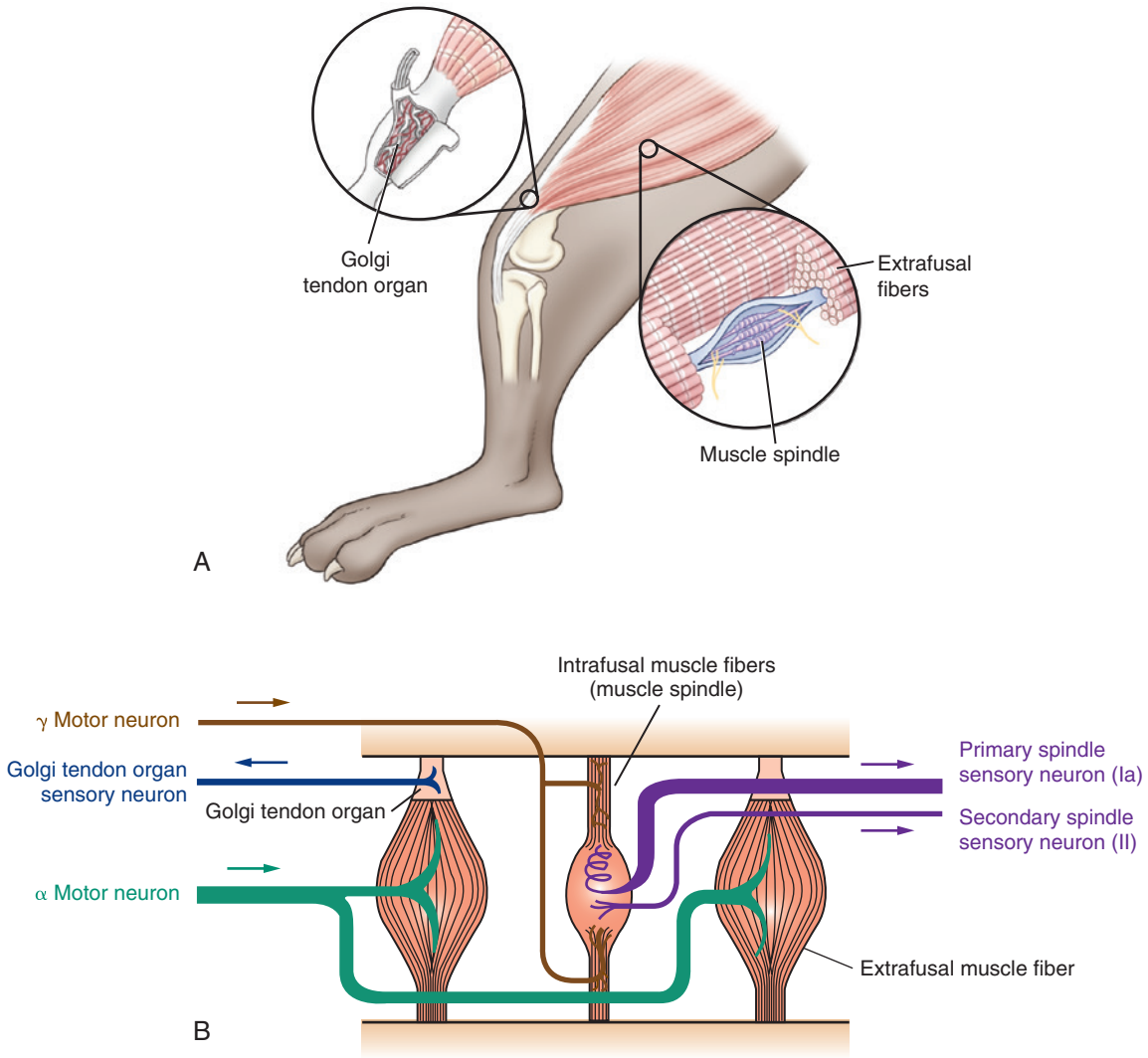
The Muscle Spindle Stretch Receptor Is an Encapsulated Organ of Specialized Muscle Fibers With Separate Motor and Sensory Innervations

The muscle spindle is an encapsulated group of about 3 to 12 small, slender, specialized skeletal muscle fibers (Figs. 8.1 and 8.2). Because their capsule is spindle-shaped, or *fusiform,* these muscle fibers are called *intrafusal muscle fibers.* The muscle fibers that cause physical shortening of the muscle (the majority of muscle fibers in a muscle belly), located outside the capsule, are called *extrafusal muscle fibers.* Extrafusal muscle fibers often span the length of the gross muscle from origin to insertion tendon; intrafusal muscle fibers and their capsules are much shorter (about 4–10 mm long). In addition, the intrafusal muscle fiber endings are attached to the extracellular matrix of, and lie in parallel to, the extrafusal muscle fibers. Therefore if the muscle is stretched, lengthening the extrafusal muscle fibers, the intrafusal fibers of the muscle spindle are also stretched.

Unlike extrafusal muscle fibers, the contractile elements of intrafusal muscle fibers are restricted to their polar ends, with none in their middle (equatorial) region. Therefore their polar ends can contract, but their equatorial region cannot. Such contraction does not directly contribute to the shortening of the gross muscle, but it can tighten the region of the intrafusal fiber that lies between the two poles. As discussed later, this can have a dramatic effect on the muscle spindle's sensitivity for transducing muscle stretch.

Spindle sensory neurons arise from the equatorial region of the intrafusal muscle fibers and carry action potentials from the spindle to the CNS by way of the peripheral nerves. These CNS afferents enter the spinal cord through the dorsal roots (Fig. 8.3). The contractile, polar regions of the intrafusal muscle fibers are innervated by motor neurons called *gamma (γ) motor neurons.* Extrafusal muscle fibers—the muscle fibers that cause the physical shortening of the muscle—are supplied by a different population of motor neurons (those that make up the motor units) called *alpha (α) motor neurons.* Although γ motor neurons go to intrafusal muscle fibers and α motor neurons go to the extrafusal muscle fibers, these CNS efferents both have their cell bodies in the ventral horn of the spinal cord, and their axons leave through the ventral roots.

Vet-ebooks.com



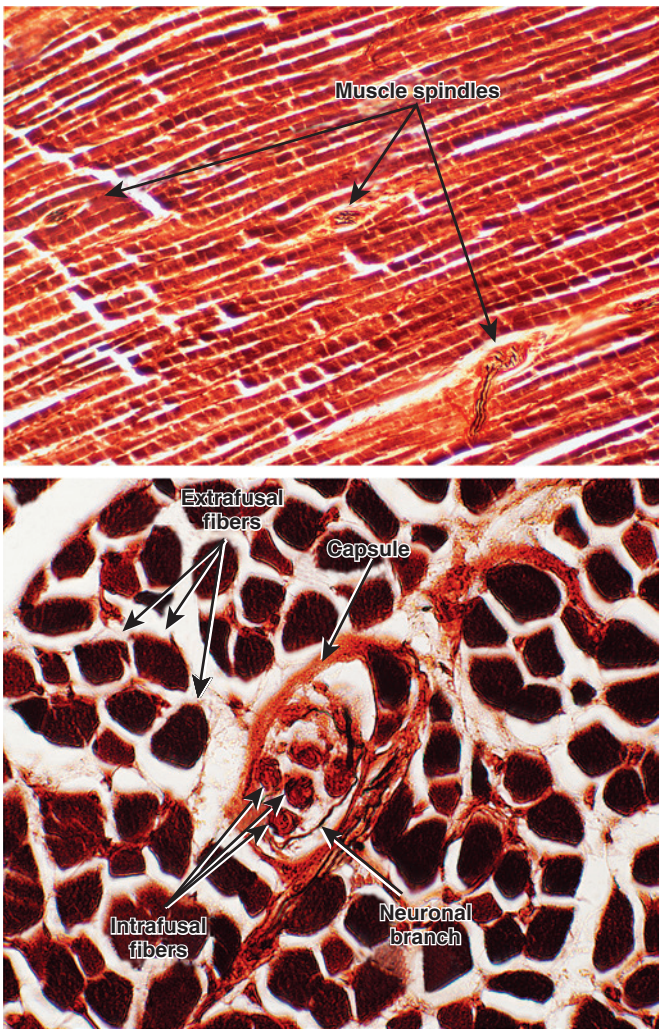
• **Fig. 8.1** Skeletal muscles have two important receptors: the muscle spindle and the Golgi tendon organ. (A) The muscle spindle, containing intrafusal fibers, is arranged in parallel with the extrafusal muscle fibers; the Golgi tendon organ is in series with the extrafusal fibers. (B) Types of neurons innervating the muscle and its sensory organs, and direction of action potential flow (*arrows*) along those neurons. (Modified from Kandel ER, Schwartz JH. *Principles of Neural Science*. 2nd ed. New York: Elsevier Science; 1985.)

The Muscle Spindle Conveys Information About Muscle Length to the Central Nervous System

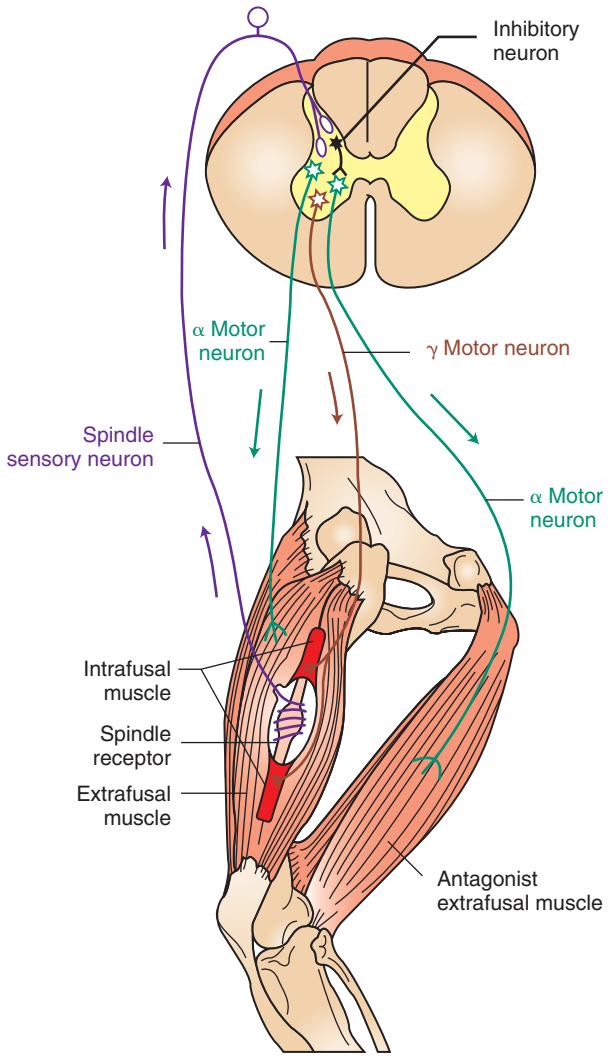
As noted, stretching a muscle can stretch the intrafusal fibers of the muscle spindle that lie parallel to the extrafusal muscle fibers. Stretching (lengthening) the equatorial segment of the intrafusal muscle fiber generates action potentials along the spindle sensory neurons. As the equatorial segment is lengthened, it is believed that stretch-sensitive ion channels open on the sensory neurons, leading to membrane depolarization and action potential generation. Action potentials are generated along the sensory neuron output of the muscle spindle in proportion to the amount of lengthening of the middle of the intrafusal muscle fibers. There are actually subclasses of intrafusal fibers (dynamic nuclear bag, static nuclear bag, and nuclear chain) and spindle sensory neurons (Ia, II). This variety allows the sensory neuron output of the spindle to detect not only a change in length during the dynamic phase of muscle lengthening, but also the rate of lengthening, as well as the steady-state length of the muscle as the animal holds the joint still (Fig. 8.4).

When a muscle is held at a constant length, such as when a joint is held still, there is usually enough stretch of the muscle's spindles to produce a steady frequency of action potential discharge in the sensory neuron output of the muscle spindle. This steady-state discharge gives the muscle spindle organ the ability to inform the CNS not only about a subsequent lengthening of the muscle, which would produce a proportional increase in action potential discharge, but about some shortening of the muscle as well, which would produce a proportional decrease in action potential discharge from the steady state (see Fig. 8.4).

Vet-ebooks.com



• **Fig. 8.2** The muscle spindle receptor is an encapsulated group of specialized (intrafusal) skeletal muscle fibers supplied with both motor and sensory innervation. *Top,* Longitudinal section through a skeletal muscle showing that the encapsulated muscle spindles are oriented parallel to the more numerous extrafusal fibers of the muscle. The ends of the muscle spindle are attached to the extracellular matrix of the extrafusal fibers. *Bottom,* Higher magnification view of a transverse section through a muscle spindle. Intrafusal fibers can be seen within the spindle's tissue capsule. These fibers are fewer, shorter, and more slender than the surrounding extrafusal fibers. A portion of the spindle's innervation can also be seen. (Images courtesy of Dr. Tom Caceci, Department of Biomedical Sciences and Pathobiology, College of Veterinary Medicine, Virginia Tech.)



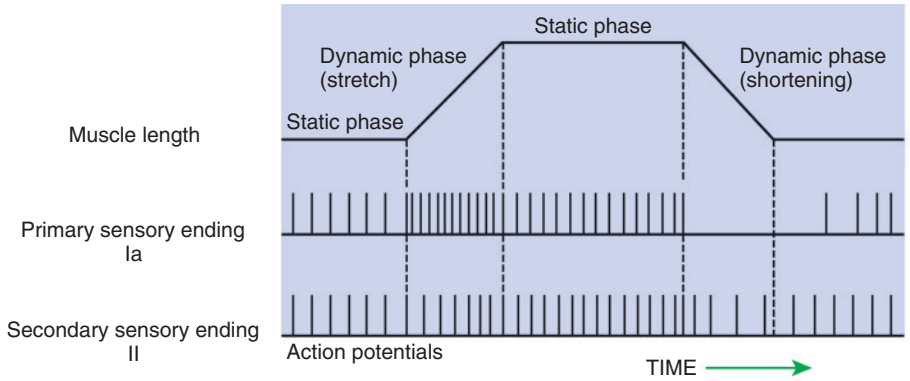
• **Fig. 8.3** The muscle spindle stretch reflex (illustrated here as the *knee jerk reflex*) begins when the spindle receptor organ is stretched. This causes action potentials on the receptor's sensory neurons, which in turn cause excitatory postsynaptic potentials on the α motor neurons returning to the extrafusal muscle fibers of that same muscle. Action potentials on the α motor neurons cause extrafusal muscle fibers to contract, and the knee extends ("jerks"). Through an inhibitory interneuron, the α motor neurons to the antagonist muscles are simultaneously inhibited.

Muscle Stretch and Action Potentials Along Spindle Sensory Neurons Lead to Reflex Contraction of the Extrafusal Muscle Fibers

The sensory output neurons of the muscle spindle enter the CNS, where they make excitatory, monosynaptic connections with α motor neurons that return to the extrafusal fibers of the same muscle (see Fig. 8.3). Therefore stretching a given muscle can lead to a rapid reflex contraction of that same muscle, bringing it back to its original length. Stretching the muscle lengthens the intrafusal muscle fibers of the spindle, increasing the frequency of action potential discharge along the sensory output neurons of the spindle. This leads to an increase in action potential frequency in the α motor neurons on which the spindle sensory neurons synapse. This produces contraction of the extrafusal fibers innervated by those α motor neurons, which results in contraction (shortening) of the muscle. Contraction of the muscle results in a shortening of the muscle spindle's equatorial region. This eventually reduces the frequency of action potentials occurring on the spindle sensory neurons to the prestretch level, terminating the response. (The cycle is a classic negative-feedback system.)

The reflex just described can be elicited by striking the patellar tendon (insertion tendon of quadriceps muscle) with a blunt object such as a reflex hammer (plexor). Because this tendon goes over a *pulley* (the patella), hitting this tendon results in a longitudinal stretch of the whole quadriceps muscle, thus also stretching the muscle spindles. Action potentials from spindle sensory neurons

Vet-ebooks.com



• **Fig. 8.4** The muscle spindle can signal the steady-state length of the muscle, as well as the onset and velocity of stretch. When the muscle is stretched, both type Ia and type II spindle sensory neurons have a higher action-potential firing rate at the muscle's new static length. During the dynamic stretching phase, the action-potential firing rate of the type Ia spindle sensory neuron increases rapidly and in proportion to the velocity of stretch. Spindle sensory neurons can also register a decreased steady-state length of the muscle, but type Ia and type II neurons display differential sensitivity to the dynamic phase of shortening. (From Brodal P. *The Central Nervous System: Structure and Function.* 2nd ed. New York: Oxford University Press; 1998.)

go to the lumbar spinal cord, by way of the dorsal roots, and cause excitatory postsynaptic potentials (EPSPs) on the α motor neurons of the motor units that return to the quadriceps muscle (see Fig. 8.3). This causes contraction of the quadriceps muscle and extension of the knee joint, and is an example of the *muscle stretch reflex,* or *myotatic reflex.* When applied to the quadriceps muscle, it is called the *knee jerk reflex,* but the mechanisms are present in almost all muscles. However, this is the muscle from which it is easiest to evoke the stretch reflex because it is one of the few whose tendon goes over a sesamoid pulley before inserting on the next bone. Because of the pulley under the tendon, a lateral deflection of the tendon, as from a reflex hammer, results in a longitudinal stretch of the muscle and thus the reflex. Hitting other tendons only moves the muscle belly laterally and does not easily result in the stretch reflex. Therefore in the clinical neurological examination of most animals, the knee jerk reflex is the most commonly evoked muscle stretch reflex.

The muscle spindle organ and stretch reflex allow the CNS to make automatic, usually unconscious, adjustments to muscle stretch imposed by small changes in body position or the weighting of a muscle. Such adjustments can return the muscle to its original length, often returning a joint to its original position. Clinical examination of the stretch reflex provides clues about the peripheral or CNS integrity of its sensory and motor components. Given that descending projections from brain to spinal cord can modulate the activity of elements of spinal reflex circuits, testing stretch reflexes may also reveal information regarding the status of those descending projections, as in hyperreflexia (exaggerated reflex response).

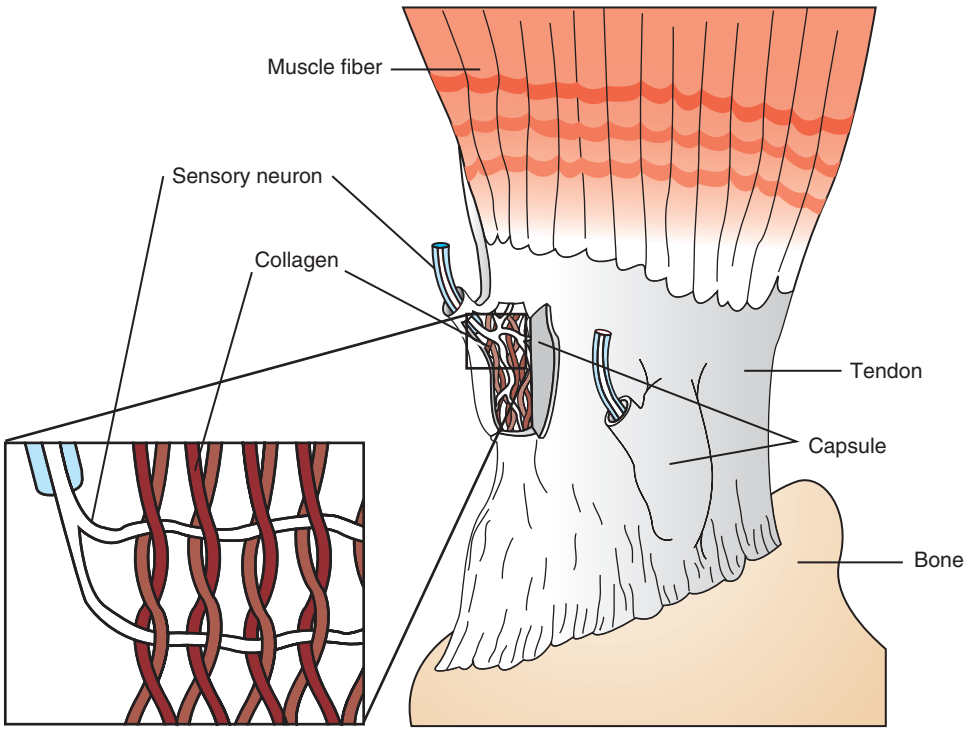
When the stretch reflex acts to return a joint to its original position, the antagonist of the stretched muscle must relax for the joint to move. Therefore in the stretch reflex, some terminal branches of individual spindle sensory neurons do not synapse directly on the α motor neurons of the stretched muscle but rather synapse on inhibitory spinal interneurons (see Fig. 8.3). These neurons, which lie completely within the spinal cord, are also activated by the stretch of the muscle. When they fire action potentials, however, they cause inhibitory postsynaptic potentials (IPSPs) on α motor neurons that innervate the antagonist of the stretched muscle. This prevents the contraction of the antagonistic muscle.

The Central Nervous System Can Control Spindle Sensitivity Directly Through the Gamma (γ) Motor Neurons

Contraction of the extrafusal muscle fibers is controlled by the larger α motor neurons; intrafusal muscle fibers of the spindle are controlled by the smaller γ motor neurons. The γ motor neurons innervate the intrafusal muscle fibers at their polar ends (see Figs. 8.1 and 8.3), the regions containing contractile protein. Action potentials on the γ motor neurons cause shortening of the polar regions of the intrafusal muscle fibers, stretching the equatorial portion.

An important function for this motor innervation of a receptor organ is to regulate the sensitivity of the muscle spindle. Shortening of a gross muscle resulting from initiation of extrafusal muscle fiber contraction has the potential to slacken the intrafusal muscle fibers given their parallel relationship to the extrafusal fibers. This would severely limit the ability of the muscle spindle to transduce stretch. However, this does not normally occur because contraction of the polar regions of intrafusal fibers, resulting from γ motor neuron activation, is initiated concurrently with shortening of extrafusal fibers caused by α motor neuron activation. This allows the spindle receptor organ to remain taut and sensitive to sudden stretches of the gross muscle over the entire range of its length. This γ motor neuron control mechanism can also function to differentially regulate the sensitivity of the muscle spindle, depending on the type of movement to be made (e.g., novel and unpredictable vs. stereotypical). There are actually two types of γ motor neurons; one regulates the sensitivity of the muscle spindle to the dynamic phase of stretch (γ_D; gamma dynamic) and one regulates sensitivity to steady-state length (γ_S; gamma static). Chapter 10 describes how coactivation of both the α and γ motor neurons allows the

Vet-ebooks.com



• **Fig. 8.5** The Golgi tendon organ is located in the tendons of skeletal muscle, in series with the extrafusal fibers. It detects tension in the tendon, produced by muscle contraction, and sends information about this tension to the central nervous system. Sensory neuron branches of the organ are intertwined among braided collagen fibrils (*inset*), which fold up and pinch the neural branches when tension develops in the tendon.

brain to test whether the amount of contraction intended by the brain was what actually occurred.

The Golgi Tendon Organ Lies in Series Between Muscle and Tendon and Detects Muscle Tension

Each Golgi tendon organ is a slender capsule, at the junction between muscle and tendon, in series with 10 to 15 extrafusal skeletal muscle fibers (Figs. 8.1 and 8.5). The capsule of each tendon organ contains a complement of braided collagen fascicles, among which the branches of a single sensory neuron are intertwined. This sensory neuron, as with those of the muscle spindle, carries action potentials to the CNS by way of peripheral nerve and dorsal root. The Golgi tendon organ has no motor innervation.

Because the Golgi tendon organ is in series with a group of extrafusal fibers and the tendon, when the extrafusal fibers shorten during contraction, tension is applied to the tendon organ. This causes the braided collagen fibrils of the organ to tighten and squeeze the endings of the sensory neuron. Action potentials are therefore generated and sent to the CNS along the sensory neuron at a frequency proportional to the tension developed by the muscle. In contrast, the muscle spindle is arranged in parallel with the extrafusal muscle fibers, and when extrafusal fibers contract, the action potential frequency of spindle sensory neurons is reduced.

When action potentials from spindle sensory neurons reach the CNS, as mentioned earlier, they monosynaptically produce EPSPs in the α motor neurons returning to the same muscle. Action potentials along sensory neurons from Golgi tendon organs have the opposite effect: they activate inhibitory interneurons, polysynaptically producing IPSPs on α motor neurons to the same muscle. This leads to a reduced extrafusal muscle fiber contraction.

Free (Non-organ) Sensory Receptors in Joints and Muscles Can Provide Information About Joint Position, Joint Movement, and Pain-Inducing Stimuli of Joints and Muscles

Input from muscle sensory organs is not the only source of muscle sensory information to the CNS and is not the only source of sensory information about movement of body parts. Non-organ peripheral endings of sensory neurons, with central projections to the spinal cord, can be found in joint capsules and in ligaments around the joints. Some of these receptors can respond to changes in tension of the joint capsule and can therefore provide information about position of the joint. Some can also respond to the velocity of joint movement. Some joint-associated sensory endings are activated by strong mechanical stimuli, or inflammatory mediators, associated with pain sensations (e.g., arthritic pain). Pain-mediating, non-organ sensory endings can also be found in muscle and may contribute to the sensation of muscle soreness. Interestingly, there is a population of non-organ sensory receptors that can be found in both muscles

Vet-ebooks.com

and joints that are thought to trigger circulatory and respiratory reflexes associated with onset of body movement.

The two skeletal muscle receptor organs discussed in this chapter provide the CNS with the most important and vital information about muscle length (the spindle) and muscle tension (the Golgi tendon organ), and through their reflex circuitry, they help to keep these parameters within optimal ranges. The information these receptors provide, along with information from receptors of skin and joints, is essential if the CNS is to coordinate posture and locomotion.

Acknowledgment

The authors thank Dr. Tom Caceci for persevering in the search through his histological specimens to capture the beautiful muscle spindle images of Fig. 8.2.

CLINICAL CORRELATIONS

Femoral Nerve Mononeuropathy

History

You examine an 8-year-old male golden retriever. The owner complains that the dog cannot bear weight on the right rear leg.

Clinical Examination

Physical examination deficits are limited to the right rear leg, where you find that the quadriceps femoris muscles are much smaller than those of the left rear leg. The dog cannot bear weight on the right rear leg because the right quadriceps femoris muscles are paralyzed. When you tap on the left patellar tendon with a reflex hammer, the knee briskly extends (the knee jerk or muscle stretch reflex). However, when you tap on the right patellar tendon, no movement occurs.

Comment

The quadriceps femoris muscle group is one of the major antigravity muscle groups of the leg causing the stifle joint (knee joint) to extend. The paralysis in this animal's quadriceps muscle is the reason why he cannot bear weight on the leg. The small size of the right quadriceps muscle is caused by atrophy, or muscle wasting, which in turn is caused by the loss of α motor neurons to the extrafusal muscle fibers in the quadriceps muscle belly (see Chapter 9). This would also cause a loss of the muscle stretch reflex, because even though the spindle detected the stretch of the muscle belly caused by either gravity or the reflex hammer, the α motor neurons returning to the quadriceps muscle are unable to signal the muscle to contract, thus severing the reflex arc. This syndrome could occur if the femoral nerve is damaged by a tumor or trauma. If the pathological lesion were in the peripheral nerve rather than only in the ventral roots, there would likely be some sensory loss in addition to the motor deficits.

Treatment

This is a *femoral nerve mononeuropathy.* Its treatment depends on the cause of the nerve damage (e.g., trauma, neoplasia, inflammation).

Obturator Nerve Paralysis in a Cow Post-Calving

History

A client calls you about a 2-year-old Holstein cow that calved a few days ago. The calving was difficult (dystocia), and the clients had to pull the calf. Since the calving, the cow has been laying down a lot, and is stiff and uncoordinated when she walks in the barn. She has almost fallen a few times.

Clinical Examination

On examination, she appears to be eating and drinking well. Her temperature, pulse, and respiratory rate are all within normal limits. You watch her get up and lie down, and she appears uncoordinated on her right hindlimb. When she is standing, she maintains a wide leg base (abducted). The next time she lies down, you test some of her muscle reflexes, including some of the muscle spindle stretch (myotatic) reflexes. Absence or reduction of such spinal reflexes, and whether the effect is in one muscle group or bilateral, can provide clues to whether there is damage to sensory or motor components of peripheral nerve or whether damage is located in the spinal cord. Exaggeration of reflexes may indicate loss of descending inhibitory control from more rostral parts of the CNS. The quadriceps stretch reflex (knee jerk reflex; see this chapter) is tested, although because the animal seems to be able to support its weight on the limb, you do not expect this to be affected. The cranial tibial stretch reflex is also tested by striking the belly of this muscle just below the proximal end of the tibia. The response of this hock flexor is used to assess integrity of the peroneal branch of the sciatic nerve. The pelvic limb flexor reflex is a response to noxious pinching of the skin of the distal limb, involves all the flexor muscles of the limb, and assesses the integrity of the sciatic nerve. All of these reflexes appear normal.

Comment

Based on her history and clinical signs, you suspect obturator nerve paralysis from the trauma during calving. The obturator nerve innervates the adductor, pectineus, and gracilis muscles, which collectively affect hindlimb adduction (drawing nearer to the midline) and hip movements. Thus the cow can stand and walk but has problems keeping the limb in a normal position, particularly if the surface is slippery or she is running.

Treatment

Clinical signs are due to inflammation and damage of the obturator nerve. Typically, cows are not treated, and the degree of recovery is based on the extent of inflammation and nerve damage present. More severe cases can be treated with anti-inflammatory drugs or steroids to limit inflammation.

Practice Questions

1. If the distance between the origin and insertion tendons is increased (the muscle is stretched), what happens to the frequency of action potentials along the sensory axons leaving the muscle spindles in that muscle?
- **a.** Increases
- **b.** Decreases
- **c.** Does not change

2. Activation of the Golgi tendon organ of a given muscle:
- **a.** Monosynaptically produces EPSPs on the α motor neuron that returns to that muscle.
- **b.** Is most effectively produced by lengthening of the extrafusal fibers of that muscle.
- **c.** Polysynaptically produces IPSPs on the α motor neuron that returns to that muscle.
- **d.** Polysynaptically produces EPSPs on the α motor neuron that returns to that muscle.
- **e.** Activates motor neurons that return to the Golgi tendon organ itself.

Vet-ebooks.com

3. Which of the following is *not* characteristic of the muscle spindle?
- **a.** Encapsulated intrafusal fibers
- **b.** Sensitivity to muscle tension
- **c.** Sensitivity to dynamic stretching of the muscle
- **d.** Lying parallel to the extrafusal muscle fibers
- **e.** Sensitivity to steady-state length of the muscle

4. Gamma (γ) motor neurons:
- **a.** Innervate and produce contraction of the equatorial (middle) region of an intrafusal fiber.
- **b.** Have their cell bodies in the dorsal horn of the spinal cord.
- **c.** Are never activated at the same time as α motor neurons.
- **d.** Can regulate the sensitivity of the muscle spindle sensory organ.
- **e.** Innervate and regulate the sensitivity of the Golgi tendon organ.

Bibliography

Bewick GS, Banks RW. Mechanoreception in the muscle spindle. *Pflugers Arch*. 2015;467:175–190.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Lorenz MD, Coates JR, Kent M. *Handbook of Veterinary Neurology*. 5th ed. Philadelphia: Saunders; 2011.

MacKay RJ, Van Metre DC. Diseases of the nervous system. In: Smith BP, eds. *Large animal internal medicine*. 5th ed. St Louis: Mosby Elsevier; 2015.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Schatzberg SJ. Neurologic examination and neuroanatomic diagnosis. In: Ettinger SJ, Feldman EC, Cote E, eds. *Textbook of Veterinary Internal Medicine*. 8th ed. St Louis: Saunders Elsevier; 2017.

Vet-ebooks.com

9

The Concept of Lower and Upper Motor Neurons and Their Malfunction

BRADLEY G. KLEIN

KEY POINTS

1. The lower motor neuron is classically defined as the alpha (α) motor neuron.
2. Disease of lower motor neurons causes stereotypical clinical signs.
3. Upper motor neurons lie completely in the central nervous system and control lower motor neurons.
4. Signs of upper motor neuron disease differ from signs of lower motor neuron disease.

The majority of veterinary patients with neurological disease display some abnormality of posture and locomotion. The abnormalities range from weakness or paralysis to spasticity, rigidity, and convulsions. The goal of the diagnostic process for such patients is to determine the location, the extent, and the cause of the lesion. Central to diagnostic logic in neurology is deciding whether the patient's lesion is located in the lower motor neurons or the upper motor neurons. (There are two other possible locations of lesions causing movement disorders: the neuromuscular junction and skeletal muscle.)

This chapter defines lower and upper motor neurons because these concepts are useful in understanding the physiology of posture and locomotion, and are essential in locating pathological processes in the nervous system. Malfunctions of these two neuron populations are also described briefly.

The Lower Motor Neuron Is Classically Defined as the Alpha (α) Motor Neuron

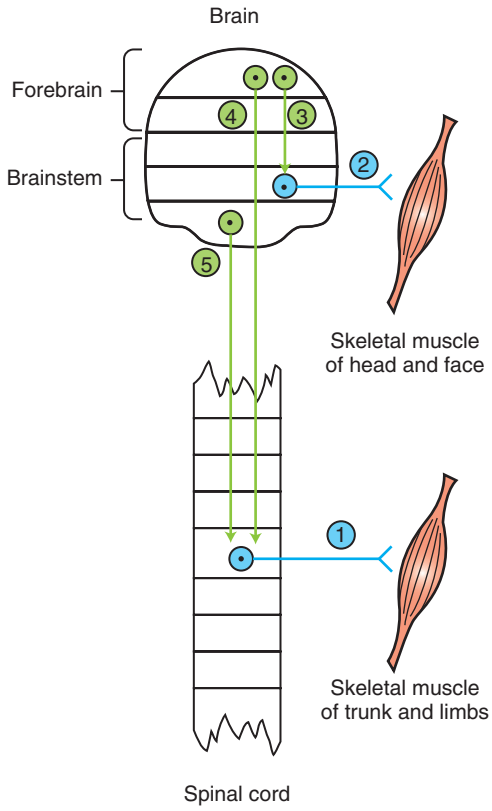
The concept of a lower motor neuron is decades old in neurology. The alpha (α) motor neuron is classically defined as a neuron whose cell body and dendrites are located in the central nervous system (CNS) and whose axon extends out through the peripheral nerves to synapse with the extrafusal skeletal muscle fibers (Fig. 9.1). The cell bodies of these neurons are located either in the ventral horn of spinal cord gray matter or in cranial nerve nuclei of the brainstem. This is the "final common path" through which CNS commands to a skeletal muscle are channeled to produce movement. Gamma (γ) motor neurons which innervate intrafusal fibers of muscle spindles, as well as pre- and postganglionic autonomic neurons (see Chapter 13), are included in the definition of lower motor neurons by some authors. However, the vast majority of clinical signs caused by lower motor neuron disease can currently be explained by the loss or malfunction of the α motor neuron.

Disease of Lower Motor Neurons Causes Stereotypical Clinical Signs

Regardless of the pathological basis for disease of lower motor neurons, a stereotypical set of clinical signs results in the skeletal muscles they innervate.

- *Paralysis* or *paresis.* Disease of the α motor neurons usually prevents the neurons' action potentials from reaching the neuromuscular junction. Therefore despite the brain's command to the muscle to contract, the message cannot reach the muscle, and paralysis is the result. In fact, such paralysis can be so complete that the adjective *flaccid* is used to describe paralysis in which no muscle contraction occurs. Because not all the α motor neuron axons of a peripheral nerve may be affected by an insult, and because muscles can be supplied by axons of more than one spinal nerve, paralysis may be incomplete. This symptom is referred to as *paresis.*
- *Atrophy.* Atrophy is the shrinking or wasting of skeletal muscle mass distal to the lower motor neuron lesion (Fig. 9.2). This occurs within days of the injury to a nerve. The exact origins of this atrophy are controversial. However, evidence indicates that the reduced frequency of muscle stimulation caused by α motor neuron insult and the resulting reduced use of the muscle trigger reductions in muscle protein synthesis and increases in muscle proteolysis. There is evidence that activation of the ubiquitin-proteasome and autophagy-lysosome proteolytic pathways is an important substrate underlying this muscle breakdown. The reduced muscle activity resulting from lower motor neuron damage appears to affect the up- or downregulation of several muscle atrophy-related genes, termed "atrogenes," and the products of some of these genes can affect elements of the aforementioned proteolytic pathways. Furthermore, there is evidence of cross-talk between elements of the proteolytic and protein synthesis pathways during muscle atrophy. Some studies have suggested that the magnitude of denervation atrophy can be reduced by direct electrical stimulation of the muscle itself or by manually imposed repetitive stretching, but their clinical effectiveness is still a matter of controversy.
- *Loss of segmental and intersegmental reflexes.* Segmental and intersegmental reflexes require a viable α motor neuron in the reflex arc for the reflex response to occur (see Chapter 7).

Vet-ebooks.com



• **Fig. 9.1** General organization of lower and upper motor neurons. *Blue,* Lower motor neurons typically originate in the ventral horn of the spinal cord (*neuron 1*) or in cranial nerve nuclei (*neuron 2*), and synapse within skeletal muscle. *Green,* Upper motor neurons typically originate in the brain and project to and control lower motor neurons. Upper motor neurons usually belong to the corticobulbar (*neuron 3*), corticospinal (*neuron 4*), or descending brainstem motor (also called bulbospinal, *neuron 5*), pathways. *Arrowheads* indicate that local synapses may be interposed between the upper and lower motor neurons. *Horizontal lines* delimit either spinal cord segments (e.g., L1, L2) or major brain divisions (e.g., medulla, pons).

Therefore such reflexes as the muscle stretch (knee jerk) reflex and the toe-pinch withdrawal (nociceptive) reflex, as well as the proprioceptive positioning reaction, fail to occur because the motor neuron portion of the arc that activates the skeletal muscle is gone.

- *Electromyographic changes.* Within a few days of damage to α motor neurons, abnormal electrical activity of the muscle can be observed on an electromyogram (see Chapter 6).

Damage of α motor neurons occurs often on insult to a peripheral nerve that also contains axons of sensory neurons. Therefore there may be an accompanying loss of sensory modalities, although this is not a cardinal sign of lower motor neuron damage.

Upper Motor Neurons Lie Completely in the Central Nervous System and Control Lower Motor Neurons

Upper motor neurons are the neurons of the CNS that influence the lower motor neurons. They are typically considered the neurons of origin of the corticospinal (cerebral cortex to spinal cord), corticobulbar (cerebral cortex to brainstem), and descending brainstem motor (brainstem to spinal cord; also called bulbospinal) pathways (see Chapter 10). Upper motor neurons send axons down to the spinal cord or into the brainstem to control the lower motor neurons (see Fig. 9.1). Voluntary movement is initiated by upper motor neurons with a cortical origin.



• **Fig. 9.2** Atrophy of the right side of the tongue in a golden retriever due to a meningioma that affected the right hypoglossal nerve roots. (From De Lahunta A, Glass E. *Veterinary Neuroanatomy and Clinical Neurology.* 3rd ed. Philadelphia: Saunders; 2009.)

Signs of Upper Motor Neuron Disease Differ From Signs of Lower Motor Neuron Disease

Lesions of upper motor neurons cause clinical signs that differ significantly from those produced by lower motor neuron disease, although paralysis/paresis may be seen in both scenarios.

- *Inappropriate movement.* Lesions of upper motor neurons can cause a variety of movement disorders depending on the location of the lesion. Spinal cord disease, affecting portions of upper motor neurons projecting to the cord, often causes various degrees of weakness below the lesion. Disease of the brain that affects upper motor neurons may cause rigidity, seizures, circling gaits, and other inappropriate movements. More specific examples of this general category are presented in Chapters 10, 11, and 12 on the central control of movement, the vestibular system, and the cerebellum, respectively.
- *No atrophy.* Because the lower motor neuron is intact, the muscle does not typically atrophy. (Modest disuse atrophy may develop much later.)
- *Retained but exaggerated segmental reflexes.* Because the neuronal circuitry of the segmental reflex arc (see Chapter 7) is not interrupted in upper motor neuron disease, reflexes such as the muscle stretch and toe-pinch withdrawal are retained, whereas in lower motor neuron disease, reflexes are depressed or lost. Because upper motor neurons are normally capable of exerting significant inhibitory control over spinal reflexes, however,

Vet-ebooks.com

damage to these neurons can decrease this inhibition, resulting in exaggeration of the reflex response (*hyperreflexia*).

- *Normal electromyogram.* Because the muscle is not atrophied and the lower motor neurons are intact, the electrical activity of the muscle appears normal.

The following clinical correlations illustrate common examples of lower and upper motor neuron disease. Before proceeding to Chapter 10, the reader should understand these concepts and why these dogs have the clinical signs mentioned.

Acknowledgement

The authors thank Dr. Karen Inzana for her thoughtful comments on material in this chapter.

CLINICAL CORRELATIONS

Lower Motor Neuron Disease

History

A 2-year-old male German short-haired pointer was admitted to the local veterinary clinic. His vaccinations were current, and the dog had no history of contributing prior illness.

A few days before admission, the dog had a fight with a skunk. For 48 hours before admission, an ascending paralysis developed, characterized initially by weakness and then by lack of voluntary movement of first the back legs, then the front legs. No barking was noticed during the illness. The dog was able to control his bladder and bowel, and could move his head.

Clinical Examination

On admission, the dog was unable to bear weight on any of his four legs. Other than an elevated respiratory rate, physical examination deficits were limited to the nervous system. The dog was able to eat, drink, and move his head. A severe paralysis was noted in all legs, and no motor response to a toe pinch or tapping of the quadriceps tendon could be elicited. There was widespread atrophy of the muscles of all four legs, as well as those of the thorax and abdomen. The dog did seem to be aware of painful stimuli (deep pain response). There were no cranial nerve deficits. Results of routine blood cell counts and serum chemistry results were within normal limits.

Comment

Generalized atrophy, paralysis, and loss of segmental reflexes indicate widespread, bilateral loss of lower motor neuron function. Fortunately, the disease has spared the muscles of the head and the diaphragm, although the dog's elevated respiratory rate indicates an attempt to compensate for paralysis of some of the respiratory muscles. A clinical diagnosis of *polyradiculoneuritis* (coonhound paralysis) was made. This disease is often preceded by the bite of another animal. Pathological changes are found predominantly in the ventral roots of the spinal cord where the axons of the lower motor neurons leave the spinal cord. The dorsal roots are usually spared, which explains this dog's apparent ability to feel pain. The clinical signs are those of widespread lower motor neuron disease. The syndrome resembles Guillain-Barré syndrome in humans, and both syndromes have been suggested to be autoimmune in origin.

Treatment

Animals with this form of paralysis usually recover spontaneously. Good nursing care is essential during the illness. A respirator may be necessary, temporarily, if respiratory paralysis occurs.

Upper Motor Neuron Disease

History

A 5-year-old male dachshund is brought to a local veterinary clinic. His vaccination history is current, and the dog has had no contributing past medical or surgical illnesses. Two days before admission, he seemed to be in pain. Throughout the next day, the dog became progressively weak in the hind legs.

Clinical Examination

Physical examination abnormalities were limited to the nervous system. The dog was bright, alert, responsive, and able to bear weight normally on the front legs. However, he was weak and unsteady on the hind legs. No atrophy was apparent. All cranial nerve reflexes were normal as were the spinal segmental reflexes of both front and hind legs. Intersegmental responses and the proprioceptive positioning reaction were normal in the front legs but absent in the hind legs (see Chapter 7). Results of a complete blood cell count and serum chemistry analysis were within normal limits.

Comment

Failure to exhibit the normal proprioceptive placing reaction of a paw indicates a lesion somewhere along the sensory or motor routing for this response. This routing includes the peripheral nerves for that limb, the spinal cord rostral to that limb on the same side, and the contralateral side of the brain. However, the absence of atrophy and the retention of segmental reflexes in the affected limbs indicate that the lower motor neurons, neuromuscular junction, and skeletal muscle are normal, and that this is an upper motor neuron disease. Because only the hind limbs are affected by weakness and exhibit a deficit in proprioceptive placing, the cervical spinal cord and brain must be normal because motor commands to the front legs are transmitted reliably. Therefore the lesion must be between the front and hind limbs. This is a typical history and a typical clinical presentation for a dog with a *herniated intervertebral disk.*

Treatment

Treatment and prognosis depend on the severity of the spinal cord trauma. Medical management is aimed at reducing edema, vasospasm, inflammation, and other metabolic consequences of the disease that worsen the damage to the spinal cord. When surgery is indicated by the severity of the trauma, the goal is to relieve spinal cord compression. With appropriate medical and surgical management, many dogs recover useful spinal function.

Practice Questions

1. Which of the following would *not* be considered an upper motor neuron?
- **a.** Bulbospinal motor neurons
- **b.** Corticospinal motor neurons
- **c.** Ventral horn α motor neurons of the spinal cord
- **d.** Corticobulbar motor neurons

2. You examine a dog that is unable to stand and bear weight on the right rear leg. The right rear leg is much smaller in diameter than the left rear leg. Pinching the toe on the left rear leg results in withdrawal of the left rear leg, but pinching the toe on the right rear leg results in no movement of the right rear leg. The proprioceptive placing response in the left rear leg is normal, but in the right rear leg it is absent. Where is this dog's pathological lesion?
- **a.** Lower motor neuron to the right rear leg
- **b.** Lower motor neuron to the left rear leg
- **c.** Upper motor neuron controlling the right rear leg
- **d.** Upper motor neuron controlling the left rear leg
- **e.** Neuromuscular synapse of the left rear leg

Vet-ebooks.com

3. You examine a dog that is bright, alert, and responsive. She can stand and bear weight on both front legs, but she cannot stand or bear any weight on the back legs. Her knee jerk and toe-pinch withdrawal reflexes are normal in all four legs. There is no atrophy. The proprioceptive positioning response is normal in the front legs but absent in both rear legs. Injecting acetylcholinesterase-inhibiting drugs causes no change in the clinical signs. Where is this dog's pathological lesion most likely located?
- **a.** Neuromuscular junction
- **b.** Cervical spinal cord (spinal cord of the neck)
- **c.** Spinal cord between the front and rear legs (thoracolumbar spinal cord)
- **d.** Lower motor neurons to the rear legs
- **e.** Brainstem

4. You examine a dog that is bright, alert, and responsive but unable to stand on any of the four legs. Toe-pinch and knee jerk local (segmental) reflexes are normal in all four legs. There is no atrophy. The proprioceptive positioning response is absent in all four legs. Injecting an acetylcholinesterase-inhibiting drug does not change the clinical signs. Where is this dog's pathological lesion most likely located?
- **a.** Cervical spinal cord (spinal cord in the neck)
- **b.** Spinal cord between the front and rear legs (thoracolumbar spinal cord)
- **c.** Lower motor neurons to all four legs
- **d.** Neuromuscular junction

5. You are presented with a horse that is unable to stand or support any weight on the hind legs. You electrically stimulate both the sciatic and the femoral nerves with a sufficient stimulus, but neither stimulation results in muscular contraction. However, direct stimulation of both the gastrocnemius and the quadriceps femoris muscles of the rear leg results in muscular contraction. From these observations, what do you logically conclude to be the location of this horse's pathological lesion?
- **a.** Upper motor neurons controlling the rear legs
- **b.** Lower motor neurons to the rear legs
- **c.** Neuromuscular synapses of the rear legs
- **d.** Muscles of the rear legs
- **e.** Either b or c

6. You examine a cat that cannot bear weight on the hind legs. The cat is bright, alert, and responsive. Atrophy is present in the back legs. Cranial nerve reflexes are within normal limits, as are segmental reflexes and the proprioceptive positioning responses of the front legs. Knee jerk and toe-pinch withdrawal reflexes are absent in the hind legs. What is the most likely location for this cat's pathological lesion?
- **a.** Brainstem
- **b.** Cervical spinal cord (spinal cord in the neck)
- **c.** Thoracolumbar spinal cord (spinal cord between the front and rear legs)
- **d.** Lower motor neurons to the front legs
- **e.** Lower motor neurons to the hind legs

Bibliography

Agata N, Sasai N, Inoue-Miyazu M, et al. Repetitive stretch suppresses denervation-induced atrophy of soleus muscle in rats. *Muscle Nerve*. 2009;39:456–462.

Bonaldo P, Sandri M. Cellular and molecular mechanisms of muscle atrophy. *Dis Model Mech*. 2013;6(1):25–39.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Saunders; 2015.

Dow DE, Dennis RG, Faulkner JA. Electrical stimulation attenuates denervation and age-related atrophy in extensor digitorum longus muscles of old rats. *J Gerontol A Biol Sci Med Sci*. 2005;60(4):416–424.

Kandarian SC, Jackman RW. Intracellular signaling during skeletal muscle atrophy. *Muscle Nerve*. 2006;33:155–165.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Lorenz MD, Coates JR, Kent M. *Handbook of Veterinary Neurology*. 5th ed. Philadelphia: Saunders; 2011.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. Sunderland, Mass: Sinauer; 2017.

Salvini TF, Durigan JLQ, Peviani SM, Russo TL. Effects of electrical stimulation and stretching on the adaptation of denervated skeletal muscle—implications for physical therapy. *Rev Bras Fisioter*. 2012;16(3):175–183.

Sandri M. Protein breakdown in muscle wasting: role of autophagy-lysosome and ubiquitin-proteasmome. *Int J Biochem Cell Biol*. 2013;45:2121–2129.

Tisdale MJ. Is there a common mechanism linking muscle wasting in various disease types? *Curr Opin Support Palliat Care*. 2007;1(4):287–292.

Vet-ebooks.com

10

The Central Control of Movement

BRADLEY G. KLEIN

KEY POINTS

1. The central nervous system structures that control movement have a hierarchical organization.
2. The spinal cord is the most caudal and simplest level of the movement control hierarchy.
3. Brainstem upper motor neuron pathways are the source of all descending motor system input to the spinal cord, except for one other major pathway.
4. Medial and lateral descending brainstem motor pathways respectively control proximal muscles of posture and more distal muscles of skilled movement.
5. The reticulospinal and vestibulospinal tracts are medial brainstem motor pathways important for keeping the body upright against the pull of gravity.
6. The rubrospinal tract is a lateral brainstem motor pathway that can control distal limb musculature associated with skilled movement.
7. The corticospinal (pyramidal) tract is a direct projection from cerebral cortex to spinal cord responsible for the most skilled voluntary movements of mammals.
8. In the spinal cord, the corticospinal tract has a massive lateral component controlling the distal musculature and a minor medial component controlling the axial and proximal musculature.
9. The motor cortices in or adjacent to the frontal lobe, the highest level of the motor control hierarchy, consist of three different functional regions.
10. Corticospinal tract coactivation of both alpha (α) and gamma (γ) lower motor neurons may help with small automatic corrections of voluntary movements.
11. The motor system shares some organizational principles with sensory systems.
12. The basal ganglia and cerebellum modulate the activity of motor system components for the respective selection and adjustment of movement.

Unlike the sensory systems, most of which transform physical energy into neural information, the motor system transforms neural information into physical energy. All movement is the result of the contraction of varying numbers of extrafusal skeletal muscle fibers within varying numbers of motor units (see Fig. 6.8). These extrafusal muscle fibers do not contract until commanded to do so by the alpha (α) lower motor neuron. The α motor neuron, in turn, does not send such an action potential command until signaled to do so by descending upper motor neurons (see Fig. 9.1) or from incoming sensory neurons (or interneurons) in a reflex arc.

Movement can be divided into two general forms. The first is a largely learned, voluntary, conscious, and skilled form, often dominated by flexor muscle activation. The second form is characterized by postural, antigravity muscle activity that is generally subconscious, involuntary, and dominated by extensor muscle contraction. The skilled movement results from fairly discrete contraction of a few muscle groups, many of which are distal to the spinal column. The maintenance of posture often includes longer-term contraction of larger groups of muscles, many of which are located closer (proximal) to the spinal column. Correspondingly, in the spinal cord gray matter, the α motor neurons that control the more distal muscles tend to be located laterally; those controlling the more proximal and axial muscles for posture are located more medially.

Initiating the learned, skilled, voluntary movement of the distal musculature is largely the responsibility of a subgroup of upper motor neuron tracts that project through more lateral regions of the spinal cord white matter and terminate in lateral regions of the spinal cord gray matter. Initiating antigravity and postural muscle activity is the responsibility of upper motor neuron tracts associated with more medial regions of the spinal cord white and gray matter, respectively. This lateral-medial distinction is a significant organizational principle in central nervous system (CNS) motor control. Skilled, voluntary movement of the distal musculature is primarily controlled by a lateral system of lower motor neurons and upper motor neuron spinal tracts. More medial systems of such neurons and tracts primarily control postural and antigravity activity of the proximal and axial musculature.

The Central Nervous System Structures That Control Movement Have a Hierarchical Organization

Another organizational principle of the neural control of movement is that it consists of a hierarchy. Generally, simpler movements or movement patterns are organized by more caudal parts of the CNS (Fig. 10.1, *bottom portions*), and more complex and skilled patterns are organized by progressively more rostral regions (see Fig. 10.1, *top portions*).

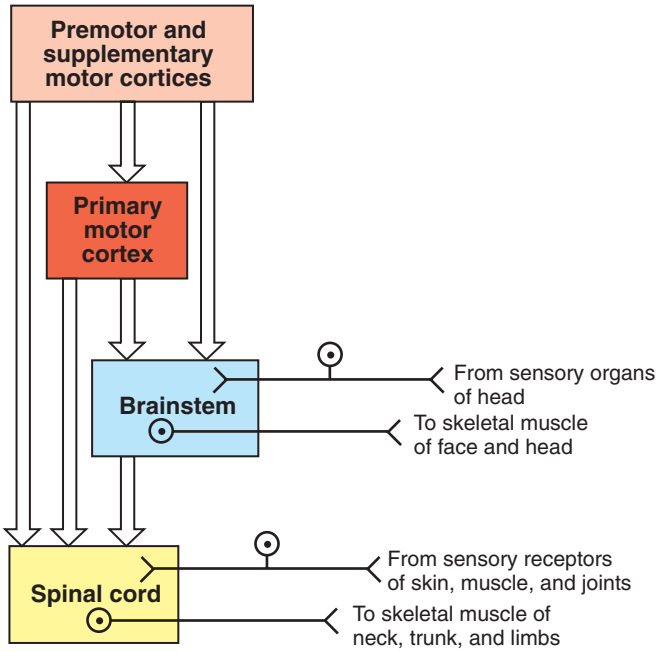
Vet-ebooks.com

The Spinal Cord Is the Most Caudal and Simplest Level of the Movement Control Hierarchy

The spinal cord contains the lower motor neurons that synapse upon the muscles of the trunk and limbs (see Fig. 10.1). As noted in Chapter 6, an α lower motor neuron innervates several extrafusal muscle fibers of a single muscle, forming a *motor unit* (see Fig. 6.8). The neuronal cell bodies of the motor units of a given muscle are clustered into a *motor neuron pool* located in the ventral horn of spinal cord gray matter. The motor neuron pool of a muscle has a cigar-shaped, longitudinal organization in the cord, often extending rostrocaudally over a few spinal cord segments (e.g., L1-L3; Fig. 10.2). These motor neuron pools have a *somatotopic organization* (or in this case, *musculotopic organization*) in the ventral horn; that is, their relative position in the CNS corresponds to the relative body position of the muscles that their neurons innervate. In other words, motor neuron pools whose neurons innervate distal muscles of the limbs tend to be located in more lateral parts of the ventral horn, whereas motor neuron pools associated with axial and proximal musculature tend to be located more medially within the ventral horn.

Lower motor neurons projecting out to muscles are often synaptically activated by *premotor neurons* whose cell bodies are usually located in the intermediate zone of spinal cord gray matter (see Fig. 10.2), but some premotor neurons can be found in the ventral horn as well. Activating a premotor neuron in the *lateral* part of the intermediate zone on one side of the body will generally activate a modest number of α motor neurons in the *lateral* part of the ventral horn, on the same side of the body. This, in turn, will result in the activation of a modest number of *distal* limb muscles that would generally be used for skilled, voluntary movement. *Medial* premotor neuron activation on one side of the body will generally activate a larger number of α motor neurons, in the *medial* part of the ventral horn, often on both sides of the body and often over more than one spinal cord segment. This, in turn, will result in the extensive activation of *axial* or *proximal* antigravity muscles on both sides of the body. Such a complement of muscles would be required for the involuntary stabilization or adjustment of posture. It can therefore be seen that more lateral parts of the spinal cord gray matter are involved in control of the distal limb musculature of skilled voluntary movement, whereas more medial parts are associated with the axial and proximal musculature of postural control.

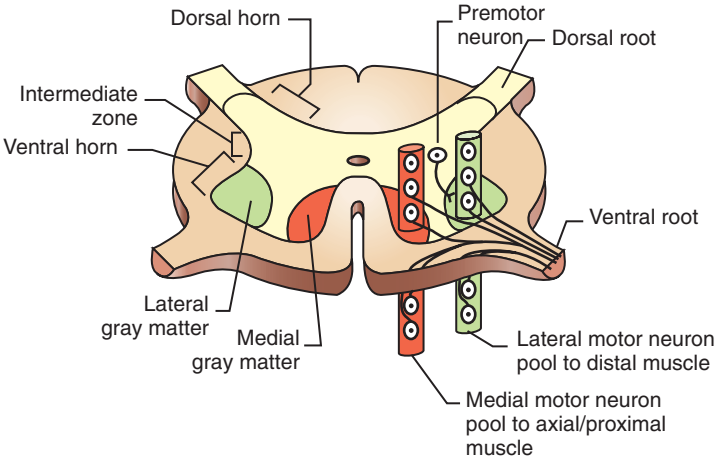
The simplest type of motor behavior, the spinal segmental reflex (e.g., the knee jerk reflex; see Chapter 7), can be organized at the level of the spinal cord, without significant control from more rostral divisions of the CNS (e.g., the brain). However, although control from the brain may not be *necessary* for such behaviors, these simple reflexes can often still be influenced by brain input. Furthermore, under different circumstances, the same spinal premotor and α motor neurons that participate in a simple spinal reflex could be activated by the brain to participate in an elegant and skilled sequence of movement.



• **Fig. 10.1** Motor system hierarchy.

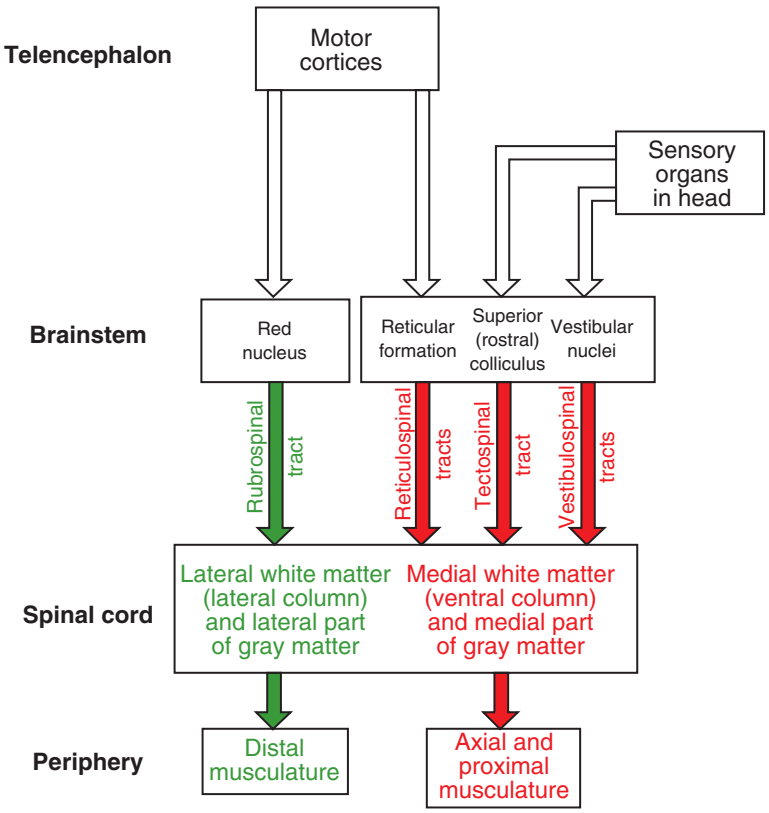
Brainstem Upper Motor Neuron Pathways Are the Source of All Descending Motor System Input to the Spinal Cord, Except for One Other Major Pathway

Four major axon routes originate in the brainstem and descend to the spinal cord to influence spinal lower motor neurons: the



• **Fig. 10.2** Somatotopic (musculotopic) organization of lower motor neurons in the ventral horn of the spinal cord that respectively supply the distal and axial/proximal musculature. Cell bodies of motor units supplying a given muscle are arranged in longitudinal columns within the ventral horn called *motor neuron pools*. Motor neuron pools to more distal muscles lie laterally to those supplying the axial and proximal musculature. Spinal premotor neurons, which synapse on the motor neurons supplying muscles, are commonly found in the intermediate zone of spinal cord gray matter and also have a somatotopic (musculotopic) organization. (Modified from Kandel ER, Schwartz JH, Jessell TM. *Principles of Neural Science,*. 3rd ed. New York: Elsevier Science Publishing; 1991.)

Vet-ebooks.com



• **Fig. 10.3** Organization of the descending brainstem motor pathways to the spinal cord. The medial brainstem motor pathways are the reticulospinal, vestibulospinal, and tectospinal tracts (*labeled red arrows*). They principally travel in more medial regions of the spinal cord white matter and synapse within more medial regions of the spinal cord gray matter controlling the axial and proximal musculature. The rubrospinal tract (*labeled green arrow*) is a lateral brainstem motor pathway that travels in more lateral regions of the spinal white matter and synapses within more lateral regions of the spinal gray matter controlling the distal limb musculature. Crossing of some of the pathways is not represented.

vestibulospinal tracts, the *reticulospinal tracts,* the *tectospinal tract,* and the *rubrospinal tract* (Fig. 10.3). Collectively, the first three are involved in the involuntary maintenance and adjustment of posture, and in reflex orientation of the head. Therefore they are principally involved in the control of axial and proximal musculature. The rubrospinal tract is mainly involved in control of distal limb musculature of the type that mediates voluntary skilled movements. These four routes (often along with components of the basal ganglia and cerebellum) are sometimes referred to as the *extrapyramidal motor system*. This is in contrast to the *pyramidal motor system* that originates in the cerebral cortex, the other major descending motor route to the spinal cord, as discussed later. Because the term *extrapyramidal* can encompass such a diverse group of structures, and because it is often applied inconsistently, it is being used less frequently. The four routes from brainstem to spinal cord are collectively referred to here as the *descending brainstem motor pathways.*

The brainstem, like the spinal cord, contains lower motor neurons that can synaptically activate skeletal muscles, in this case the face and head muscles (see Fig. 10.1). The cell bodies of these α motor neurons reside in various cranial nerve nuclei (e.g., facial motor, hypoglossal, oculomotor). The brainstem also receives direct input from sensory organs in the face and head (e.g., eye, vestibular apparatus). Therefore as in the spinal cord, some fairly simple segmental reflexes can be organized at the brainstem level without the necessity of significant control from other levels of the motor system. Because the brainstem contains the previously noted descending motor pathways to the spinal cord, the brainstem provides a means by which input from sensory organs in the face and head can reach and control lower motor neurons of the spinal cord that operate muscles of the trunk and limbs (see Fig. 10.3). Some of the descending brainstem motor pathways also provide a means by which more rostral regions of the motor system (e.g., motor cortex) can indirectly influence spinal lower motor neurons.

Medial and Lateral Descending Brainstem Motor Pathways Respectively Control Proximal Muscles of Posture and More Distal Muscles of Skilled Movement

The descending brainstem motor pathways to the spinal cord can be divided into a medial group and a lateral pathway. The vestibulospinal, reticulospinal, and tectospinal tracts constitute the *medial brainstem motor pathways,* whereas the rubrospinal tract represents *the lateral brainstem motor pathway* (see Fig. 10.3). The groupings are generally based on the relative position of these tracts within the spinal cord white matter. The axons of the tracts that represent the medial brainstem motor pathways (vestibulospinal, reticulospinal, tectospinal) will mainly travel in more medial regions of the spinal cord white matter (e.g., ventral column) and will synapse within more medial regions of the spinal cord gray matter. Those medial regions of the spinal gray matter contain medial premotor neurons and medial α motor neurons that control the axial and proximal extensor musculature primarily involved in involuntary maintenance and adjustment of posture. Axons of the lateral brainstem motor pathway (rubrospinal) run in a more lateral region of the spinal white matter (lateral column) and synapse in the more lateral spinal gray matter. The premotor and α motor neurons of this region principally control the distal flexor musculature involved in voluntary skilled movement.

Vet-ebooks.com

Thus medial brainstem motor pathways project to medial regions of the spinal cord gray matter whose neurons control the more medially located (axial and proximal) extensor muscles of posture, whereas the lateral brainstem motor pathway projects to lateral regions of the spinal gray matter whose neurons control the more laterally located (distal) flexor muscles of skilled movement.

The Reticulospinal and Vestibulospinal Tracts Are Medial Brainstem Motor Pathways Important for Keeping the Body Upright Against the Pull of Gravity

A major responsibility of the medial descending brainstem motor pathways is to maintain the body subconsciously in an upright position against the pull of gravity. The reticulospinal and vestibulospinal tracts play a major role in this involuntary control of the axial and proximal extensor musculature that prevents the animal from falling to the ground. The reticulospinal tracts are particularly important in controlling the magnitude of the steady-state contraction level, or *muscle tone,* of these antigravity muscles. Components of the vestibulospinal tracts play an essential role in activating the antigravity muscles in response to destabilization of the body with respect to gravity. Keep in mind that subconscious control of the postural musculature is an integral part of the ability to execute skilled voluntary movement of the distal musculature successfully, because voluntary movement requires a stable platform on which it can proceed.

The reticulospinal tracts originate from cell bodies in the *reticular formation* of the brainstem (see Fig. 10.3). This is a complex of less aggregated and smaller nuclei, and somewhat more loosely organized axonal projections, located near the midline. Once thought to be a diffuse and fairly nonspecific system, the reticular formation is now known to contain a number of functionally specific nuclei. In addition to being the origin of medial descending brainstem motor pathways to the spinal cord, ascending projections of the reticular formation play an important role in modulating consciousness, arousal, and attention. The reticular formation receives a vast array of sensory information and plays an important role in pain perception, respiration, and circulatory function.

Axons of the reticulospinal tracts synapse within medial regions of the spinal cord gray matter that primarily control the axial and proximal extensor musculature (see Fig. 10.3). Collectively, the tracts project to virtually all rostrocaudal levels of the cord. The *pontine reticulospinal tract* originates from cells in the reticular nuclei of the *pons* and, after traveling in the ventral column of the spinal cord, tends to have an *excitatory* effect on lower motor neurons to the antigravity muscles. The *medullary reticulospinal tract,* coming from the reticular nuclei of the *medulla,* tends to have an *inhibitory* effect on the lower motor neurons to antigravity muscles. In some species, the medullary reticulospinal tract axons, although running close to the ventral column, actually travel in the ventral part of the lateral column close to the ventral horn. However, the axons still synapse within the more medial regions of spinal cord gray matter. These opposing portions of the reticulospinal tracts interact to regulate antigravity muscle tone. Influences from other regions of the brainstem, the cerebellum, and the spinal cord endow the pontine reticular nuclei, and their pontine reticulospinal tract axons, with a high level of spontaneous activity. The effects of such spontaneous excitatory activity on antigravity muscle tone can be tempered by activation of the inhibitory medullary reticular nuclei and their medullary reticulospinal tract axons. Descending projections from the cerebral cortex to the brainstem represent a significant forebrain source of relative control over the two reticulospinal tracts. This cortico-reticulospinal route emphasizes the point that some of the descending brainstem motor pathways provide an indirect way for more rostral levels of the motor system hierarchy to influence spinal lower motor neurons (see Fig. 10.3).

The descending cortical projections to the origins of the reticulospinal tracts endow those tracts with two important motor functions, in addition to their critical role in the subconscious modulation of antigravity muscle tone. The first function is related to skilled voluntary movement requiring a stable postural background, as previously noted. Just *before* the execution of such a voluntary movement, reticulospinal tract components *subconsciously* activate the appropriate axial and proximal musculature that will compensate for the postural destabilization that will be produced by the intended voluntary movement (usually of the distal musculature). Reticulospinal tract components under the control of descending cortical projections also play a role in the voluntary execution of crude (nonskilled), often stereotypical movements of the proximal limb musculature, such as those involved in simple pointing or locomotion.

As noted in Chapter 8, γ motor neurons are usually activated along with α motor neurons so that muscle spindles maintain their sensitivity to stretch even when the muscle is shortened during contraction. This α-γ *co-activation* is a principle common to the excitation of lower motor neurons by upper motor neurons. Under certain circumstances, however, it appears that this process can be dissociated, such that the γ motor neuron–mediated sensitivity of the muscle spindle, and thus the sensitivity of the stretch reflex, can be adjusted apart from extrafusal muscle contraction. Although reticulospinal tract components participate in α-γ coactivation of lower motor neurons, they appear to be strongly associated with the ability to regulate γ motor neuron activity independently. It is likely that this ability of reticulospinal tract components to independently modulate the sensitivity of the stretch reflex underlies their significant role in adjusting antigravity muscle tone.

The vestibulospinal tracts originate from cell bodies in the *vestibular nuclear complex,* a bilateral group of four distinct subnuclei occupying a substantial portion of the medulla, and part of the pons, next to the lateral wall of the fourth ventricle (see Fig. 11.8). The subnuclei of this complex receive their principal synaptic input from the eighth cranial nerve fibers carrying sensory input from the vestibular apparatus of the inner ear (see Chapter 11). The vestibular apparatus provides sensory information about the position of the head with respect to gravity and about acceleration of the head through space, thus indicating body position and disturbances of balance. The vestibular nuclear complex also receives significant input from the cerebellum but not from forebrain levels of the motor system hierarchy.

As in the reticulospinal tracts, axons of the vestibulospinal tracts synapse within medial regions of the spinal cord gray matter that primarily control the axial and proximal extensor musculature (see Fig. 10.3). The *lateral vestibulospinal tract* (see Chapter 11) projects to all rostrocaudal levels of the spinal cord, and when the vestibular apparatus detects a disturbance of balance, the tract excites the antigravity extensor musculature in an attempt to counteract the disturbance. The less substantial *medial vestibulospinal tract* only

Vet-ebooks.com

reaches as far as the cervical and upper thoracic levels of the spinal cord. It principally controls the neck musculature to produce compensatory head movements in response to rotational disturbances of balance (see Chapter 11). Although the vestibulospinal tracts principally function to produce compensatory adjustments to postural disturbances, components seem to make some contribution to antigravity muscle tone as well.

Some aspects of the functions of the reticulospinal and vestibulospinal descending brainstem motor pathways can be better understood by considering the clinical state called *decerebrate rigidity.* This condition occasionally results from severe forebrain disease. It also results from surgical transection of the brain at the rostral midbrain level, as discovered by the British neurophysiologist Charles Sherrington. As noted earlier, the neurons of the pontine reticulospinal tract, which excite lower motor neurons to antigravity muscles, have a high degree of spontaneous activity. Excitation of the neurons of the medullary reticulospinal tract inhibits the lower motor neurons to antigravity muscles. When the forebrain is disconnected from the brainstem, descending projections from the cerebral cortex cannot excite these medullary reticular neurons projecting to the spinal cord, and thus a significant source of inhibition to the lower motor neurons of antigravity muscles is removed. The excitation of the lower motor neurons produced by the spontaneous activity of the pontine reticulospinal neurons has now lost a significant source of opposition, and therefore much greater muscle tone exists in the antigravity muscles. The animal now assumes a hobbyhorse-like posture, often so rigid that the animal stands in a fixed position. Subsequently cutting a portion of the lateral vestibulospinal tract reduces some of this rigidity, so the tract apparently plays some normal role in regulating antigravity muscle tone, in addition to its principal role in responding to postural destabilization with respect to gravity.

As noted earlier, the reticulospinal and vestibulospinal tracts make important contributions to the control of the axial and proximal musculature to keep the body upright. However, the to-and-fro rhythmicity of walking and running is organized by circuits of spinal interneurons that control the lower motor neurons in a repetitive, oscillating manner. Although these spinal neural networks are capable of producing this simple oscillatory behavior without control by more rostral parts of the motor system hierarchy, reticulospinal neurons play an important role in initiating this locomotor behavior and in controlling its speed.

The *tectospinal tract* is a medial brainstem motor pathway principally involved in reflex orientation of the head toward environmental stimuli. The cells of origin of the tectospinal tract are located in the superior colliculus of the midbrain (often called *rostral colliculus* in quadrupeds; see Fig. 10.3). Like the reticulospinal and vestibulospinal brainstem motor pathways to the spinal cord, the tectospinal tract axons synapse within medial regions of the spinal cord gray matter that primarily control the axial and proximal musculature. However, these axons only project as far as the upper cervical regions of the cord. This is consistent because the tectospinal tract principally controls the musculature that moves the head. The superior colliculus processes visual, auditory, and somatosensory information about the relative position of stimuli in the environment with respect to the organism. The superior colliculus can also control rapid reflex movements (saccades) of the eyes to the stimulus. The tectospinal tract is involved in producing a movement of the head toward the stimulus that corresponds with the rapid eye movement so that the animal's gaze is fixated directly on the stimulus.

The Rubrospinal Tract Is a Lateral Brainstem Motor Pathway That Can Control Distal Limb Musculature Associated With Skilled Movement

As noted earlier, the reticulospinal, vestibulospinal, and tectospinal tracts are medial descending brainstem motor pathways whose axons run rostrocaudally mainly within more medial portions of the spinal white matter and synapse in more medial portions of the spinal gray matter. This region of the spinal gray matter exerts extensive, often bilateral control of the axial and proximal musculature involved in postural control and head orientation. In contrast, the *rubrospinal tract* is a lateral descending brainstem motor pathway whose axons course within more lateral regions of the spinal white matter and synapse in more lateral portions of the spinal gray matter (see Fig. 10.3). This region of the spinal gray matter exerts unilateral control over a limited complement of muscles of the distal limbs, often flexors, associated with skilled movements of the extremities.

The rubrospinal tract axons originate in cells of the *red nucleus* (nucleus ruber) of the mesencephalon. The red nucleus receives very significant descending input from higher levels of the motor system hierarchy in the cerebral cortex. This cortico-rubrospinal route provides a means for the motor cortices to indirectly influence the spinal lower motor neurons that operate the distal limb flexor musculature. Therefore the cortico-rubrospinal route is involved in the voluntary control of musculature that participates in skilled, often manipulative movements of the extremities (although not in the most dexterous movements of the digits). The rubrospinal tract is more important for these types of movements in quadrupeds compared with primates. In primates, direct projections from motor cortices to the spinal cord (the corticospinal tract, described next) are more important than the rubrospinal tract in the control of voluntary skilled movement of the extremities. As for most nuclei giving rise to tracts that play a direct role in movement, the red nucleus also receives significant input from the cerebellum. The role of the cerebellum in motor control is briefly described later and in more detail in Chapter 12.

The Corticospinal (Pyramidal) Tract Is a Direct Projection From Cerebral Cortex to Spinal Cord Responsible for the Most Skilled Voluntary Movements of Mammals

The motor cortices of the forebrain constitute the portion of the motor system hierarchy above that of the brainstem and represent the most complex level. As mentioned earlier, these cortical regions are collectively capable of operating spinal lower motor neurons indirectly through some of the descending brainstem motor pathways to the spinal cord (e.g., cortico-reticulospinal route, cortico-rubrospinal route). In mammals, a more efficient system exists for the cortical control of spinal lower motor neurons: a direct projection from cells in the motor cortices to the gray matter of the spinal cord. This direct *corticospinal tract,* also referred to as the *pyramidal tract,* is responsible for the most elaborate and dexterous voluntary movement sequences of which mammals are capable, especially movements involving the extremities. However,

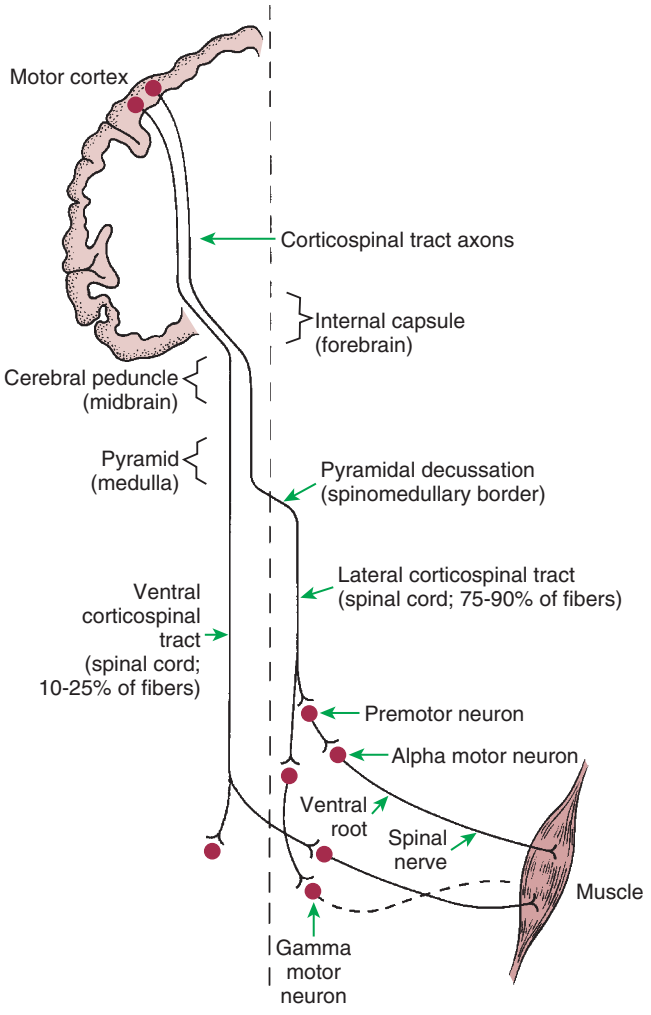
Vet-ebooks.com

this tract also participates in less elaborate voluntary movements of the distal musculature and can exert some voluntary control over the postural muscles as well.

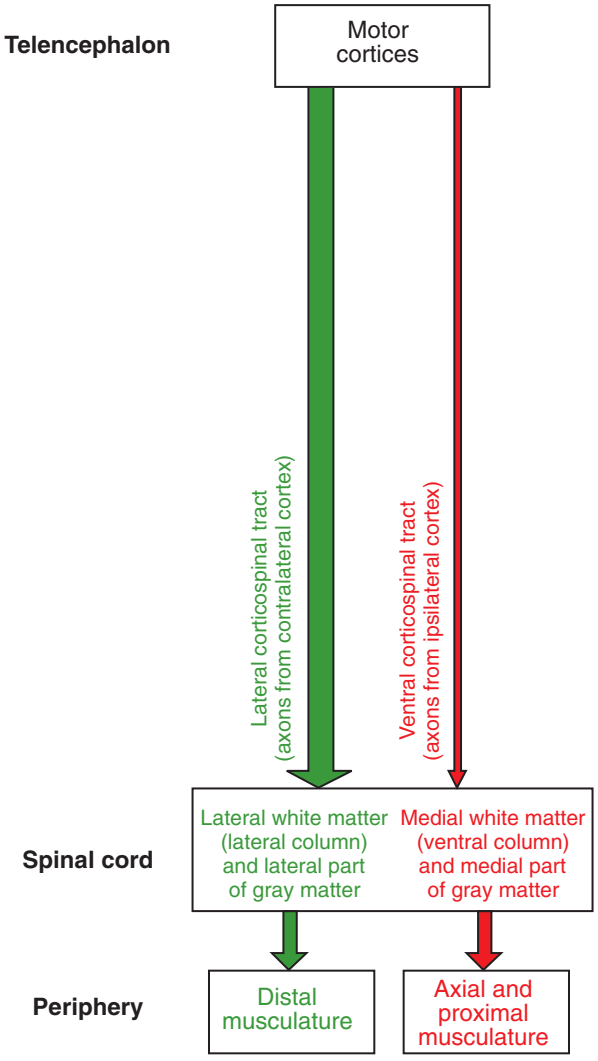
In the Spinal Cord, the Corticospinal Tract Has a Massive Lateral Component Controlling the Distal Musculature and a Minor Medial Component Controlling the Axial and Proximal Musculature

The corticospinal tract axons primarily originate from cells located in the motor cortices in, or adjacent to (depending upon species), the frontal lobe of the cerebral hemisphere (Fig. 10.4). All cells contributing to the tract are located in layer five of the six histological layers of cortical tissue (see Fig. 16.1). Along their route from the cerebral cortex, these corticospinal axons pass through the internal capsule of the forebrain, the cerebral peduncles on the ventral surface of the midbrain, and the pontine nuclei within the ventral pons. They emerge on the ventral surface of the medulla, adjacent to the midline, as the pyramids. These appear pyramid-shaped in cross-section, partly inspiring the name *pyramidal tract* for axons that pass through them.

As the corticospinal tract axons reach the spinomedullary border, the vast majority (75% in canines to 90% in primates) cross the midline at a structure called the *pyramidal decussation* (see Fig. 10.4). The crossing axons then form the *lateral corticospinal tract,* located in the lateral spinal cord white matter, and synapse within lateral regions of the spinal cord gray matter (Fig. 10.5). As noted



• **Fig. 10.4** The corticospinal tract is a direct route primarily from the motor cortices to the contralateral spinal cord gray matter. Most axons of the tract synapse on premotor neurons of the intermediate zone, but some, depending on species phylogeny, synapse directly on α and γ lower motor neurons. About 75% to 90% (again depending upon phylogeny) of the axons of the tract cross the midline at the spinomedullary border to form the lateral corticospinal tract, and about 10% to 25% remain on the same side to form the ventral corticospinal tract. Some prominent anatomical structures formed by corticospinal tract axons within the brain, and their locations, are also indicated.



• **Fig. 10.5** Somatotopic (musculotopic) relationship of corticospinal tract components. Like the descending brainstem motor pathways, the corticospinal tract can be divided into components that respectively travel in more lateral or more medial regions of the spinal cord white matter. The massive and laterally located lateral corticospinal tract synapses in more lateral regions of the spinal cord gray matter that control the distal limb musculature. The axons of this tract originate from the contralateral motor cortices. The much smaller ventral corticospinal tract, whose axons originate from the ipsilateral motor cortices, travels in more medial regions of the spinal white matter and synapses in more medial regions of the spinal gray matter that control the axial and proximal musculature.

Vet-ebooks.com

earlier, the lateral regions of the spinal gray matter contain premotor and α motor neurons that primarily control the distal flexor musculature of the extremities that participate in skilled, manipulative, usually voluntary movements. Given this organization, damage to the motor cortices on one side of the body has devastating effects on voluntary skilled movement of the distal flexor musculature on the opposite side of the body. A much smaller percentage of axons traveling in the medullary pyramid do not cross the midline at the pyramidal decussation and remain on the same side of the body to form the much smaller *ventral corticospinal tract* (see Fig. 10.4). The axons of this tract are located in more medial regions of the spinal white matter and synapse in more medial regions of the spinal gray matter that control the axial and proximal postural musculature (see Fig. 10.5). Many axons of the ventral corticospinal tract actually cross the midline, locally, just before synapsing in the spinal gray matter. The ventral corticospinal tract provides a direct means of voluntary control over muscles normally involved in a subconscious antigravity function.

The ability of the corticospinal tract to control the most dexterous, skilled movements of the body derives from the synaptic termination pattern of several of its axons. The greater the number of synapses between a neuron in the motor cortices and an α motor neuron in the spinal cord ventral horn, the greater the number of α motor neurons activated and the less precise the control of the musculature. This is true because each neuron that is excited in the pathway usually activates several postsynaptic neurons. Corticospinal axons bypass synapsing with neurons of the brainstem motor pathways to the cord, but more significantly, some corticospinal axons can bypass synapsing with premotor neurons of the spinal gray matter, contacting α motor neurons directly. Therefore a given corticospinal neuron can ultimately control smaller numbers of α motor neurons and a smaller complement of the musculature. This permits increased *fractionation* of movement, the increased independence of the actions of different muscles (e.g., the ability to move individual fingers instead of all the fingers together). The proportion of corticospinal neurons making monosynaptic connections with spinal α motor neurons is related to phylogeny. There are no such connections in cats, a small proportion in monkeys, a larger proportion in the anthropoid apes, and a still larger proportion in humans, where the most skilled, fractionated, manipulative movements take place.

As noted earlier, cranial nerve nuclei in the brainstem contain lower motor neurons that travel through cranial nerves to control muscles of the face and head (see Fig. 10.1). A significant complement of the axons that leave the motor cortices to form the corticospinal tract will not continue to the spinal cord but will leave the tract to synapse at the cranial nerve motor nuclei of the brainstem. This complement of axons is referred to as the *corticobulbar tract* ("bulb" being an archaic term for portions of the brainstem; also called the *corticonuclear tract*). A given cranial nerve nucleus generally receives significant corticobulbar input from both cerebral hemispheres.

The Motor Cortices in or Adjacent to the Frontal Lobe, the Highest Level of the Motor Control Hierarchy, Consist of Three Different Functional Regions

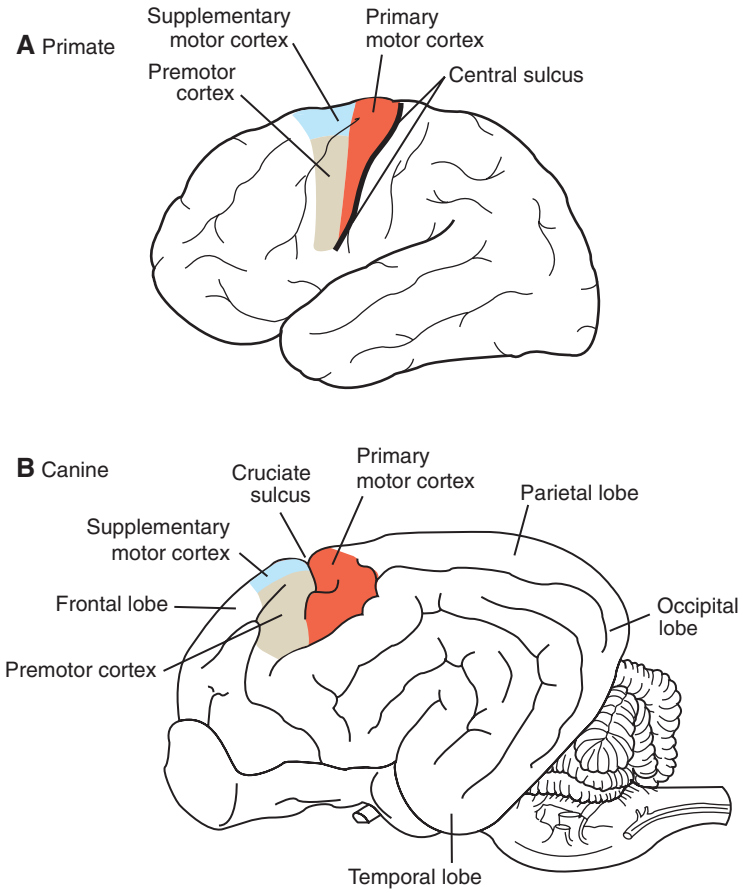
The motor cortices in or adjacent to the frontal lobe, the origin of most of the corticospinal tract axons, are composed of the primary motor cortex, the supplementary motor cortex, and the premotor cortex (Fig. 10.6; see also Fig. 10.1). Although these forebrain regions collectively represent the highest level of the motor control hierarchy, the areas appear to differ with respect to the complexity of motor functions controlled.

In primates, the *primary motor cortex* (MI) is located just rostral to the prominent central sulcus and therefore lies along the precentral gyrus (see Fig. 10.6A). In many nonprimate mammalian species, a central sulcus is not present, and MI appears to lie near the cruciate sulcus (see Fig. 10.6B). Low-level electrical stimulation of a very small region of MI is capable of activating a small number of functionally related muscles. Furthermore, an orderly relationship exists between the region of the body where the muscles are activated and the region of MI stimulated. In this *somatotopic map* (or *musculotopic map* in this case) of the body musculature in MI, muscles in the caudal part of the body (or the feet in bipeds) can be most easily activated from more dorsomedial parts of MI, whereas muscles in the rostral part of the body (or head in bipeds) can be most easily activated from more ventrolateral parts of MI, with a fairly orderly representation of the other regions of the body between those parts of MI. As shown in Fig. 10.7, the musculature of different parts of the body is not equally represented in the somatotopic map. Regions depicted as larger have a larger area of MI devoted to their voluntary muscular control, and thus the movements of that region will generally be that much more precise and fractionated. In the somatotopic map of MI in humans, the hand and the mouth musculatures have a very large proportional representation, reflecting the respective importance of these areas in manipulating objects with the fingers and in articulating speech. The proportional representation of the musculature of the different body parts in MI varies with phylogeny, but the somatotopic maps of the primates tend to be the most detailed, reflecting the most precise control over skilled, voluntary movements.

The *supplementary motor cortex* and the *premotor cortex* are located in the frontal lobe, just rostral to MI, with supplementary motor cortex positioned dorsomedial to premotor cortex (see Fig. 10.6). Both areas also have a somatotopic map of the body musculature, but it is less precise than in MI. In addition to corticospinal and corticobulbar tract axons, both areas also collectively give rise to axonal projections to nuclei of origin of some descending brainstem motor pathways (see Fig. 10.1). Most significantly, however, the supplementary motor and premotor cortices send axons to synapse within MI and thus may represent "supramotor" areas, with an even higher status in the motor control hierarchy than MI; these areas may instruct MI to organize its fairly discrete muscle actions into more elaborate movement patterns. This concept is supported by the fact that, for voluntary movement, neurons in these supramotor areas become active before those of MI.

Evidence indicates that the supplementary motor cortex is particularly important in the *planning and organizing of complex sequences* of the discrete movements normally carried out by MI. For example, the supplementary motor cortex appears to be particularly active when an individual mentally rehearses a specific sequence of finger movements. The supplementary motor cortex also appears to be important for instructing the limbs (particularly the forelimbs) on the two sides of the body to work together, simultaneously, to accomplish a task. The premotor cortex appears to play an important role in *preparatory orientation of the body* for the execution of a particular motor task. An example in primates would be rotation of the shoulders and movement of the arms toward a target to be manipulated by the hands. Both these areas receive integrated sensory input and visuospatial information (from the posterior

Vet-ebooks.com



• **Fig. 10.6** Motor cortices. Vicinity of primary motor, supplementary motor, and premotor cortices in the human (A) and canine (B) brain. Portions of both primary and supplementary motor cortices extend onto the medial surface (not shown).

parietal cortex), which most likely plays a role in their respective functions.

Using the analogy of playing the piano, we could view MI as being responsible for the simplest muscle activation necessary to press a single piano key, the supplementary motor cortex as responsible for planning and organizing the sequence of such finger movements necessary to play a melody, and the premotor cortex as responsible for orienting the arms and hands to the correct region of the keyboard to play the various sequences. Of course, the interaction among these areas in determining the appropriate corticospinal (and corticobulbar) tract activity necessary to produce the voluntary movement is certainly more complex than this, and the functional role of these areas in motor control, and how they work together, is still under investigation.

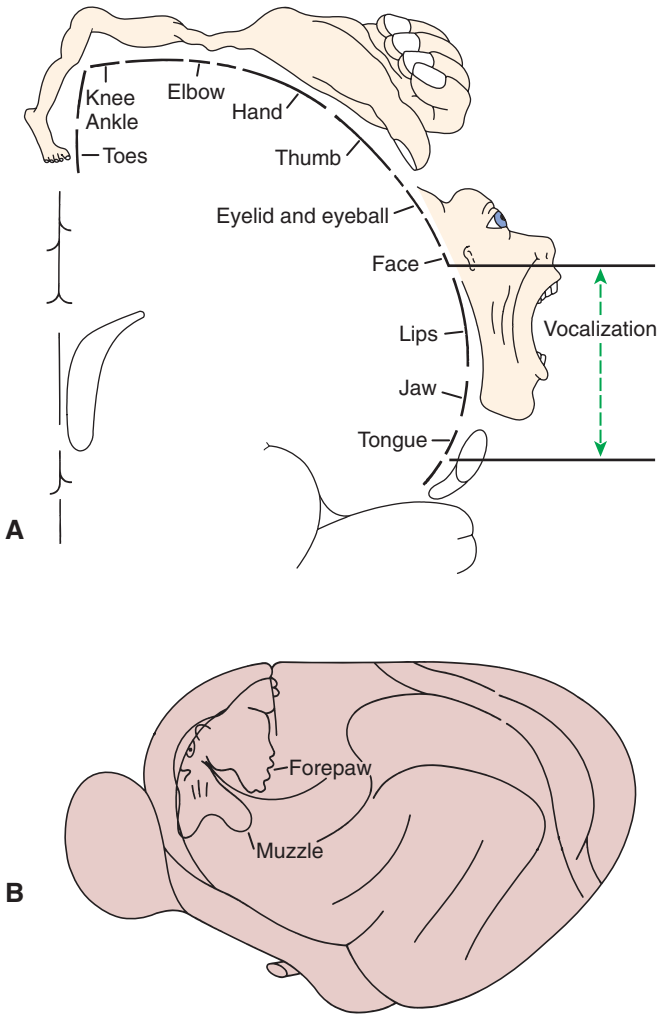
The severity of the deficits resulting from lesions of the corticospinal (pyramidal) system varies with the phylogenetic status of the animal. In primates, such as humans, in whom the pyramidal system is developed extensively, corticospinal tract lesions rostral to the pyramidal decussation cause a dense weakness of the contralateral side of the body. Such one-sided weakness is called *hemiparesis* and is most extensive in the hand and facial muscles (symptoms common in "stroke" in humans). These symptoms are understandable because a huge percentage of the corticospinal tract axons in primates cross the midline at the spinomedullary border (the location of the pyramidal decussation), and the hand and face have the largest proportional representation in primary motor cortex (see Figs. 10.4 and 10.7).

In most veterinary species, the corticospinal system is not as well developed as in humans, and supraspinal lesions of this system cause much less severe contralateral weakness and little alteration of gait. However, clinical examination can reveal more subtle deficits in voluntary movements of the contralateral limbs. An example is the *proprioceptive positioning reaction,* the ability of an animal to return its paw to a normal, pads-down posture after the clinician flexes it to make the dorsal surface touch the floor or tabletop. This response requires the animal's conscious awareness that the paw is in the flexed position (conscious proprioception) and then requires that the animal be able to respond consciously by returning the paw to its normal posture. This latter motor response is affected by the integrity of the upper motor neurons of the corticospinal tract. When these corticospinal tract neurons are damaged, the animal is slow to return its paw to a normal posture. In addition, toes tend to be dragged on the ground as the leg is drawn forward in normal gait. It should be noted that such deficits could also be produced by damaged corticorubral axons (axons to the red nucleus) originating in the motor cortices. Noting these conscious positioning deficits and other subtle gait changes is important in localizing lesions within the CNS.

Corticospinal Tract Coactivation of Both Alpha (α) and Gamma (γ) Lower Motor Neurons May Help With Small Automatic Corrections of Voluntary Movements

As noted earlier, α-γ coactivation is a principle common to the excitation of lower motor neurons by upper motor neurons. It has been suggested that such coactivation may permit the muscle spindle to function as an "automatic error correction system" when voluntary

Vet-ebooks.com



• **Fig. 10.7** Somatotopic (musculotopic) map of primary motor cortex (MI) showing the origins of axons controlling different skeletal muscles of the body. Body parts represented as proportionally larger have a larger area of MI devoted to their voluntary control, and the movement of that part will generally be that much more precise and fractionated. (A) In the human, muscles controlling the hand/digits and mouth are disproportionately represented because these muscles are needed for the critical and precise movements of grasping/manipulation and speech. (B) Analogous somatotopic (musculotopic) map for a cat. (A redrawn from Penfield W, Rasmussen T. *The Cerebral Cortex of Man*. New York: Macmillan; 1950; from Berne RM, Levy MN. *Physiology*. 2nd ed. St Louis: Mosby; 1988; B from Prosser CL. *Comparative Animal Physiology*. 3rd ed. New York: Wiley; 1988.)

movement against a load results in a small deviation from the intended result.

As discussed in Chapter 8, the activation of γ motor neurons along with α motor neurons ensures that the intrafusal muscle fibers remain taut enough to transduce stretch even as the muscle reaches a shorter length on contraction of the extrafusal fibers. The γ motor neuron activation tightens the intrafusal fibers by causing contraction of their polar ends, resulting in adjustment of the muscle spindle sensitivity to the new length of the muscle. It is thought that the α-γ coactivation resulting from a voluntary motor command is meant to produce a contraction of intrafusal fibers concordant with extrafusal fiber contraction, such that the muscle spindle is made just sensitive enough to transduce stretch at the new muscle length. Under these circumstances, if the load is more than expected, the α motor neuron activity will not have produced enough extrafusal fiber contraction to shorten the muscle to the new desired length. However, the γ motor neuron activity will have produced the appropriate intrafusal fiber contraction to adjust the muscle spindle sensitivity for the new desired length. This mismatch, where the spindle sensitivity has been adjusted for the new muscle length but the extrafusal fibers have not contracted enough to reach that length, results in a stretching of the muscle spindle and activation of segmental stretch reflex mechanisms. That is, the stretching of the muscle spindle results in more excitatory postsynaptic potentials (EPSPs) on the α motor neurons to the muscle, increasing their action potential firing and increasing the extrafusal fiber contraction to assist with reaching the new desired length.

This type of error correction, in which segmental stretch reflex mechanisms help to accomplish the intended muscle shortening when the corticospinal pathway has not produced the sufficient α motor neuron activity, is called a *servo-assist function.* Thought to result from α-γ coactivation, this servo-assist function is analogous to the power steering in a car, where a compressor in the motor adds power to the driver's turning of the steering wheel when significant resistance is encountered by the tires.

The Motor System Shares Some Organizational Principles With Sensory Systems

With most of the major components and pathways of the motor system now described, it appears that the motor system shares principles of organization common to other brain systems (e.g., sensory systems). One such organizational principle is the existence of topographic maps of the body. As noted earlier, there are organized somatotopic maps of the body's musculature in the motor cortices. Topographic organization also exists in many sensory systems, except it is the peripheral receptor surface that is topographically mapped. For example, CNS components of the somatosensory (touch) system, such as primary somatosensory cortex, contain an organized somatotopic map of the different regions of the skin surface.

Two other principles of organization shared by the motor system and sensory systems are serial and parallel processing of nervous system information. In sensory systems, *serial processing* generally refers to the passage of information from the periphery to successively more rostral regions of the nervous system, in a serial fashion. For example, in the visual system, axons of cells in the retina synapse in the lateral geniculate nucleus of the thalamus, and these thalamic neurons in turn send their axons to synapse in primary visual cortex. Often, in serial processing within the sensory systems, the information collected at successively more rostral levels of the nervous system is organized into a more sophisticated form. Serial processing can also be observed in the motor system, although in a different direction: from more rostral regions to more caudal regions. The cortico-reticulospinal route is an example of this. However, in motor system serial processing, commands often proceed from areas with more complex organization to those with simpler organization (see Fig. 10.1).

Parallel processing refers to the different pathways within a given sensory system operating in parallel, respectively, to carry qualitatively different forms of information. Again, using the somatosensory

Vet-ebooks.com

system as an example, there are separate pathways to the cerebral cortex to carry information about a gentle touch of the skin and about an intense skin contact usually perceived as painful. In the motor system, an example of parallel processing is the respective control of the proximal antigravity musculature by one set of descending brainstem motor pathways (vestibulospinal, reticulospinal) and control of the distal flexor musculature by a different descending brainstem motor pathway (rubrospinal).

Undoubtedly, a combination of both serial and parallel processing is necessary for the integrated function of sensory as well as motor systems.

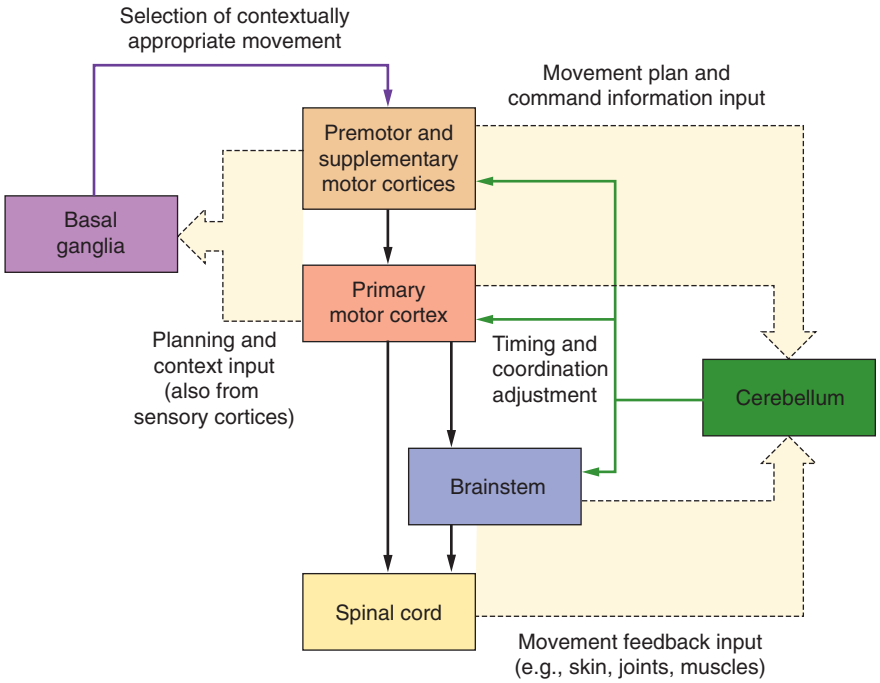
The Basal Ganglia and Cerebellum Modulate the Activity of Motor System Components for the Respective Selection and Adjustment of Movement

Portions of the motor system are important for proper motor function but do not appear to be directly involved in initiating movement. These structures—the basal ganglia and the cerebellum—serve primarily to modulate the activity of other motor system structures without directly producing movement (Fig. 10.8).

The *basal ganglia* are a group of nuclei, the majority of which are deep within the cerebral hemispheres. They include the caudate nucleus and putamen (known collectively as the striatum), the globus pallidus, the substantia nigra, and the subthalamic nucleus. The internal neural circuitry of this multinuclear functional unit is extremely complex and participates in several parallel pathways running through the basal ganglia. The basal ganglia receive input from the motor cortices and many other areas of the cerebral cortex and, by way of the thalamus, project output back to the motor cortices, particularly the supplementary motor and premotor cortices (see Fig. 10.8). Again, these regions are important in the planning and preparation for movement. Some basal ganglia output projects directly to brainstem nuclei controlling movement.

Generally it is thought that the basal ganglia use the information received from the cortex, including information about the movement plan and the context of the situation, to *help select the appropriate movement pattern while suppressing less appropriate, competing patterns.* Two principal circuits within the basal ganglia play an important role in this process. One circuit acts to facilitate inhibitory output of the basal ganglia, presumably acting to suppress the inappropriate, competing movement pattern. The other circuit acts to reduce inhibitory output of the basal ganglia, presumably "removing the brakes" from the appropriate movement pattern. Dopamine-containing neurons that project from the substantia nigra of the basal ganglia to the striatum of the basal ganglia play an important role in regulating these two circuits. When these dopamine-containing neurons degenerate in humans with Parkinson's disease, severe motor deficits develop, such as difficulty beginning an appropriate movement, slowness of movement, rigidity, and resting tremor. Parkinson's disease does not occur naturally in veterinary species, but some toxins can selectively destroy these dopamine-projecting neurons in nonhuman species, producing similar motor deficits as seen in the human disease. In horses, ingestion of the yellow star thistle can produce damage of the basal ganglia. Some of the abnormal movements resulting from this damage, involving the lips and tongue of the horse, are reminiscent of such abnormal movements seen in the fingers of human patients with Parkinson's disease. Interestingly, the respective structures in the two species are both used in grasping movements.

The structure and function of the *cerebellum* and its role in motor control are discussed in Chapter 12 and are mentioned only briefly here. The cerebellum's importance in motor control is indicated by the earlier observation that virtually all the nuclei giving rise to the brainstem motor pathways receive output from the cerebellum. Also, the cerebellum indirectly receives input (through pontine nuclei) from the motor cortices (MI, supplementary motor cortex, premotor cortex). As with the basal ganglia, the cerebellum not only receives information from the motor cortices but indirectly sends information back to them as well. Importantly, the cerebellum receives much sensory information from the skin,



• **Fig. 10.8** Modulatory roles of the cerebellum and basal ganglia in relation to the motor system hierarchy. Interposed synaptic relays are not represented.

Vet-ebooks.com

joints, muscles, vestibular apparatus, and even the visual system. Therefore the cerebellum receives information about the planning and initiation of movement, as well as continuous sensory feedback about the progress of the movement (see Fig. 10.8). The cerebellum, in turn, can influence activity in the motor cortices and in the brainstem motor pathways to the spinal cord.

Through this organization, it is thought that the cerebellum acts to *compare information about the movement plan with information about how the movement is actually being carried out.* It can then presumably *make adjustments to the ongoing movement itself, or even adjust the movement plan* for future execution. Within this framework, the cerebellum appears particularly concerned with gathering sensory feedback about, and with adjusting the control of, the timing of movement. Both experimental and clinical studies have shown that cerebellar damage produces significant deficits in the coordination and smoothness of complex movements. These deficits presumably arise because of problems in the timing of muscle contraction components of the movement. If the muscle contraction components are not properly timed, the movement can appear jerky and uncoordinated, may exhibit improper force, and may not stop at the appropriate time.

CLINICAL CORRELATIONS

Focal Lesion of the Motor Cortex

History

You examine an 11-year-old female boxer dog. Her vaccination history is current. She had an adenocarcinoma of the mammary gland removed 6 months before your examination.

The owner states that over the past few days the dog has become progressively weaker in the left front and left rear legs, and occasionally stands with the left front paw flexed such that the dorsal surface touches the ground. On the previous day, the dog had a seizure.

Clinical Examination

On physical examination of the patient, you find several routine old-age changes and the results of the mammary surgery. You find also that the dog seems drowsy and is weak on the left front and left rear legs. She has a proprioceptive positioning reaction deficit of both the left front and the left rear leg. Radiographic study of the chest reveals metastatic, neoplastic lesions in the lungs.

Comment

The proprioceptive positioning reaction (or response) is tested by flexing the animal's paw, dorsal side down, while gently supporting her weight. A normal dog senses (conscious proprioception) that the paw is upside down and returns it to the normal pads-down posture (motor response). This is called a response (or reaction), rather than a reflex, because it involves a degree of conscious control. This particular response requires normal function of skin and joint receptors and the peripheral nerve in the tested leg and of the sensory neuron tracts that ascend toward the brain along the ipsilateral (same) side of the spinal cord. Traveling along a multisynaptic pathway, the sensory information crosses to the contralateral (opposite) side of the brain in the brainstem and reaches the contralateral (with respect to the side of the original stimulus) cerebral cortex. As the animal becomes consciously aware that the paw is in an unusual position, action potentials are sent back, down the corticospinal tract, to the lower motor neurons of the muscles of the leg, causing the paw to return to the normal position.

With the wiring diagram of this response in mind, you can see that a deficit in the proprioceptive positioning reaction of the left front and left rear legs could be caused by a lesion of the left cervical spinal cord, the right motor cortex, or supraspinal portions of the right corticospinal tract (before crossing at the pyramidal decussation). This dog's seizure (a manifestation of brain disease) at about the same time suggests that the lesion is in the right cerebral cortex. The brain is a common site for metastasis, and the radiographic lung lesions suggest that the mammary tumor has spread to both the lung and the right side of the brain. The lung contains the first capillary bed that a metastatic cancer cell is likely to encounter when it enters the venous system of the mammary gland. Some cells stop here and grow.

Treatment

Dogs with metastatic mammary carcinomas are usually not treated except to make them more comfortable.

Cow With Hypomagnesemia

History

It is May in Southwest Virginia, and a client calls you with a 2-year-old Angus (beef) cow that just calved 2 weeks ago. Early in the morning, they noticed that the cow had been acting a bit abnormal, and then she started to become agitated and hyperexcitable. Within the last 20 to 30 minutes, the cow has gone down and appears unable to stand. The calf appears to be normal but appears to be having difficulty nursing due to the cow's agitated state and current recumbency.

Clinical Examination

You rush there immediately, and the cow remains recumbent. She appears unable to stand, and you notice that she has nystagmus, as well as muscle spasms. The temperature, heart rate, and respiratory rate are all increased. You are suspicious that she may progress to convulsions fairly quickly if not treated appropriately.

Comment

You suspect hypomagnesemia based on clinical signs, the recent calving, and that there is lush grass, which is often high in potassium and nitrogen but low in sodium and magnesium. The lush pasture also increases the pH of the rumen and decreases the transit time in the intestines, both of which cause decreased absorption of magnesium. Magnesium is a cofactor for a vast number of enzymatic reactions and is required for virtually all enzymatic processes involving adenosine triphosphate (ATP). It is therefore critical for a huge array of metabolic pathways, including controlling muscle contractions and normal nerve conduction. Thus deficiency in magnesium limits basic metabolism and activity of the body, and depolarization of nerve and muscle cells is altered. Because both are divalent cations, calcium and magnesium often compete directly in physiological processes. Thus with hypomagnesemia, the resulting increased calcium binding alters nerve and muscle activity, including neurotransmitter release and cardiac conduction. Within the CNS, the decreased magnesium results in decreased impedance to neuronal calcium influx, and this is further exacerbated by receptor-mediated actions of the neurotransmitter glutamate. The excessive influx of calcium leads to abnormally prolonged activation of intraneuronal calcium-mediated signaling, thus causing neurotoxic damage from excessive enzyme activation that can lead to cell death. As a result, signaling from the motor cortex through the ventral and lateral corticospinal tracts, as well as the premotor neurons, α motor neurons, and the muscle, are all affected. This can account for the prominent motor signs in this cow.

Treatment

Immediate treatment with intravenous magnesium is critical. Additional magnesium supplementation is given orally to decrease chances of relapses. Treated animals should not be disturbed for at least 30 minutes. Relapses are common.

Vet-ebooks.com

Practice Questions

1. A motor neuron pool located most laterally in the ventral horn of the spinal cord is most likely to operate a muscle controlling movement of the:
- **a.** Proximal limb
- **b.** Neck
- **c.** Distal limb
- **d.** Abdomen

2. Which of the following is *true* regarding decerebrate rigidity?
- **a.** It can result from severe forebrain disease.
- **b.** Disruption of the cortical control of medullary reticulospinal neurons is a major contributor to the condition.
- **c.** It can result in a fixed, rigid, hobbyhorse-like posture in the quadruped.
- **d.** Removal of normal inhibition to some of the antigravity muscles contributes to the condition.
- **e.** All of the above are true.

3. Which of the following descending brainstem motor pathways controls distal limb musculature associated with skilled movement?
- **a.** Medial vestibulospinal tract
- **b.** Rubrospinal tract
- **c.** Pontine reticulospinal tract
- **d.** Tectospinal tract
- **e.** All of the above play a major role in such control

4. The corticospinal (pyramidal) tract, in general, initiates what form of movement?
- **a.** Antigravity movement
- **b.** Postural adjustment
- **c.** Skilled, voluntary, mostly flexor movement
- **d.** Tremulous, jerky movement
- **e.** None of the above

5. You are presented with a dog with a weakness, and proprioceptive placing reaction deficit, of his left front and left back legs. A single pathological site could cause these signs if it were located in the:
- **a.** Left side of the cervical spinal cord
- **b.** Left cerebral cortex
- **c.** Right cerebral cortex
- **d.** Either a or b
- **e.** Either a or c

6. The corticospinal tract simultaneously co-activates both the α and the γ lower motor neurons. If the initial coactivation fails to be sufficient to cause the intended shortening of the muscle, sensory neuron activity of the muscle spindle of that muscle will have what influence on the α motor neurons to the same muscle?
- **a.** Addition of EPSPs
- **b.** Addition of IPSPs
- **c.** No influence
- **d.** Decrease in action potential frequency
- **e.** Either b or d

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Chang HT, Rumbeiha WK, Patterson JS, et al. Toxic equine parkinsonism: an immunohistochemical study of 10 horses with nigropallidal encephalomalacia. *Vet Pathol*. 2012;49(2):398–402.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Fletcher TF. Spinal cord and meninges. In: Evans HE, de Lahunta A, eds. *Miller's Anatomy of The Dog*. 4th ed. Philadelphia: Saunders; 2013.

Fletcher TF, Beitz AJ. The brain. In: Evans HE, de Lahunta A, eds. *Miller's Anatomy Of the Dog*. 4th ed. Philadelphia: Saunders; 2013.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Jennings DP. Supraspinal control of posture and movement. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Lemon RN. Descending pathways in motor control. *Annu Rev Neurosci*. 2008;31:195–218.

Lorenz MD, Coates JR, Kent M. *Handbook of Veterinary Neurology*. 5th ed. Philadelphia: Saunders; 2011.

McFarlane D. Endocrine and metabolic diseases. In: Smith BP, eds. *Large Animal Internal Medicine*. 4th ed. St Louis: Mosby Elsevier; 2009.

Nicholls JG, Martin AR, Fuchs PA, et al. *From Neuron to Brain*. 5th ed. Sunderland, Mass: Sinauer; 2012.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Stewart AJ. Magnesium disorders. In: Reed SM, Bayly WM, Sellon DC, eds. *Equine Internal Medicine*. 2nd ed. St Louis: Saunders; 2004.

Vet-ebooks.com

11

The Vestibular System

BRADLEY G. KLEIN

KEY POINTS

1. The vestibular system is a bilateral receptor system located in the inner ear.
2. Specialized regions of the vestibular system contain receptors.
3. The semicircular ducts detect rotary acceleration and deceleration of the head.
4. The utricle and saccule detect linear acceleration and deceleration and static tilt of the head.
5. The vestibular system provides sensory information for reflexes involving spinal motor neurons, the cerebellum, and extraocular muscles of the eye.
6. Vestibular reflexes coordinate eye and head movements to maximize visual acuity.

To coordinate posture and locomotion, the brain needs to know not only what movement it intends to command but also the orientation of the body and what movement the body is actually performing. Chapter 8 describes the muscle spindle, an important source of information for the brain about body position and movement. Another important source of information is the vestibular system. This bilateral receptor system is located in the inner ear and informs the brain about the position and motion of the head.

The vestibular system provides the organism with its sense of equilibrium or balance. It supplies information about the body's orientation or tilt with respect to gravity and about acceleration of the body through space. It does this by detecting (1) static tilt of the head (e.g., the head is held stationary at 5 degrees from vertical), (2) linear acceleration of the head (e.g., the head accelerates in a straight line as an organism begins to run or as your elevator begins to rise), and (3) rotary acceleration of the head (e.g., the head accelerates in a circular fashion as an organism begins to turn its head toward a target or as someone begins to spin you in an office chair). This information is used most often to make automatic, unconscious postural adjustments to keep the organism from falling over as a result of self-imposed or environmentally imposed changes in body orientation. The vestibular system also helps to keep the eyes fixated on a relevant target in the face of such changes in body orientation.

The vestibular system is a common site of pathological lesions. In most veterinary species, lesions of the vestibular system cause a syndrome characterized by head tilt, compulsive rotary movements such as circling or rolling, and spontaneous nystagmus, which is an oscillating movement of the eyes.

To understand how such clinical signs arise and the importance of the vestibular system to the physiology of movement, you need to study its anatomy and function first.

The Vestibular System Is a Bilateral Receptor System Located in the Inner Ear

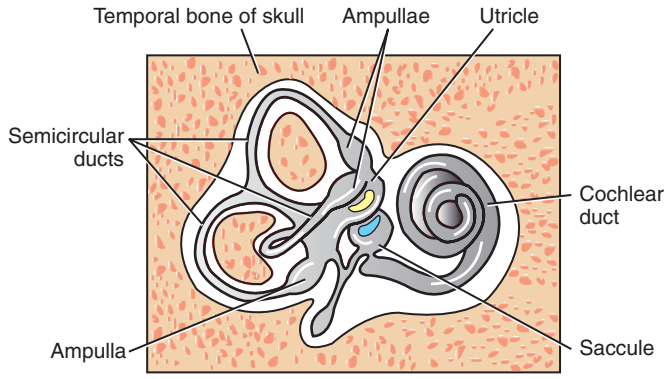
The inner ear, or *labyrinth,* is made up of two parts: the bony labyrinth and the membranous labyrinth. The bony labyrinth is a system of caverns and tunnels through the petrous temporal bone of the skull. The *bony labyrinth* houses the receptor organs of the vestibular system as well as the receptor organ for hearing, the *cochlea* (Fig. 11.1) (see Chapter 17). These receptor organs are part of the *membranous labyrinth,* which consists of thin membranes of epithelium and lies within the bony labyrinth. This epithelial membrane is specialized at some locations to become the sensory receptor cells that form the vestibular and auditory receptor organs. The membranous labyrinth is filled with a fluid called *endolymph* and is separated from the bony labyrinth by a fluid called *perilymph.* The vestibular portion of the membranous labyrinth consists of two major sets of structures: (1) three *semicircular ducts,* located at approximately right angles to each other, and (2) a pair of saclike structures called the *utricle* and *saccule,* sometimes called the *otolith organs.* As discussed later, each major set of structures is respectively involved in transducing a different major class of vestibular stimulus.

Specialized Regions of the Vestibular System Contain Receptors

Each vestibular structure of the membranous labyrinth has a region of epithelial lining that has become specialized into a set of secondary receptor cells (see Chapter 7) called *hair cells* (Fig. 11.2). These hair cells form the basis of a sensory receptor organ within each vestibular structure. Each hair cell has several cilia at its apex arranged in order according to size. At its base, the hair cell synapses with a sensory neuron that carries action potentials to the brainstem. The cell bodies of these sensory neurons are located in Scarpa's ganglia, and their axons collectively form the vestibular portion of the vestibulocochlear nerve (cranial nerve VIII). The cilia from all the hair cells within any one vestibular structure project into a gelatinous mass; displacement of this gelatinous mass in a given direction causes all the hair cell cilia to bend in that direction.

At rest, when the cilia are not deflected, the sensory neurons that synapse with the vestibular hair cells transmit action potentials spontaneously at about 100 per second (Fig. 11.3). When the hair

Vet-ebooks.com



• **Fig. 11.1** The bilateral inner ear contains receptor systems for hearing (cochlea) and for detecting the orientation and acceleration/deceleration of the head (vestibular system). The peripheral vestibular apparatus on each side of the head contains a utricle, a saccule, and three semicircular ducts, each with an ampulla at one end.

cell cilia are bent in a direction toward the largest cilium, the hair cells depolarize, the release of transmitter from the hair cells onto the sensory neurons increases, and the action potential frequency of the neurons increases. When the cilia are bent in the opposite direction toward the smaller cilia, hair cell membranes hyperpolarize, transmitter release decreases, and the action potential frequency of sensory neurons decreases. Therefore displacement of the hair cell cilia in either of these directions can be detected by the brain as an increase or a decrease from the resting action potential frequency. Deflections in other directions are much less effective. The depolarization or hyperpolarization of the hair cell membrane, depending on the direction of cilia deflection, is regulated by mechanically sensitive, gated ion channels at the tips of the cilia (Fig. 11.2A). Elastic links between the cilia and these ion channels result in the channels increasing their opening (compared with the resting state) when the cilia are deflected in one direction and decreasing their opening when deflected in the opposite direction. How the brain uses this information from the hair cells to detect the direction of head movement is described later.

The Semicircular Ducts Detect Rotary Acceleration and Deceleration of the Head

Three membranous semicircular ducts are located within corresponding semicircular canals of each bony labyrinth (Fig. 11.4). They are positioned at approximately right angles to each other, and both ends of each fluid-filled duct terminate in the utricle. Each semicircular duct has an enlargement at one end, called the *ampulla,* near its junction with the utricle. The ampulla contains a hair cell receptor organ called the *crista ampullaris* (see Fig. 11.2). This is a ridge of hair cells that, at their base, synapse on sensory neurons projecting to the central nervous system (CNS) and whose cilia are embedded in an overlying gelatinous mass. This mass, called the *cupula,* attaches to the roof of the ampulla (Fig. 11.5). All hair cells of a given crista ampullaris are oriented in the same direction with respect to their cilia. Together, the hair cell ridge and overlying cupula span the diameter of the ampulla.

The semicircular ducts, together with the ampullae and its contents, are involved in transducing rotary acceleration and deceleration of the head. When the head begins to accelerate in a rotary fashion, the semicircular duct and its receptor organ rotate with the head, but the endolymph's acceleration lags behind because of inertia. This relative difference in the rate of acceleration of the semicircular duct and its enclosed endolymph causes the crista ampullaris to "crash into" the slower-moving endolymph. This results in a displacement of the gelatinous cupula in the direction opposite that of head rotation with a corresponding bending of the hair cells. This in turn changes the firing rate of the sensory neurons projecting to the CNS. The opposite happens with deceleration because the semicircular duct and crista ampullaris slow immediately along with the head while inertia continues to carry the endolymph forward. Stimulation of the crista ampullaris occurs on rotary acceleration or deceleration of the head, not during constant rotational velocity. During the latter, the movement of the endolymph will eventually catch up with the movement of the semicircular ducts and the hair cells will no longer be bent.

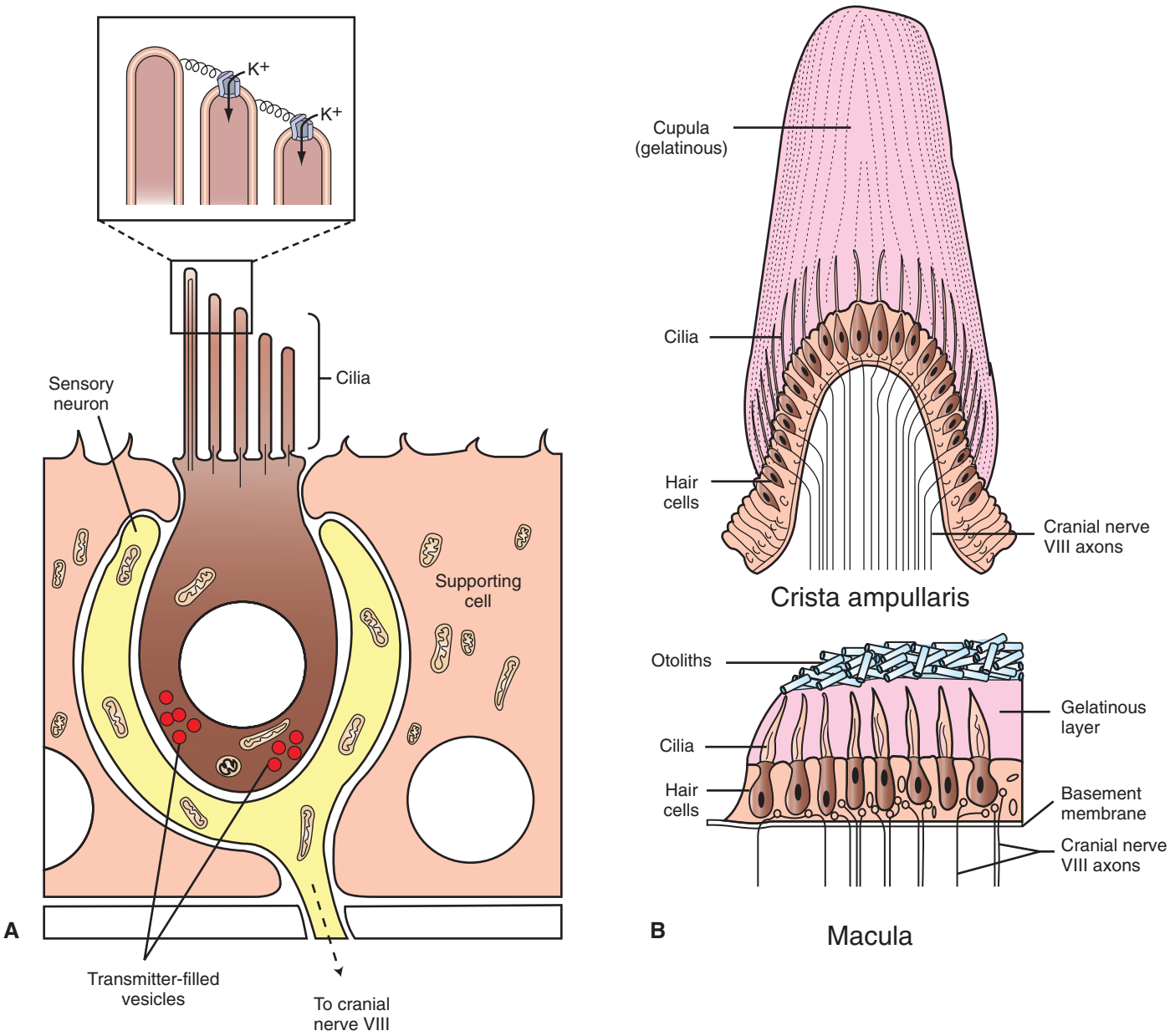
Semicircular ducts located on opposite sides of the head, but in approximately the same plane (co-planar), work as a pair to provide the brain with information about the direction and nature of head movement. For instance, a clockwise rotary acceleration of the head would cause bending of the directionally sensitive hair cell cilia in each member of a co-planar pair of semicircular ducts on opposite sides of the head. However, the sensory axons leaving the crista ampullaris from the duct on one side of the head would carry an increased action potential frequency, whereas those from the duct of the other side would carry a decreased action potential frequency. The brain interprets such reciprocal changes in sensory action potential frequency as resulting from clockwise or counterclockwise acceleration or deceleration in a given plane of movement. In reality, rotary acceleration/deceleration in any given plane usually affects all three sets of paired semicircular ducts but each pair to different degrees. In this way, the bilateral system of six semicircular ducts detects the direction of both rotary acceleration and deceleration of the head and activates or inhibits particular CNS structures to produce the appropriate reflex response.

The Utricle and Saccule Detect Linear Acceleration and Deceleration and Static Tilt of the Head

In the utricle and saccule, the hair cell receptor organ is called the *macula* (see Fig. 11.2). It is an oval patch of hair cells with a primarily horizontal orientation on the floor of the utricle and a primarily vertical orientation on the wall of the saccule. The cilia of the macular hair cells extend into a gelatinous layer atop the hair cells. Embedded at the top of this gelatinous mass is a layer of calcium carbonate crystals called *otoliths.* This otolith layer is heavier and denser than the endolymph and other surrounding materials.

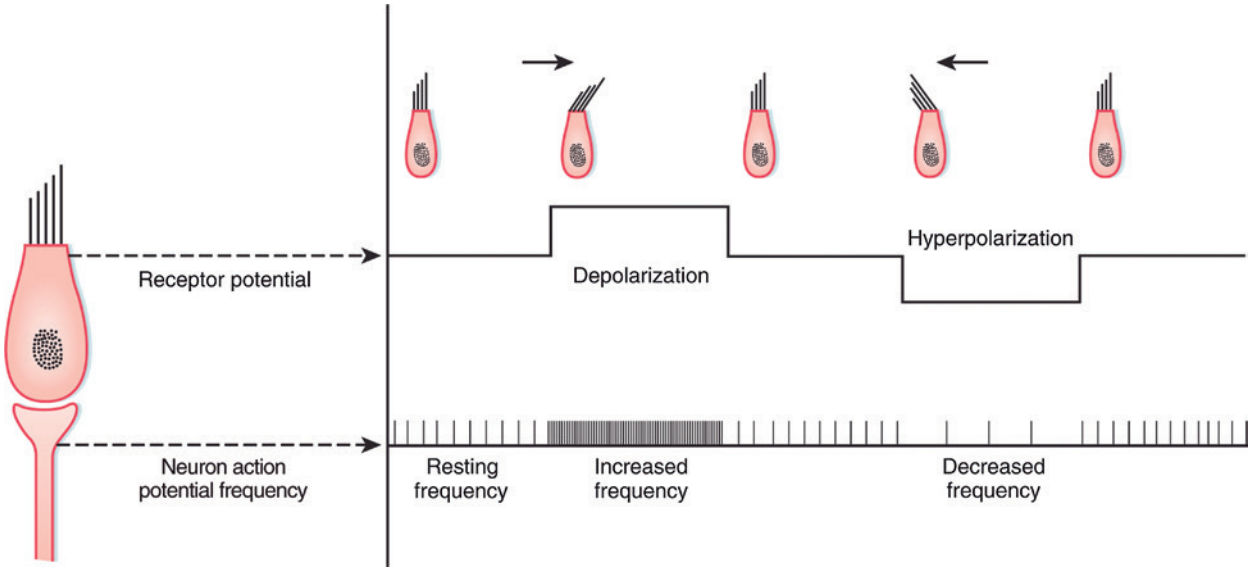
Given the organization of the receptor apparatus that lies within the utricle and saccule, these vestibular structures can transduce linear acceleration and deceleration of the head, as well as static head tilt (Fig. 11.6). Considering the horizontally oriented macula of the utricle, if the head is accelerated forward in a straight line, movement of the dense otolith layer lags behind that of the macular hair cells. This "drag" produces a shearing force that bends the tips of the hair cell cilia, by way of the gelatinous layer, until constant velocity is achieved and the otolith layer catches up with the hair cell layer. In contrast to the crista ampullaris, not all the hair cell clusters of a given macula are oriented in the same direction with respect to their cilia (Fig. 11.7). In addition, as previously noted, the macula of the utricle is oriented in the horizontal plane, whereas the macula of the saccule is in the vertical plane. Therefore linear acceleration in a specific direction will bend hair cells of a

Vet-ebooks.com

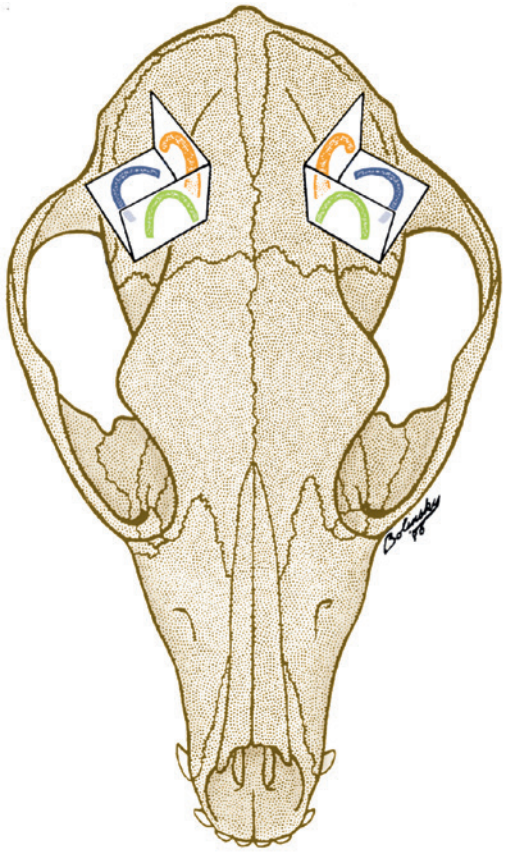


• **Fig. 11.2** Each structure of the peripheral vestibular apparatus contains a region of hair cells that form the basis of a sensory receptor organ. (A) Each hair cell has several cilia at its apex, arranged in size order, and synapses on a sensory neuron of cranial nerve VIII at its base. *Inset,* Mechanically sensitive, elastically linked, gated ion channels at the tips of the cilia regulate depolarization or hyperpolarization of the hair cell membrane depending on the direction in which the cilia are deflected. (B) In each ampulla of the semicircular ducts, there is a crest of hair cells whose cilia project into a gelatinous mass called the *cupula,* forming a receptor organ called the *crista ampullaris.* The hair cell receptor organ in the utricle and saccule is the macula, a layer of hair cells whose cilia project up into a gelatinous layer, on top of which lies a layer of calcium carbonate crystals called otoliths. (Inset modified from Bear MF, Connors BW, Paradiso, MA. *Neuroscience: Exploring the Brain.* 4th ed. Philadelphia: Wolters Kluwer; 2016.)

Vet-ebooks.com



• **Fig. 11.3** At rest, sensory neurons on which vestibular hair cells synapse transmit action potentials spontaneously at a rate of about 100 per second. When hair cell cilia are deflected in one direction, the action potential frequency increases; when cilia are deflected in the opposite direction, the frequency decreases.



• **Fig. 11.4** Three semicircular ducts, each positioned at approximately right angles to the other two, are located on each side of the head and work to detect rotary acceleration and deceleration of the head.

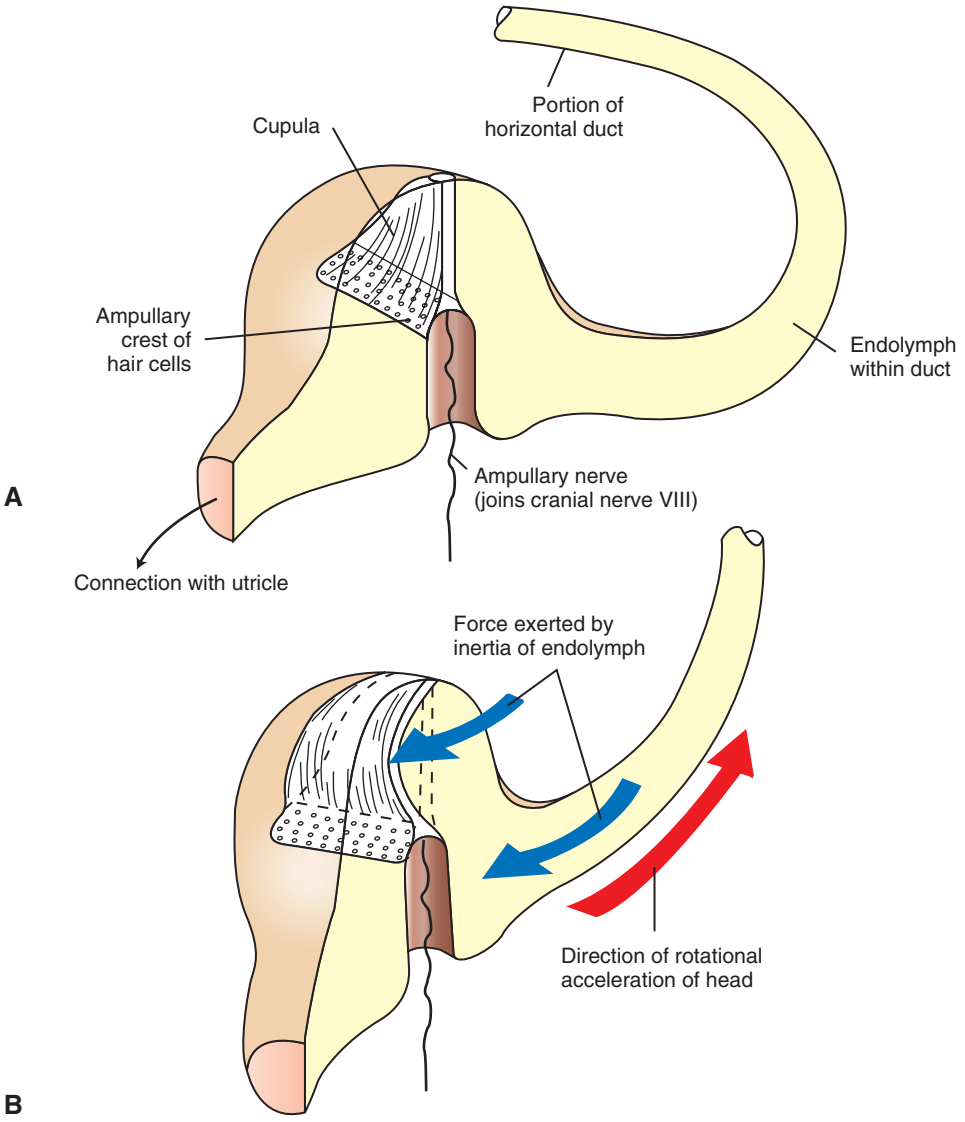
particular location and orientation in a way that will transiently increase the action potential firing rate of their associated sensory neurons; those of another location and orientation will be bent in a way that transiently decreases the firing rate; and those of yet another location and orientation will be bent in a way that has little or no effect on firing rate. This topographical pattern of hair cell bending and associated transient changes in action potential firing will be different for linear acceleration in a different direction. The CNS can decipher these various patterns of neural activity to determine the onset and direction of linear acceleration and to initiate an appropriate compensatory response.

Again considering the horizontally oriented macula of the utricle, when the head tilts from the upright position, the heavy and dense otolith layer effectively "falls over" as it is pulled by gravity (see Fig. 11.6B). This bends the hair cell cilia, by way of the gelatinous layer, and keeps them bent as long as the head is tilted. This sustained bending (compared with the transient bending during linear acceleration) is translated into sustained changes in action potential firing frequency (compared with transient changes during linear acceleration) in particular populations of associated sensory neurons. In this way, the utricle and saccule can inform the brain about a stationary tilt of the head. The direction of the tilt is detected by a similar mechanism as that for detecting the direction of linear acceleration. Astronauts in low gravitational settings receive relatively little information from their utricles and saccules about their stationary head position and must rely more heavily on visual and other sensory cues to detect head position.

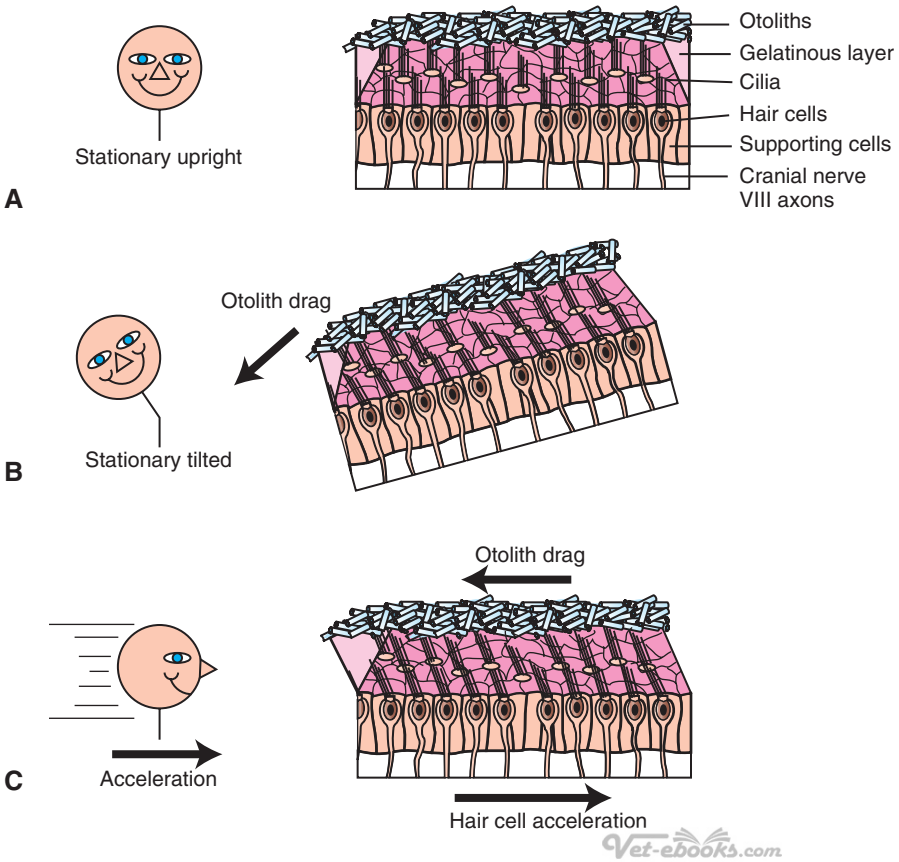
The Vestibular System Provides Sensory Information for Reflexes Involving Spinal Motor Neurons, the Cerebellum, and Extraocular Muscles of the Eye

As noted earlier, vestibular hair cells synapse on sensory neurons whose axons form part of the eighth cranial (vestibulocochlear)

Vet-ebooks.com

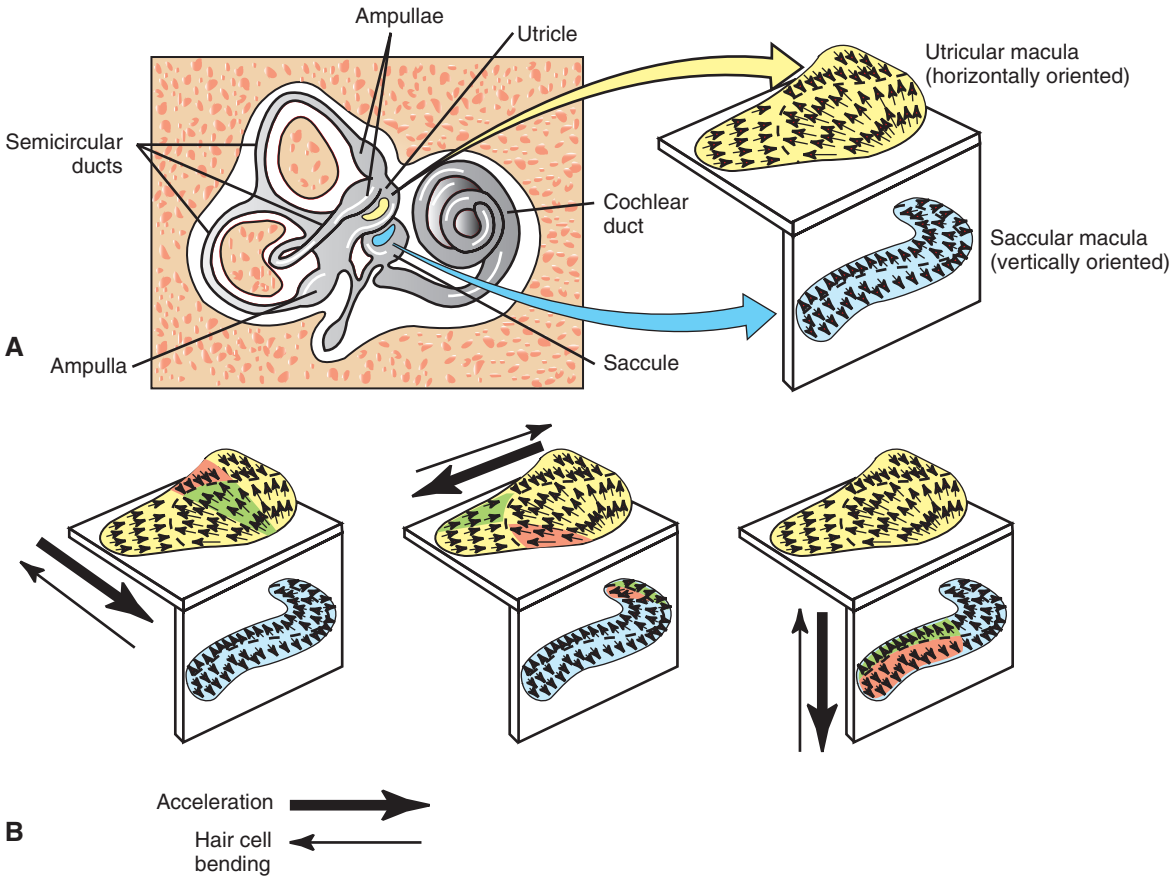


• **Fig. 11.5** Ampullae of the semicircular ducts contain a crista ampullaris, which transduces rotational acceleration/deceleration of the head. The ducts are filled with endolymph. (A) Crista ampullaris of the horizontal canal when the head is at rest. (B) On rotational acceleration of the head in the indicated direction, the relative inertia of the endolymphatic fluid displaces the cupula, and thus the hair cell cilia, in the opposite direction. (Modified from Kandel ER, Schwartz JH, eds. *Principles of Neural Science.* 2nd ed. New York: Elsevier Science Publishing; 1985.)



• **Fig. 11.6** Utricle and saccule each contain a macula, which transduces static head tilt and linear acceleration/deceleration of the head. Macula of the utricle is horizontally oriented. (A) When the head is stationary and upright, there is little or no bending of the hair cell cilia. (B) When the head tilts and remains tilted, the heavy otolith layer "falls over," producing a drag. This bends the hair cell cilia, by way of the interposed gelatinous layer, in the direction of the tilt. (C) When the head accelerates in a straight line, the hair cells accelerate in the same direction, but the heavy otolith layer lags behind, producing a drag in the opposite direction. This bends the hair cell cilia, by way of the interposed gelatinous layer, in the direction opposite the acceleration. (Portions modified from Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience.* 3rd ed. Sunderland, Mass: Sinauer; 2004.)

Vet-ebooks.com



• **Fig. 11.7** (A) Macula of the utricle is horizontally oriented, and macula of the saccule is vertically oriented. *Small arrows* in a macula represent the approximate orientation of the hair cells in that region, with respect to their cilia. For a given hair cell, the *arrow tip* represents the position of the largest cilium, and the *arrow tail* represents the shortest cilium. (B) Acceleration in a given direction (*large thick arrows*) results in bending of hair cell cilia in the opposite direction (*large thin arrows*) caused by otolith drag. Hair cells whose cilia are bent directly toward the largest cilium (*green regions*) will be depolarized the most and will produce the greatest increase in action potential frequency in their associated sensory neurons. Conversely, hair cells whose cilia are bent directly away from the largest cilium (*red regions*) will be hyperpolarized the most and will produce the greatest decrease in action potential frequency in their associated sensory neurons. Hair cells whose cilia are bent along other axes will be less significantly affected. (Portions modified from Fuchs AF. Peripheral motor control: the vestibular system. In: Patton HD, Fuchs AF, Hille B, et al., eds. *Textbook of Physiology,* 21st ed. Philadelphia: Saunders; 1989.)

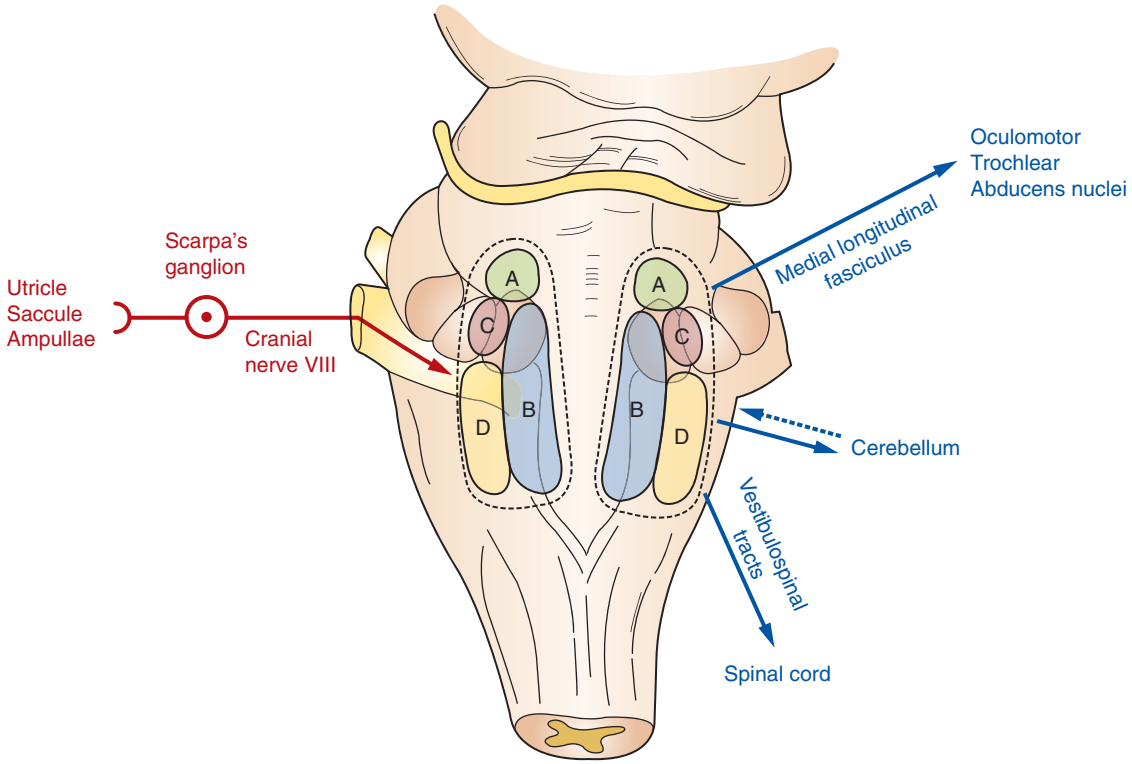
nerve and that carry action potentials to the medulla. Almost all these axons synapse in the *vestibular nuclear complex,* a bilateral group of four distinct nuclei occupying a substantial portion of the medulla, and part of the pons, next to the lateral wall of the fourth ventricle (Fig. 11.8). From here, second-order neurons (those on which cranial nerve VIII axons synapse) project to three important areas of the nervous system as described below.

Some of the neurons of the vestibular nuclear complex receive significant input from the utricle and saccule (the otolith organs), and their axons in turn form the lateral vestibulospinal tract. This tract provides excitatory facilitation to gamma (γ) and alpha (α) motor neurons of antigravity muscles of the trunk and limbs in response to linear acceleration/deceleration or static tilt of the head (see Chapter 10). Other neurons of the vestibular nuclear complex receive significant sensory input from the crista ampullaris of the semicircular ducts, and their axons in turn form a pathway that projects to cranial nerve nuclei that control eye movements. This pathway, called the *medial longitudinal fasciculus* (MLF), produces compensatory eye movements in response to rotary acceleration/deceleration of the head. The vestibular nuclear complex also sends projections to, and receives projections from, the cerebellum, especially the flocculonodular lobe. Through these reciprocal connections, the cerebellum can fine-tune the coordination of postural and oculomotor reflexes controlled by the vestibular system. Finally, some of the projections leaving the vestibular nuclear complex participate in neural circuits leading to cerebral cortex, resulting in conscious vestibular sensations.

Vestibular Reflexes Coordinate Eye and Head Movements to Maximize Visual Acuity

Vestibular reflex control of the extraocular muscles of the eye, known as the *vestibuloocular reflex* (VOR), coordinates eye and

Vet-ebooks.com



• **Fig. 11.8** Dorsal view of the brainstem (with cerebellum removed) and rostral spinal cord showing the vestibular nuclear complex (within *dashed borders*) with its major afferents (*left*) and efferents (*right*). The vestibular nuclear complex is comprised of the rostral (*A*), medial (*B*), lateral (*C*), and caudal (*D*) vestibular nuclei. The complex spans portions of both the medulla and pons. Different subsets of these nuclei receive afferents from particular portions of the vestibular apparatus, and other subsets give rise to particular efferent pathways, although this characteristic has not been specified in the figure. The *dashed arrow* denotes that the projection to the cerebellum is reciprocal. Note that the afferents and efferents depicted actually exist on both sides of the midline. (Portions from De Lahunta A, Glass E. *Veterinary Anatomy and Clinical Neurology,*. 3rd ed. Philadelphia: Saunders; 2009.)

head movements so that, as the head turns (rotates), the eyes remain fixed on the original field of vision for as long as possible. Imagine that a dog is seated on a piano stool and you rotate her clockwise to the right. As you rotate her slowly to the right, her eyes rotate in her head slowly to the left so that the eyes remain fixed on the same field of vision as long as possible. As the eyes reach the limit of their leftward excursion, they swiftly move to the right, in the direction of the head movement, until they fix on a new field of vision. If the head continues to rotate, the cycle repeats until constant velocity is achieved. This allows the animal time to interpret a field of vision despite rotary acceleration of the head. When this pattern of eye movement occurs during the VOR, it is referred to as *normal nystagmus* or *physiological nystagmus:* the slow drift opposite head rotation followed by a fast flick back in the direction of head rotation. A transient *postrotatory nystagmus,* with the converse eye movement pattern to that just noted, can be seen if a spinning animal or person is suddenly stopped. Inertia of the endolymph causes it to continue rotating in the semicircular duct, pushing on the crista ampullaris, even though the head and duct have stopped moving.

These reflex eye movement patterns require normal sensory input from the semicircular ducts, an intact MLF in the brainstem, and normally functioning motor units of the extraocular muscles (as well as an intact cerebellum). The VOR occurs on rotation of the head in the horizontal plane or the vertical plane, or on clockwise or counterclockwise torsional rotation of the head. Voluntary control of the eyes is independent of vestibular reflexes and is controlled by the cerebral cortex.

Nystagmus may appear occasionally under pathological conditions of the vestibular system, even when the head is at rest. This condition is known as *spontaneous nystagmus.* A persisting head tilt, falling, and compulsive circling or rolling often accompany spontaneous nystagmus in animals with acute vestibular disease. These actions often tend to be oriented in a consistent pattern with respect to the side of a peripheral lesion of the vestibular system. This is presumably a result of abnormal, asymmetrical action potential inputs to the brainstem from the vestibular apparatus on the two sides of the head.

Another compensatory reflex that can be elicited by rotational acceleration is the *vestibulocollic reflex* (VCR). This reflex acts to stabilize the head by activating elements of the neck musculature. An extreme example of this reflex is if your dog or cat were in a rowboat with you and the boat suddenly began to roll to one side; the animal's head would move in the direction opposite the roll to try to recapture its original position. Interestingly, the VCR is more effective in species that make minimal eye movements, such as pigeons and owls. In these species, the VCR may serve a significant role in stabilizing gaze.

Vet-ebooks.com

CLINICAL CORRELATIONS

Vestibular Syndrome in a Dog

History

A 3-year-old male cocker spaniel is brought to your clinic. The owner states that, for the previous 2 days, the dog has held his right ear lower than his left ear. He also tends to walk in circles, clockwise to the right. You have treated this dog previously for an infection of the outer right ear.

Clinical Examination

On physical examination of the dog, you find that the outer ear infection persists. You also confirm that the dog persistently tilts his head with the right ear down and he circles to the right; you find that he has a spontaneous horizontal nystagmus. Results of the remaining physical and neurological examination are within normal limits.

Comment

Head tilt, circling, and spontaneous nystagmus constitute a common constellation of clinical signs often called the *vestibular syndrome.* It results from an abnormality in the vestibular system, usually in the membranous labyrinth. It is frequently caused by the extension of an infection from the outer and middle ear to the labyrinth of the inner ear. This results in an abnormal balance of action potential frequencies between the normal and abnormal sides of the vestibular system, causing asymmetrical stimulation of the ocular and postural reflex mechanisms normally controlled by the vestibular nuclei.

Treatment

When such labyrinthitis is caused by bacterial infection, treatment with appropriate antibiotics is often effective in eliminating the clinical signs by returning the peripheral receptor to its normal function. In cases of idiopathic vestibular syndrome in older cats and dogs, spontaneous recovery without treatment is common.

Horse With Vestibular Disease

History

A client calls you about a 6-year-old Quarter Horse mare that has not been eating or drinking well for the last few days. The horse seems depressed and has been less active than usual in the field. This is the client's favorite mare, and she just returned from training a few months ago. During the spring, she had an episode of *strangles* (also called *distemper*), but she appeared to have recovered uneventfully. Strangles is caused by the bacterium *Streptococcus equi.* She has had no other health issues.

Clinical Examination

The horse appeared to be quieter than expected. She was not very responsive to external stimuli. She had an increased temperature (102.5° F), with normal heart rate and respiratory rate. The mare also had a right-sided head tilt, positional nystagmus (nystagmus that occurs when the head is held in a non-neutral position) that changed with different head positions, and ventrolateral strabismus (eyes not focused on the same focal point) on the right side. The mare's strength was normal, but she had conscious proprioceptive deficits (on the right side worse than the left side). Because of this ataxia, she tended to drift to the right when she was standing still and as she walked.

Comment

Based on the deficits, the mare probably has central vestibular disease. Although head tilt toward, and strabismus on, the side of the lesion can also be seen with peripheral vestibular disease, positional nystagmus that changes direction with the position of the head is more characteristic of a central vestibular insult (mass or focus of infection). The conscious proprioception deficit, not ordinarily seen with peripheral vestibular defect, also suggests a CNS locus. To determine the cause of the disease, radiographs of the temporomandibular joint (TMJ) should be taken (there is some association between vertigo and TMJ abnormalities), and a cerebrospinal fluid (CSF) tap with culture should be submitted. Complete white blood cell count and biochemical profile should also be performed to help identify a cause. With the previous history of *St. equi,* the clinical signs could be caused by an abscess in the CNS/spinal cord.

Treatment

The horse has an increased white blood cell count with an increase in neutrophils and fibrinogen. The chemistry panel is normal. Radiographs are normal. The CSF tap shows increased neutrophils and protein present. Culture of the CSF demonstrates *St. equi.* Bacterial meningitis is the cause of the central vestibular signs. The horse will be treated with intravenous antibiotics, followed by oral antibiotics for an extended period of time. The prognosis is guarded, based on the severity of infection.

Practice Questions

1. The receptor organ detecting rotary acceleration and deceleration of the head is located in the:
- **a.** Utricle
- **b.** Saccule
- **c.** Ampulla of the semicircular duct
- **d.** Scala media of the cochlea
- **e.** Vestibular nuclear complex

2. Which *two* of the following are *not* generally associated with the macula?
- **a.** Otoliths
- **b.** Cupula
- **c.** Detection of linear acceleration of the head
- **d.** Hair cells
- **e.** Normal nystagmus

3. You are presented with a dog with a head tilt, compulsive circling, and spontaneous nystagmus. The most likely site of this dog's pathological lesion is the:
- **a.** Oculomotor nucleus
- **b.** Cerebral cortex
- **c.** Vestibular system
- **d.** Cervical spinal cord
- **e.** Spinal accessory (eleventh cranial) nerve

4. Which one of the following statements is *false?*
- **a.** In a given utricle, the cilia of all of the hair cells are oriented in the same horizontal direction.
- **b.** In a single vestibular hair cell, displacement of the cilia toward the largest cilium increases the firing rate of the hair cell's associated sensory neuron.
- **c.** The axons of sensory neurons synaptically associated with vestibular hair cells form the eighth cranial nerve.
- **d.** A gelatinous layer is associated with the vestibular macula.
- **e.** The vestibular nuclear complex is located in the brainstem.

Vet-ebooks.com

5. If a normal dog is sitting on a piano stool and I start to spin (accelerate) the stool to the right, which two of the following will be *false* regarding the observed nystagmus?
   a. The pattern of nystagmus observed at the start of rotation will be seen in reverse briefly after the spinning is abruptly stopped.
   b. An intact medial longitudinal fasciculus (MLF) is important for producing the nystagmus.
   c. The nystagmus will continue long after constant velocity is achieved.
   d. The eyes will drift slowly to the left, as far as they can go, and then flip rapidly back to the right.
   e. Nystagmus will often be observed long after the spinning has stopped, while the dog is stationary.

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Goldberg JM, Cullen KE. Vestibular control of the head: possible functions of the vestibulocollic reflex. *Exp Brain Res*. 2011;210(3–4):331–345.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

MacKay RJ, Van Metre DC. Diseases of the nervous system. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby Elsevier; 2015.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Rush BR. Vestibular disease. In: Reed SM, Bayly WM, Sellon DC, eds. *Equine Internal Medicine*. 4th ed. St Louis: Saunders Elsevier; 2018.

Vet-ebooks.com

12

The Cerebellum

BRADLEY G. KLEIN

KEY POINTS

1. The cerebellum constantly compares the intended movement with the actual movement and makes appropriate adjustments.
2. Cerebellar histology and phylogeny give clues to cerebellar function.
3. The vestibulocerebellum helps coordinate balance and eye movements.
4. The spinocerebellum helps coordinate muscle tone as well as limb movement.
5. The cerebrocerebellum helps with planning coordinated, properly timed movement sequences.
6. The cerebellum plays a role in motor learning.
7. Cerebellar disease causes abnormalities of movement and further illuminates cerebellar function.

The preceding chapters, which describe the physiology of movement, discuss the function of lower motor neurons through which the central nervous system (CNS) can initiate and control movement by initiating contraction of skeletal muscle. The corticospinal system and the descending brainstem motor system are described in those previous chapters as major subgroups of upper motor neurons that influence the lower motor neurons. More medial portions of those systems coursing through the spinal cord are primarily responsible for the control of axial and proximal antigravity extensor muscles. The more lateral portions primarily control more skilled, learned, voluntary movements caused by contraction of distal flexor muscles. This chapter describes the function of the cerebellum, part of another subgrouping of upper motor neurons critical for proper movement.

The *cerebellum* (Latin, "little brain") is caudal to the cerebral cortex and dorsal to the brainstem (Fig. 12.1A). Although it constitutes only about 10% of the gross brain volume because of its highly folded structure, the cerebellum contains more than half of all CNS neurons. The raised, midline region of the cerebellum is called the *vermis* and is bounded laterally, on both sides, by the *cerebellar hemispheres* (Fig. 12.1B). The hemispheres can be further subdivided into intermediate (closer to the vermis) and lateral (farther from the vermis) zones with respect to functional considerations discussed later. A flocculonodular lobe is normally curled up under the caudal portion of the cerebellum.

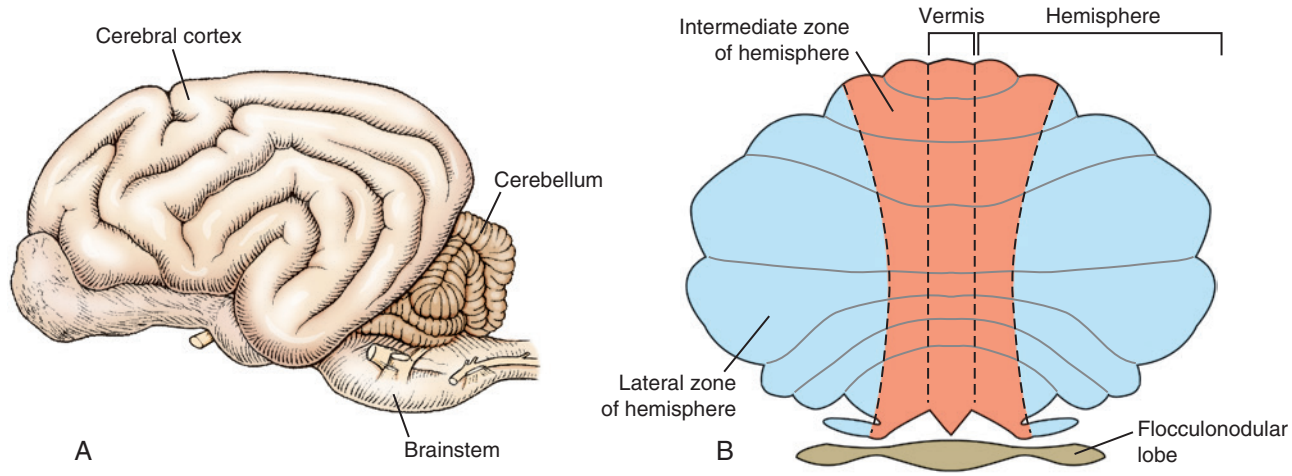
The outer layer of cerebellar gray matter, the *cerebellar cortex,* has a highly regular, three-layered, histological appearance, which suggests that all cerebellar regions may perform a common underlying task. Like the cerebral cortex, the particular inputs to a given region of cerebellar cortex, and the particular output targets that it influences, in large part account for the functional differences between cerebellar regions. In addition to the cerebellar cortex and the cerebellar white matter axons entering and leaving the cortex, a group of deep *cerebellar nuclei* is embedded within the cerebellar white matter (Fig. 12.2). The cells of these nuclei are a principal origin of the axons leaving the cerebellum. Two large pairs of white matter stalks, the *rostral* and *middle cerebellar peduncles,* respectively carry axons out from and into the cerebellum. A third, smaller pair of cerebellar peduncles, the *caudal cerebellar peduncles,* carry axons both into and out from the cerebellum.

The cerebellum is not necessary for the initiation of movement. Muscle strength remains largely intact with complete destruction of the cerebellum. However, the cerebellum plays a crucial role in the timing and coordination of movement initiated by the parts of the motor system hierarchy discussed in Chapter 10. It does so by adjusting and modulating the output of the motor cortices, corticospinal tract, descending brainstem motor pathways, and spinal cord. Lesions of the cerebellum lead to major clinical deficits in the precision and grace with which movement is accomplished.

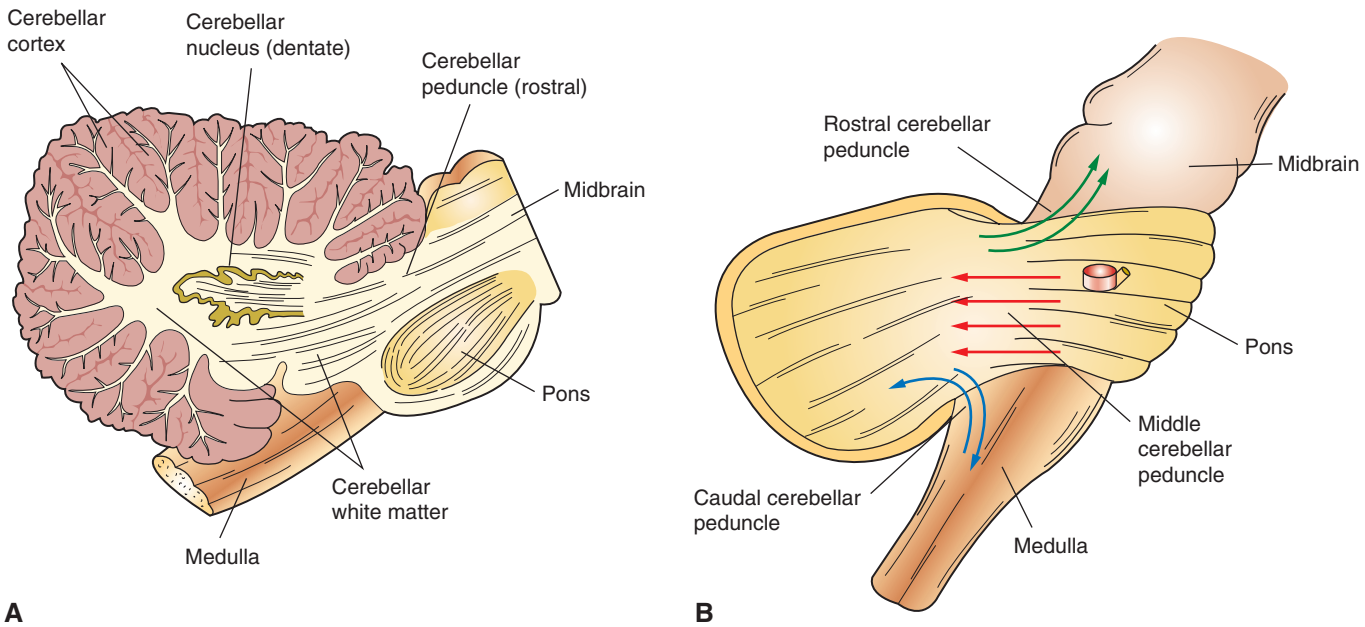
The Cerebellum Constantly Compares the Intended Movement With the Actual Movement and Makes Appropriate Adjustments

In performing the essential role of adjusting the timing and coordination of movement, the cerebellum first receives information from components of the motor system hierarchy about the movement it has commanded. It also receives information from muscle spindles, the vestibular and visual systems, and other sensory receptors about the movement the body is actually performing. When the intended movement and the actual movement are not the same, the cerebellum's job is to perform the adjustments necessary to make them the same. For example, if a cat's brain intends that its mouth move to a piece of food in a dish, but sensory receptors inform the cerebellum that the trajectory of the head will cause the mouth to miss the dish, the cerebellum makes appropriate adjustments in the components of the motor system hierarchy to correct the head's trajectory. The correction can be made to the movement in progress and to the plan for subsequent movement.

Vet-ebooks.com



• **Fig. 12.1** (A) The cerebellum (Latin, "little brain") is caudal to the cerebral hemispheres and dorsal to the brainstem. (Redrawn from Miller ME, Christiansen GC, Evans HE. *The Anatomy of the Dog.* Philadelphia: Saunders; 1964.). (B) View of dorsal surface of the unfurled cerebellum showing two principal anatomical regions: the normally raised midline vermis region and the more lateral hemispheres. The hemispheres can be further subdivided into lateral and intermediate zones with respect to functional considerations discussed later. The flocculonodular lobe is normally curled up under the caudal portion of the cerebellum. (Modified from Kandel ER, Schwartz JH, Jessell TM, eds. *Principles of Neural Science.* 2nd ed. New York: Elsevier Science & Technology; 1985.)



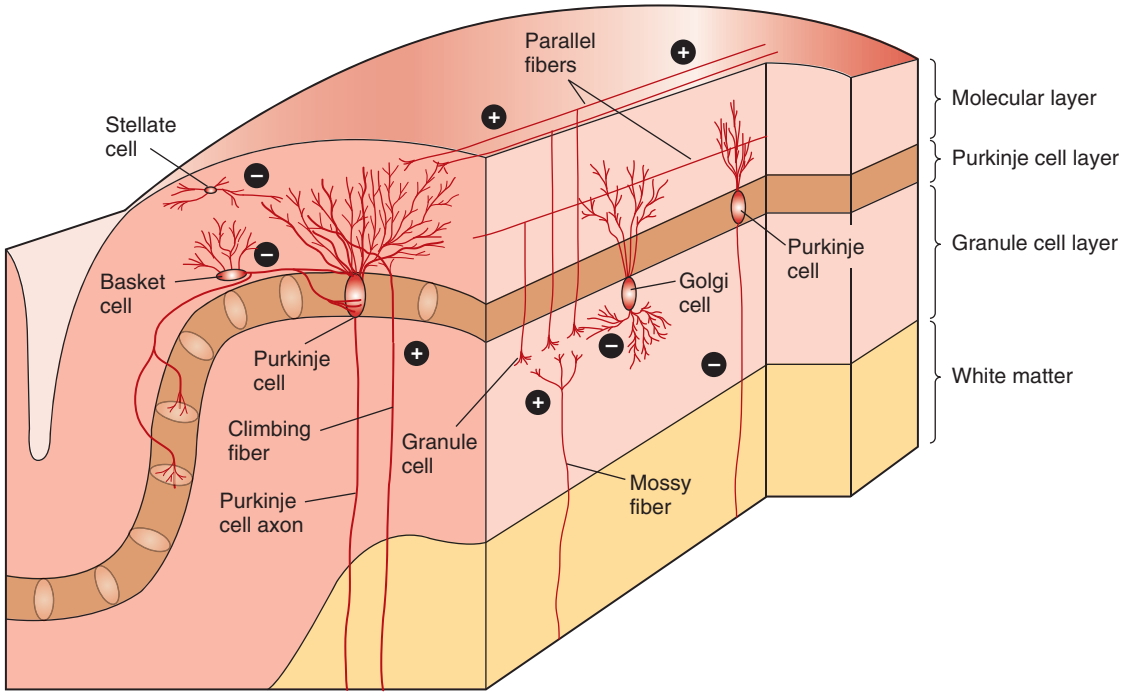
• **Fig. 12.2** (A) Midsagittal section through the brainstem region showing the internal organization of the cerebellum. (B) A lateral view of the brainstem region emphasizing the cerebellar peduncles and the principal directions that axons travel within them.

Cerebellar Histology and Phylogeny Give Clues to Cerebellar Function

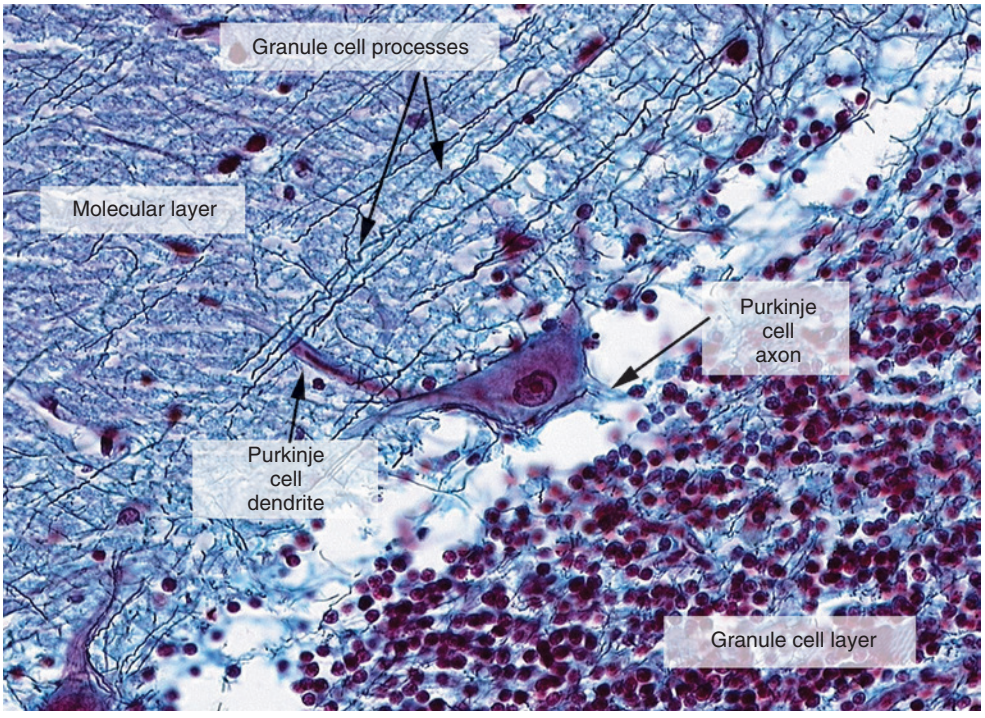
The cortex throughout the cerebellum is quite uniform and consists of three layers and five major types of neurons: stellate, basket, Golgi, granule, and Purkinje cells (Fig. 12.3). The outermost layer is the *molecular layer* and consists primarily of granule cell axons, known as *parallel fibers* (Fig. 12.4). The molecular layer also contains dendrites of neurons located in deeper layers, as well as scattered inhibitory interneurons. The latter are the *stellate* and *basket cells.*

The middle *Purkinje cell layer* of cerebellar cortex consists of the large cell bodies of Purkinje neurons, which have a flat but extremely expansive dendritic field that extends into the molecular layer (see Figs. 12.3 and 12.4). This dendritic field is oriented at right angles to the parallel fibers. Therefore a Purkinje cell is contacted by an expansive array of excitatory parallel fiber axons of granule cells, and an individual parallel fiber contacts the dendrites of many Purkinje cells. The stellate and basket cell inhibitory interneurons, noted earlier, can act to refine, or prune, this extensive spatial pattern of Purkinje cell activation by parallel fibers.

Vet-ebooks.com



• **Fig. 12.3** Five major types of neurons are organized into three layers in the cerebellar cortex. A single, vertically oriented cerebellar folium is sectioned in both sagittal and transverse planes to illustrate the general organization of the cerebellar cortex. A positive sign denotes an excitatory effect of a neural element on its postsynaptic target. A negative sign denotes an inhibitory effect of a neural element on its postsynaptic target. (Modified from Kandel ER, Schwartz JH, eds. *Principles of Neural Science,* 4th ed. New York: McGraw-Hill; 2000.)



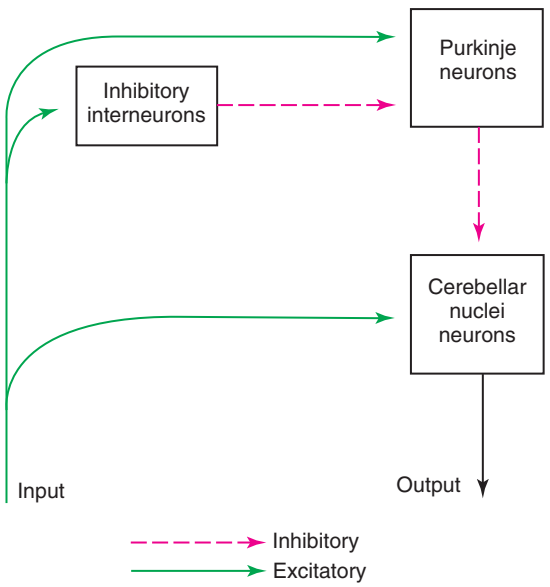
• **Fig. 12.4** High-power photomicrograph of the three layers of the cerebellar cortex: granule cell layer, Purkinje cell layer (not labeled), and molecular layer. The image provides a good example of how the parallel fibers, labeled *granule cell processes,* cross the dendritic region of the Purkinje cell in the molecular layer. However, the elaborate branching pattern of the Purkinje cell dendrites is not apparent with this stain. (Image courtesy Dr. Tom Caceci, Department of Biomedical Sciences and Pathobiology, College of Veterinary Medicine, Virginia Tech.)

Vet-ebooks.com

The innermost *granule cell layer* of cerebellar cortex contains the vast number of granule cell somas that give rise to the parallel fibers (see Figs. 12.3 and 12.4). This layer also contains occasional Golgi cell bodies. These are inhibitory interneurons that can regulate the overall level of excitation of the Purkinje cells by the granule cell parallel fibers.

Axons of the Purkinje neurons go to the deep cerebellar nuclei, located outside of the cerebellar cortex, embedded in the cerebellar white matter (see Fig. 12.2). The Purkinje cells are the only output neurons of the cerebellar cortex and are all inhibitory. They can inhibit the spontaneously active neurons of the deep cerebellar nuclei, whose axons leave the cerebellum. This selective inhibition represents a sensitive temporal refinement of cerebellar processing that supplements the spatial refinement noted earlier. The cerebellar output neurons participate in regulating the activity of brainstem motor pathways and motor cortices involved in the execution and planning of movement.

The two primary groups of input axons to the cerebellum are the *mossy fiber* and *climbing fiber* axons (see Fig. 12.3). Both are excitatory; they cause excitatory postsynaptic potentials (EPSPs) within the cerebellar cortex and, through collateral axons, within the deep cerebellar nuclei (Fig. 12.5). The mossy and climbing fibers collectively carry information from components of the motor system hierarchy and from peripheral sensory receptors regarding the planning, initiation, and execution of the movement. The shorter input/output circuit of the cerebellum (deep excitatory loop) consists of the climbing and mossy fiber stimulation to the deep cerebellar nuclei, whose output in turn leaves the cerebellum to modify components of the motor system hierarchy. However, this output of the deep cerebellar nuclei is refined by inhibition from Purkinje cell axons that originate in cerebellar cortex. The Purkinje cell inhibition of deep cerebellar nuclei is based on the cerebellar cortex's own integration of mossy and climbing fiber inputs (cortical inhibitory loop). In other words, the same information coming into the cerebellum that drives the cerebellar nuclei is also processed by cerebellar cortex, whose resulting Purkinje cell output refines or "sculpts" the output of the cerebellar nuclei that project to components of the motor system. Within the cerebellar cortex, inhibitory interneurons help to refine or "sculpt" the Purkinje cell output of cerebellar cortex.



• **Fig. 12.5** Input/output organization of the cerebellum. See text for explanation.

Although the cortical synaptology is understood, just how the cerebellum integrates movement feedback with the motor plan and then modifies the output of the deep nuclear neurons is not clear. As noted earlier, because the histological appearance of the cortex is similar throughout the cerebellum, it seems likely that a similar underlying processing mechanism exists in the cortex, regardless of the cerebellar region. However, within the cerebellum, regional variation of inputs from and outputs to different parts of the nervous system renders different motor results from different cerebellar regions. This can be thought of as effectively partitioning the cerebellum into different functional modules. The cerebellum can be divided into three distinct regions from both a functional perspective and a phylogenetic perspective: the *vestibulocerebellum,* the *spinocerebellum,* and the *cerebrocerebellum* (Fig. 12.6).

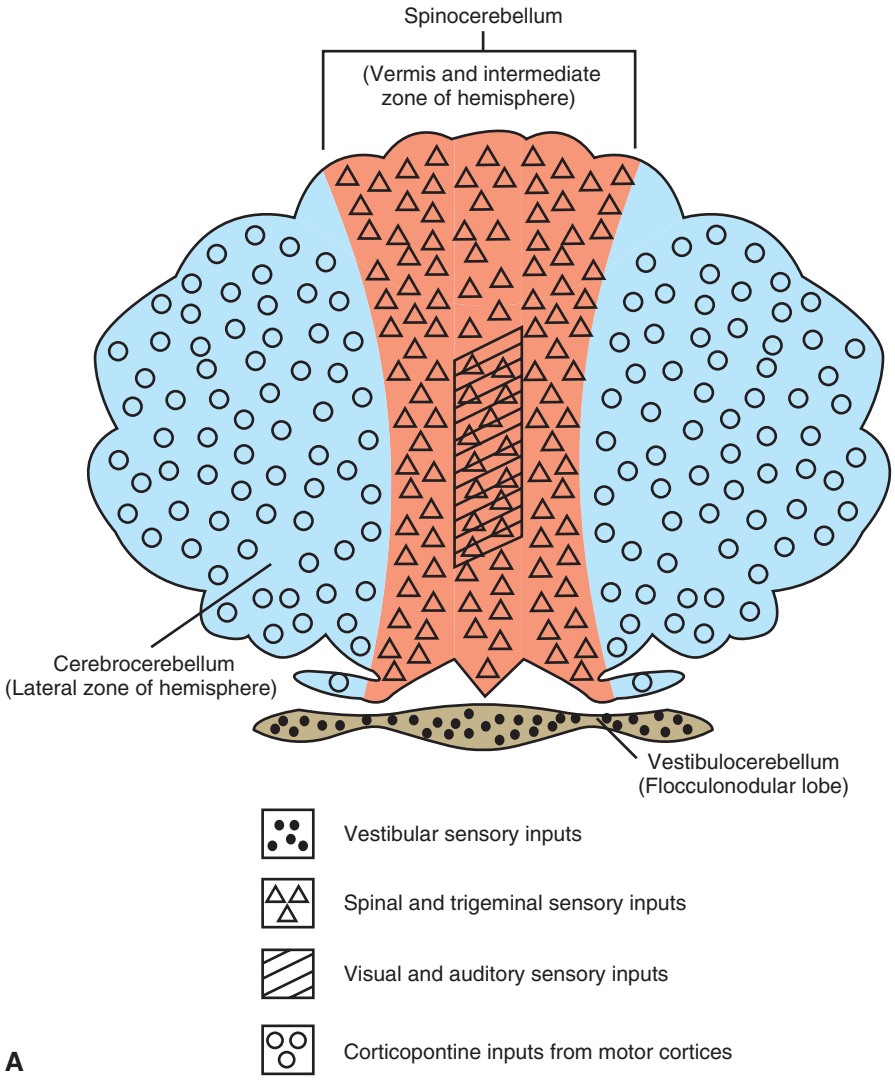
The Vestibulocerebellum Helps Coordinate Balance and Eye Movements

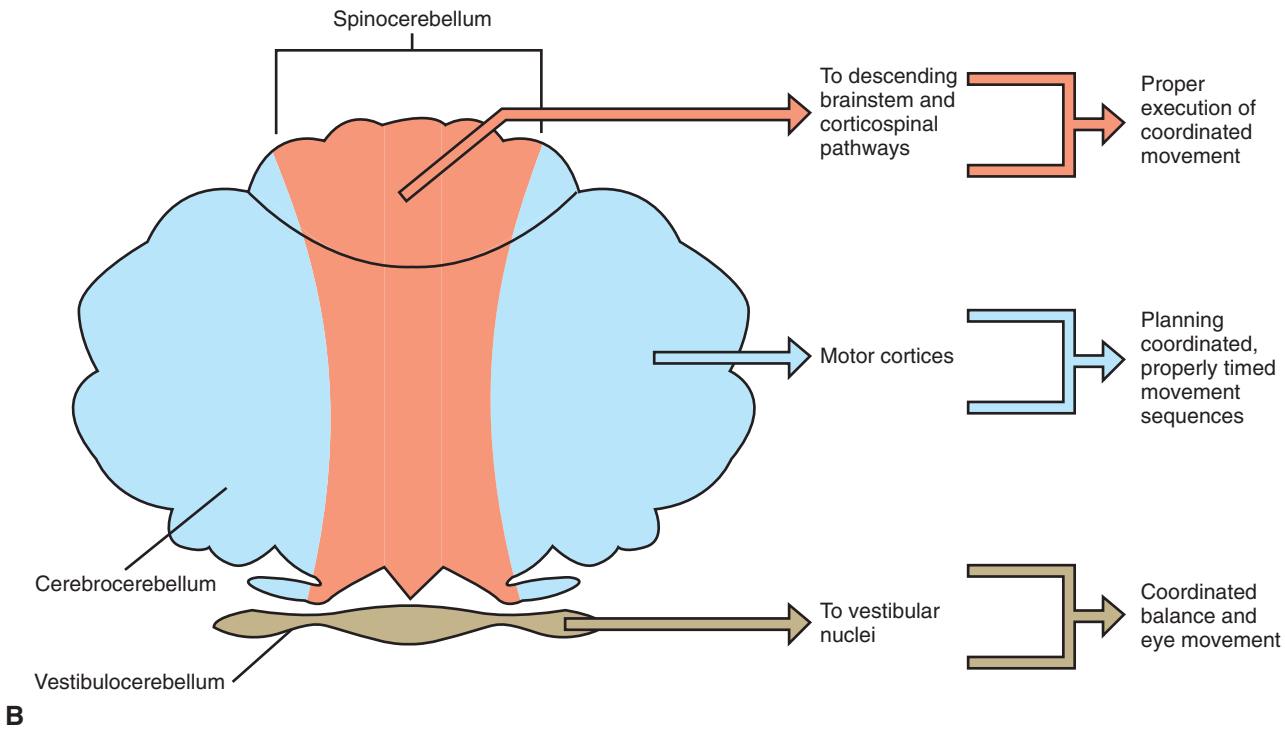
The vestibulocerebellum occupies the flocculonodular lobe and receives most of its afferent input from the vestibular system, by way of the caudal cerebellar peduncles (see Figs. 12.1, 12.2, and 12.6). Its efferent output returns to the vestibular nuclei through these same peduncles either directly from cerebellar cortex or by way of the deep cerebellar nuclei (specifically the fastigial nucleus). The cerebellar output to the vestibular nuclei helps to coordinate the axial and proximal muscles controlling balance, by way of the vestibulospinal tract, and helps to coordinate head and eye movements by way of the medial longitudinal fasciculus (see Chapter 11). In short, the vestibulocerebellum adjusts the coordination of vestibular reflexes. Because this part of the cerebellum was the first to appear in vertebrate evolution, it is sometimes called the *archicerebellum.*

The Spinocerebellum Helps Coordinate Muscle Tone as Well as Limb Movement

The spinocerebellum extends rostrocaudally along the vermis and the intermediate zone of the hemisphere (see Figs. 12.1 and 12.6). It receives sensory inputs from muscle and cutaneous receptors through the spinal cord and trigeminal nuclei. It also receives input from neurons in spinal reflex circuits, some of which receive commands from corticospinal or descending brainstem motor pathways. The cerebellar afferents coming from the spinal cord form several spinocerebellar tracts, which mostly enter the cerebellum through the caudal cerebellar peduncle (see Fig. 12.2). Some input to the spinocerebellar region also comes directly from the primary motor and primary somatosensory cortices. The spinocerebellum therefore receives information about commands for movement and significant feedback information about the execution of the movement itself. Its outputs travel, by way of its deep cerebellar nuclei (specifically the fastigial and interpositus), to brainstem nuclei controlling the antigravity musculature (e.g., reticular nuclei), as well as to a brainstem nucleus controlling distal limb musculature (e.g., red nucleus). Some of the spinocerebellar output travels to the primary motor cortex, particularly important for voluntary limb movement, by way of the thalamus. Through these output projections, which leave the cerebellum through the caudal and rostral peduncles, the spinocerebellum can adjust the timing and coordination of "in progress" movement and muscle tone. Such adjustments are presumably based on a comparison of spinocerebellar input regarding the movement command (e.g., from primary motor

Vet-ebooks.com





• **Fig. 12.6** (A) The cerebellum can be divided into three distinct functional regions (or modules), illustrated here with their respective major inputs. The relation between these functional regions and the anatomical designations illustrated in Fig. 12.1B is also indicated. (B) Major output targets and general roles of the three functional regions of the cerebellum. (Modified from Kandel ER, Schwartz JH, eds. *Principles of Neural Science.* 2nd ed. New York: Elsevier Science & Technology; 1985.)

Vet-ebooks.com

cortex) with feedback about the ongoing movement itself (e.g., from muscle, joint, and skin inputs). Because this portion of the cerebellum appeared next in evolution, it is sometimes called the *paleocerebellum.*

The Cerebrocerebellum Helps With Planning Coordinated, Properly Timed Movement Sequences

The cerebrocerebellum occupies the lateral zone of the cerebellar hemispheres (see Figs. 12.1 and 12.6). This region also receives input from the primary motor cortex, but more important, receives a substantial input from premotor and supplementary motor cortices. These cortical inputs reach the cerebellum by way of the corticopontine-cerebellar system, which enters the cerebellum through the massive middle cerebellar peduncles. The cerebrocerebellum does not have direct access to information from peripheral receptors like the spinocerebellum. Its outputs, which travel through the rostral cerebellar peduncles, return to the motor cortices by way of the thalamus. Therefore the cerebrocerebellum is part of a communication loop with regions of motor cortex involved in the planning of, and preparation for, movement. Whereas the spinocerebellum helps coordinate the "in progress" execution of movement, the cerebrocerebellum helps the motor cortices with planning ahead for the next appropriate movement so there will be smooth and appropriately timed transitions between components of a movement sequence. The dramatic growth of the cerebrocerebellum and cerebral cortex was the major phylogenetic addition to the brain during primate evolution, and thus it is often called the *neocerebellum.* Presumably this is linked to the primate's ability to perform graceful, intricate, appropriately timed voluntary movements, such as coordinated finger movements as well as mouth and tongue movements that can produce speech.

The Cerebellum Plays a Role in Motor Learning

Several lines of evidence suggest that the cerebellum plays a significant role in motor learning. For example, functional magnetic resonance imaging (fMRI) studies have shown that the cerebellum is very active when learning a new sequence of movements, but it is not as active when the movement becomes relatively automatic. This suggests that the cerebellum is involved in the *transition* from concentrating on learning a new motor skill, such as where to place individual fingers on piano keys to form a chord, to being able to perform that skill automatically, with limited thought. Some reflex behaviors, such as the vestibuloocular reflex (see Chapter 11), although automatic, need to be fine-tuned or adjusted (e.g., with respect to the amount of eye rotation necessary to counteract a given amount of head rotation to keep the gaze fixed on a target) as the proportions of the head change during growth. Damage to certain regions of the cerebellum can prevent this type of adaptive adjustment. In addition, some forms of associative learning, such as some classically conditioned responses, can be abolished after cerebellar lesions. The ability to make motor adaptations after alterations of the visual world, such as learning to throw darts accurately after wearing prism glasses, can also be severely impaired in individuals with cerebellar damage.

Structural and functional changes in cerebellar circuitry have also been observed during motor learning. For example, increases in the number of parallel fiber and climbing fiber synaptic contacts on Purkinje cells have been observed following the learning of complex motor behavior. Furthermore, simultaneous activation of these two types of fibers synapsing on a Purkinje cell, where it is presumed that the climbing fiber is carrying a motor error signal, can produce a long-term depression of Purkinje cell activity making the cell less responsive to subsequent mossy fiber signals. Such depression can have a profound effect on the activity of neurons of the deep cerebellar nuclei that leave the cerebellum to control components of the motor hierarchy.

Cerebellar Disease Causes Abnormalities of Movement and Further Illuminates Cerebellar Function

As discussed, the cerebellum constantly compares the intended movement with the actual movement and makes appropriate adjustments. In cerebellar disease, these appropriate adjustments are not made, resulting in a variety of movement disorders. Affected animals often place their paws far apart (*wide-based gait*) and walk in an uncoordinated manner (*ataxia*), which reflects the inability of the vestibulocerebellum and spinocerebellum to coordinate balance and movement of the axial skeleton. Affected animals also have various degrees of *dysmetria* (inappropriate measure of muscular contraction), in which movements either continue too long or not long enough. This is often manifested as difficulty in bringing the muzzle to a fixed point in space, such as a food dish, and as exaggerated "goose stepping" walking movements. *Asynergia,* a failure of the components of a complex, multiple-joint movement to occur in a coordinated fashion, may also be seen. It is particularly characteristic of damage to the cerebrocerebellum. *Intention tremor* (action tremor), an oscillating movement disorder that is worse when the animal is moving, especially near the end of the movement, is common in cerebellar disease as well. Intention tremors are much less severe when the animal is relaxed and not moving and worsen when a movement is being performed. In animals, intention tremors seem worse in the head and axial (proximal) antigravity muscles. If the vestibulocerebellum is damaged, nystagmus may also be seen (see Chapter 11). These commonly associated clinical signs resulting from cerebellar disease exemplify how the mechanism of disease can be understood through knowledge of normal physiology.

Some clinical studies in humans suggest that the cerebellum may play a role in cognitive functioning as well. Individuals with cerebellar damage have problems making rhythmic movements, which might be expected, but they also appear to have problems judging or perceiving rhythm. Problems estimating the velocity of visual stimuli have also been reported. Some individuals with cerebellar damage also appear to have problems rapidly shifting their attention from one stimulus to another. Some imaging studies have shown increased activity in the cerebellum, in normal individuals, when counting silently or while imagining movement. The study of the role of the cerebellum in cognitive function is still fairly recent compared with the study of its motor function. Finally, several of the motor and cognitive deficits discussed above seem to reflect a relationship to timing. The cerebellum seems well suited for a timing role given its cellular circuitry: long parallel fibers of granule cells that can perpendicularly intersect and excite the dendritic trees of many Purkinje cells in a fixed temporal sequence (see Fig. 12.3).

Vet-ebooks.com

CLINICAL CORRELATIONS

Cerebellar Hypoplasia

History

An 11-week-old female barn kitten is brought to your clinic for examination. The owner states that this kitten and several others in the litter have been uncoordinated since they began to walk.

Clinical Examination

Physical examination abnormalities are limited to the nervous system. The kitten is bright, alert, and responsive, and seems to be of normal size for her age. All cranial nerve and spinal segmental reflexes and intersegmental responses are within normal limits. There is no atrophy. The kitten is uncoordinated (ataxic) when she moves and tends to raise her front paws higher than normal when walking ("goose stepping" hypermetria). She holds her paws far apart when walking. There are coarse, rhythmic movements of her head and proximal antigravity muscles that are absent at rest and severe when she is attempting a precise movement, such as getting her head to a food dish (intention tremor). Her complete blood count and serum chemistry results are within normal limits.

Comment

This kitten demonstrates classic signs of cerebellar disease. The cerebellum constantly compares the intended movement with the actual movement and, when these are not the same, makes the appropriate adjustments. When the cerebellum cannot do this, movement disorders characterized by wide-based gaits, ataxia, dysmetria, asynergia, and intention tremor occur. These movement disorders are worse with precise movement and nearly absent at rest.

This kitten's clinical signs are likely caused by *cerebellar hypoplasia,* in which the cerebellum never developed completely in utero. The in utero infection of *feline panleukopenia virus* results in destruction of the actively dividing granule cells (neurons), with an underdevelopment (hypoplasia) of the granule cell layer of the cerebellum. Purkinje cells may also be affected. Barn cats are often not vaccinated for this disease, and often several kittens in a litter are affected.

Treatment

There is no treatment for cerebellar hypoplasia caused by such in utero viral infection. It is not a progressive disease, and if affected kittens are kept in a fairly safe environment, they can have a normal life span.

Newborn Calf Unable to Rise

History

A producer calls to ask about an Angus heifer calf born earlier today that has not stood. The calf makes efforts but does not seem coordinated enough to stand. The producer has fed the calf with colostrum by tube and wants her examined. This is the second calf this season that has had this problem. They euthanized the other calf after she had not improved over 2 to 3 days. The calves are very valuable, and the producer would like to keep this calf in the herd. Further questioning of the owner reveals an increased percentage of abortions this year. She also bought several new replacement cows last fall that have been introduced to the herd.

Clinical Examination

The calf has a normal temperature, pulse, and respiration. She appears responsive to noise, almost hyperexcitable. There is no evidence of trauma. When the calf is placed in a standing position, she sways back and forth; she tries to maintain a base-wide stance but sometimes falls over or backs up. She appears extremely uncoordinated and hypermetric (movement continuing too long). She scores a 4/5 on ataxia, with 5 being the most severe. Other abnormalities include a greatly delayed menace response and bumping into things when she tries to walk. When she is laid back down and her reflexes are assessed, the calf is hyperreflexic in all her responses.

Comment

Based on this history of the herd and the calf, this herd most likely has a problem with *bovine viral diarrhea virus* (BVDV). This was likely introduced by replacement cows. BVDV would explain the abortions as well as the two affected calves. With BVDV, the virus infects the germinal cells within the cerebellum and kills the Purkinje cells. Infection of these cells results in local inflammation, cell death, hemorrhage, and necrosis. Because of the damage to the Purkinje cells, inhibitory function is disrupted, which affects the vestibulocerebellum, spinocerebellum, and cerebrocerebellum. The deficits in these areas are associated with clinical signs of abnormalities in balance (vestibulocerebellum), eye movement (vestibulocerebellum), ataxia and base-wide stances (vestibulo- and spinocerebellum), and motor coordination and sequencing (spino- and cerebrocerebellum).

Treatment

Because BVDV causes irreversible cell damage, the prognosis for this calf is poor. Even if treatment were available, the calf most likely has BVDV and would shed virus if reintroduced into the herd. Euthanasia is the best option for this calf. The owner should screen the herd and identify infected and persistently infected (PI) animals. Additionally, vaccination with a live versus a killed BVDV vaccine may improve overall outcome.

Practice Questions

1. Which of the following is principally involved in planning ahead for the next appropriate movement?
- **a.** Vestibulocerebellum
- **b.** Spinocerebellum
- **c.** Cerebrocerebellum
- **d.** Archicerebellum
- **e.** Both a and b

2. Loss of the cerebellum causes immediately obvious sensory deficits and prevents the initiation of movement.
- **a.** True
- **b.** False

3. Which of the following is *true* regarding cerebellar Purkinje cells?
- **a.** They are located in the cerebellar cortex.
- **b.** They have large cell bodies.
- **c.** They have an extensive dendritic tree.
- **d.** When active, they inhibit the activity of cells in deep cerebellar nuclei, whose axons leave the cerebellum.
- **e.** All the above.

4. Loss of the cerebellum causes loss of the muscle stretch reflex.
- **a.** True
- **b.** False

Vet-ebooks.com

5. Cats with congenital malformations of the cerebellum often have ataxia, intention tremor, and wide gait.
a. True
b. False

6. Which one of the following regarding the spinocerebellum is *false*?
a. It receives inputs from muscle and cutaneous receptors.
b. It occupies the lateral zone of the cerebellar hemispheres.
c. It receives some input from primary somatosensory cortex.
d. It is sometimes called the paleocerebellum.
e. It receives some input from primary motor cortex.

7. With regard to symptoms of cerebellar dysfunction, movements that continue either too long, or not long enough, are most correctly referred to as which one of the following?
a. Ataxia
b. Asynergia
c. Hypoplasia
d. Dysmetria

Bibliography

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Jennings DP. Supraspinal control of posture and movement. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Vet-ebooks.com

13

The Autonomic Nervous System

BRADLEY G. KLEIN

KEY POINTS

1. The peripheral autonomic nervous system differs from the somatic motor system in a number of important ways.
2. The peripheral autonomic nervous system has two subdivisions that originate in the central nervous system and one that does not.
3. The sympathetic nervous system originates from the thoracolumbar spinal cord.
4. The parasympathetic nervous system arises from the brainstem and sacral spinal cord.
5. Most sympathetic and parasympathetic neurons secrete either acetylcholine or norepinephrine as a neurotransmitter.
6. Acetylcholine and norepinephrine have different postsynaptic receptors.
7. Neurotransmitters other than acetylcholine and norepinephrine play some role in peripheral autonomic function.
8. There are general differences in sympathetic and parasympathetic function.
9. Visceral afferent (sensory) neurons play an important role in autonomic nervous system function.
10. The autonomic nervous system participates in many homeostatic reflexes.
11. Preganglionic neurons are influenced by many regions of the brain.
12. There are interactions between the autonomic nervous system and the immune system.

The autonomic nervous system (ANS) is a part of the nervous system that is generally not under voluntary control, nor is the organism usually conscious of its operation. The ANS is commonly defined as a *peripheral motor system* innervating smooth muscle, cardiac muscle, glandular tissue, and the organs of the body cavity, known as *viscera* (e.g., stomach, urinary bladder), that these tissues often comprise. It should be kept in mind, however, that these peripheral targets and their motor innervation are usually part of reflex pathways that also include visceral afferents (see Chapter 3) and central nervous system (CNS) structures (e.g., hypothalamus), both of which are sometimes included in broader definitions of the ANS.

A principal function of the ANS is maintaining the stability of the body's internal environment, or *homeostasis.* Toward this end, it regulates such functions as blood pressure, heart rate, intestinal motility, bladder emptying, sweating, and the diameter of the eye's pupil. The ANS has unique anatomy, synaptic transmission, and effect on its various target organs. This chapter describes the general anatomy and function of the ANS. It mainly focuses on its peripheral motor aspects given their relevance to understanding the actions of the large number of drugs that affect the ANS. However, visceral afferents and CNS regulation of autonomic function are touched upon as well. The specific effect of the ANS on particular target organs is described in the chapters for each of the body's systems.

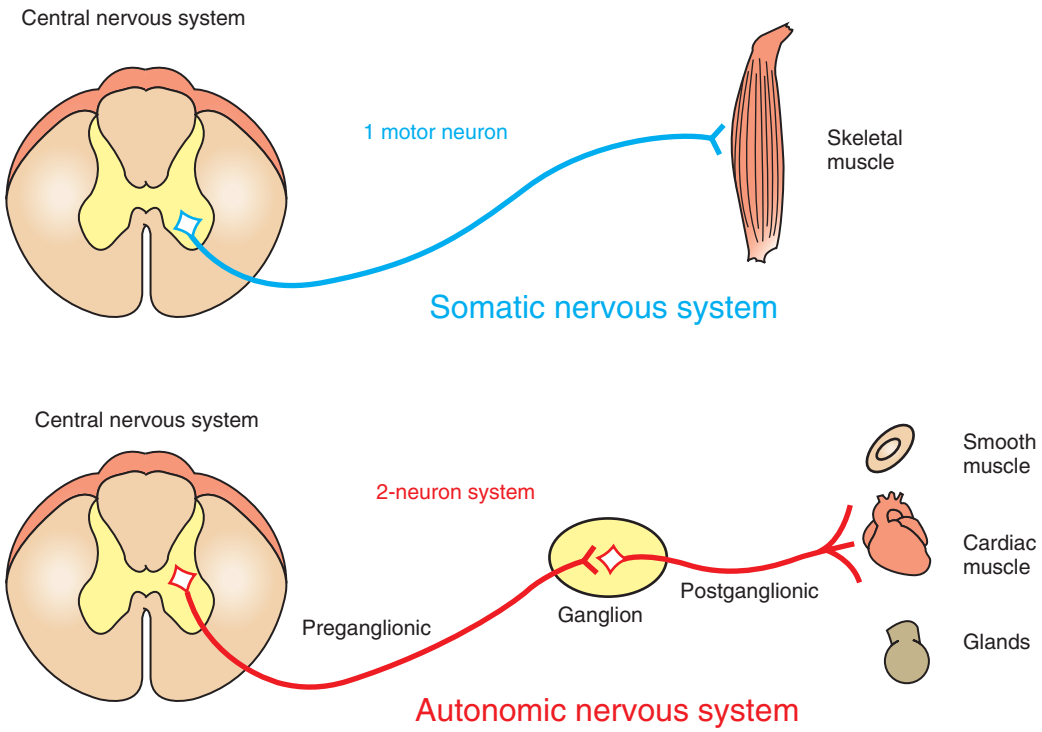
The Peripheral Autonomic Nervous System Differs From the Somatic Motor System in a Number of Important Ways

The ANS differs from the somatic motor system in its target organs, in the number of neurons in its peripheral circuit, and in the nature of the synapse at the target organ. The somatic motor system innervates skeletal muscle, which is the muscle responsible for all movements of the body, as described in Chapters 5 and 6. In contrast, the ANS innervates smooth muscle, cardiac muscle, and glandular tissue (Fig. 13.1). Cardiac muscle is the muscle of the heart (see Chapter 19). Smooth muscle is the muscle in blood vessels, in most of the gastrointestinal tract, in the bladder, and in other hollow visceral structures. Gland cells can also be part of visceral organs, as well as composing nonvisceral glands (e.g., salivary glands, lacrimal gland).

The ANS also differs in the number of neurons it has in the peripheral nervous system pathway to its target (see Fig. 13.1). The somatic nervous system has one neuron whose cell body is located in the CNS and whose axon extends, uninterrupted, to the skeletal muscle, where the peripheral chemical synapse occurs. In contrast, the ANS has two peripheral neurons. The first, called a *preganglionic neuron,* also has its cell body in the CNS, but its axon innervates a second neuron in the chain, called the *postganglionic neuron.* The latter's cell body is in a peripheral structure called a *ganglion,* a collection of neuronal cell bodies outside the CNS. There are chemically mediated synapses both between the preganglionic and postganglionic neurons, and between the postganglionic neuron and the cells of its target organ.

The ANS also differs from the somatic motor system in the amount of *myelin* along the peripheral axons; the autonomic postganglionic neurons usually have slowly conducting, unmyelinated axons. In addition, somatic motor neurons always excite their skeletal muscle targets, whereas the autonomic postganglionic neurons can either excite or inhibit their targets. Furthermore, unlike the narrow synaptic cleft at the focal neuromuscular junction of a skeletal muscle cell, ANS target cells are often activated at a greater distance by a highly branched postganglionic neuron with synaptic boutons (called *varicosities*) distributed all along the length

Vet-ebooks.com



• **Fig. 13.1** The autonomic nervous system (ANS) differs from the somatic motor system in the number of neurons that it has in the peripheral nervous system pathway to its target. The somatic motor system has one neuron, whose cell body is located in the central nervous system (CNS) and whose axon extends, uninterrupted, to the skeletal muscle, where the peripheral chemical synapse occurs. In contrast, the ANS has two neurons in the path from the CNS to the target. The first, called a *preganglionic neuron,* also has its cell body in the CNS, but its axon innervates a second neuron in the chain, called the *postganglionic neuron.* Its cell body is in a peripheral structure called a *ganglion.*

of these branches. This can contribute to a longer latency for, and greater spatial distribution of, postsynaptic cell activation by autonomic postganglionic neurons. However, this latency and spatial distribution can vary depending on the nature of the particular autonomic target tissue, with some distances between varicosities and their targets being quite similar to that found at the skeletal neuromuscular junction. Some aspects of the neural activation of autonomic target tissue were discussed in Chapter 6.

The Peripheral Autonomic Nervous System Has Two Subdivisions That Originate in the Central Nervous System and One That Does Not

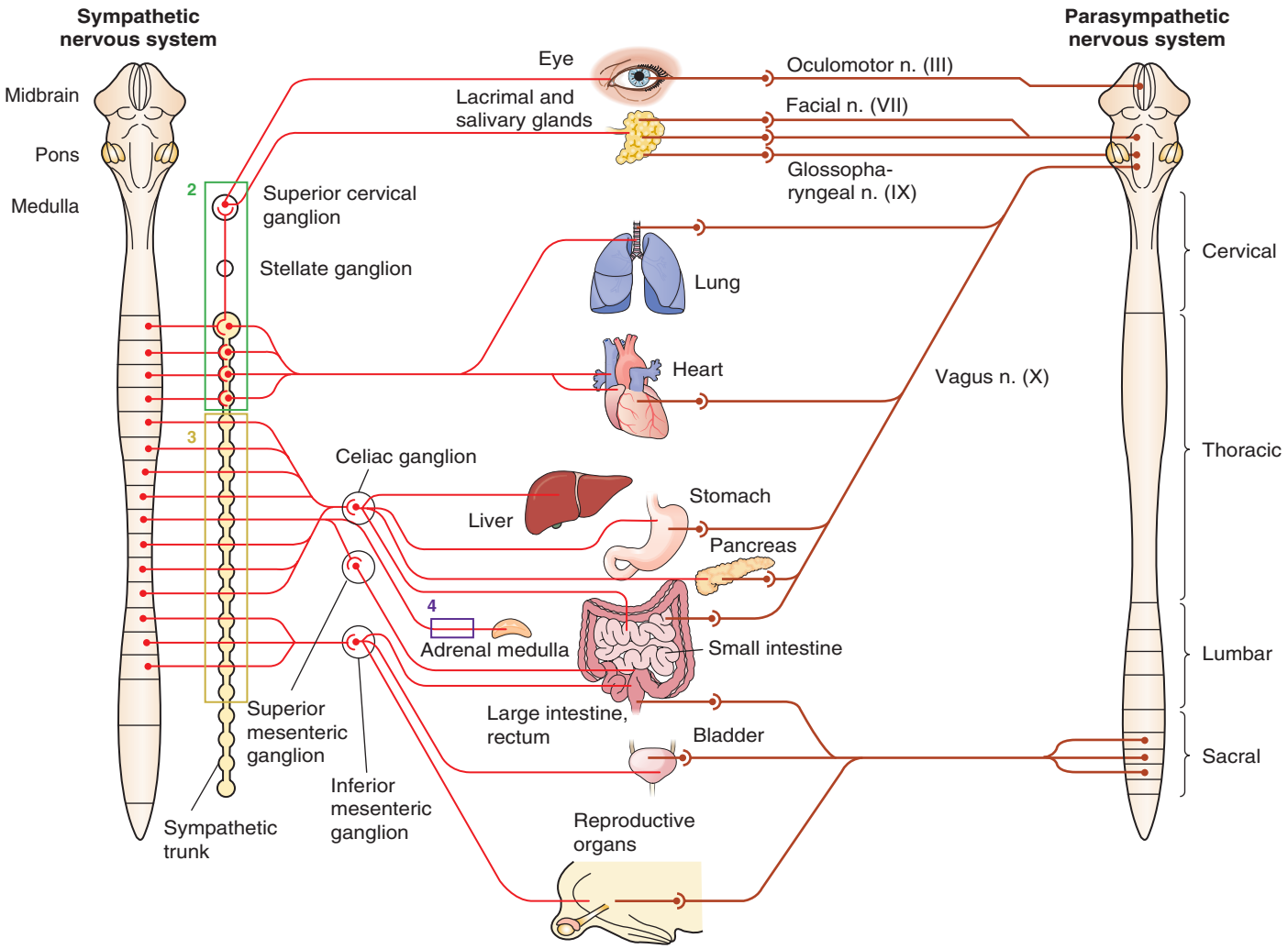
The peripheral ANS is divided into two major subdivisions based on the respective CNS origin of their preganglionic neurons and on their synaptic transmitters at the target organ. These two subdivisions are the *sympathetic nervous system* and the *parasympathetic nervous system.* The *enteric nervous system* can be considered a third subdivision of the peripheral ANS. It is an extensive network of interconnected sensory, motor, and interneurons within the gut (gastrointestinal tract) wall that can control gut function independently of the CNS. However, these neurons can also be influenced by the CNS through input from the sympathetic and parasympathetic subdivisions. The enteric nervous system will be discussed in more detail in reference to the regulation of gastrointestinal function in Chapter 27.

The Sympathetic Nervous System Originates From the Thoracolumbar Spinal Cord

The sympathetic nervous system generally has short preganglionic and long postganglionic axons. Preganglionic axons of the sympathetic nervous system leave the spinal cord by way of the ventral roots of the first thoracic through the third or fourth lumbar spinal nerves (Fig. 13.2). For this reason, the sympathetic nervous system is often called the *thoracolumbar system.* The preganglionic axons pass through the ventral root and then a communicating branch (*white ramus*) to enter the *paravertebral sympathetic ganglion chain* (also called the *sympathetic trunk*), where most synapse with a postganglionic neuron (Fig. 13.3A). The ganglion chain actually extends from cervical to sacral regions, and some of the thoracolumbar preganglionic neurons extend their axons rostrally or caudally within the chain to reach these cervical and sacral ganglia (see Fig. 13.3A, asterisk). A large complement of postganglionic axons from each of the chain ganglia enter nearby spinal nerves, through a different communicating ramus (*gray ramus*), and travel to the body wall or extremities to control blood vessels, sweat glands, or hair erector muscles (see Fig. 13.3A, #1). Another complement of these postganglionic neurons, mainly from thoracic or cervical chain ganglia, does not enter spinal nerves but forms separate nerves that travel respectively to thoracic viscera (e.g., heart, bronchi) or to organs and glands of the head (e.g., eye, lacrimal gland; see Fig. 13.3A, #2).

Some of the thoracolumbar preganglionic axons simply pass through the sympathetic chain ganglia without synapsing there. These axons form *splanchnic nerves* that synapse with postganglionic

Vet-ebooks.com



• **Fig. 13.2** The site of origin of preganglionic neurons in the central nervous system for both the sympathetic nervous system (*left*) and the parasympathetic nervous system (*right*). Several sites of projection of postganglionic axons are also shown. The *colored rectangles* highlight some of the different paths that the sympathetic system may take to its targets after leaving the central nervous system. Those paths, with corresponding numbers, are presented in more detail in Fig. 13.3. (Modified from Kandel ER, Schwartz JH, Jessell TM. *Principles of Neural Science.* 4th ed. New York: McGraw-Hill; 2000.)

neurons in *prevertebral ganglia* (see Fig. 13.3A, #3), usually named for neighboring blood vessels (e.g., celiac, mesenteric). Postganglionic neurons of the prevertebral ganglia innervate abdominal and pelvic visceral organs. Some of the aforementioned splanchnic nerve fibers bypass the prevertebral ganglia and continue all the way to the *adrenal medulla* (atop the kidney), where they synapse with rudimentary postganglionic neurons that make up the adrenal medullary secretory cells (see Fig. 13.3A, #4). These vestigial postganglionic neurons, called *chromaffin cells,* secrete their transmitter substance directly into the circulating blood. The transmitter substance, acting as a true hormone, is carried by the blood to all tissues of the body.

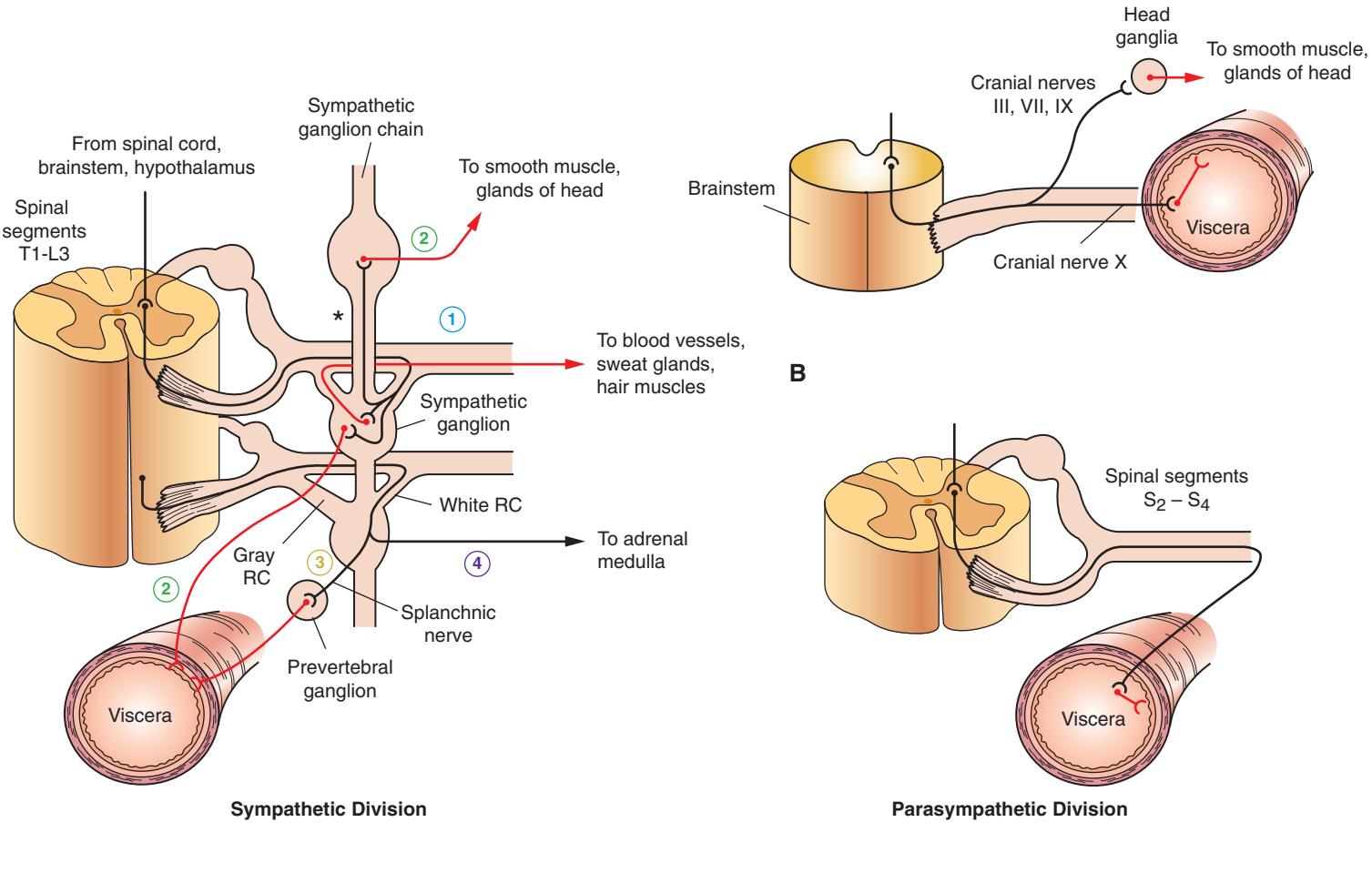
The Parasympathetic Nervous System Arises From the Brainstem and Sacral Spinal Cord

The parasympathetic nervous system generally has long preganglionic and short postganglionic axons. Preganglionic axons of the parasympathetic system leave the CNS by way of cranial nerves III (oculomotor), VII (facial), IX (glossopharyngeal), and X (vagus), and through several sacral spinal nerves. For this reason, it is called the *craniosacral system* (see Fig. 13.2). The parasympathetic preganglionic axons leaving through cranial nerves III, VII, and IX synapse in well-defined ganglia outside the skull (e.g., otic, submandibular; see Fig. 13.3B, *top*). The parasympathetic postganglionic neurons project to smooth muscle and glandular targets in the head (e.g., ciliary muscle, parotid gland).

Preganglionic axons leaving through cranial nerve X travel all the way to the body cavity to synapse in more diffuse parasympathetic ganglia located close to, or within, thoracic and abdominal viscera (see Fig. 13.3B, *bottom*). The short postganglionic neurons control the smooth muscle, cardiac muscle, and glandular cells of these organs.

Parasympathetic preganglionic axons leaving through sacral spinal nerves depart to form pelvic nerves that synapse in diffuse parasympathetic ganglia residing close to, or within, pelvic viscera (e.g. rectum, bladder; see Fig. 13.3C). The short postganglionic neurons control these organs, as well as erectile tissue of the genitals.

Vet-ebooks.com



• **Fig. 13.3** Synaptic organization of preganglionic and postganglionic neurons of the sympathetic (A) and parasympathetic (B, C) divisions of the autonomic nervous system. Anatomical locus of each numbered scheme in part A can be seen in Fig. 13.2, except for #1. See text for further description. *RC,* Ramus communicans (communicating branch). (Modified from Ganong WF. *Review of Medical Physiology.* 13th ed. Norwalk, Conn: Appleton & Lange; 1987.)

Most viscera receive both sympathetic and parasympathetic innervation (see Fig. 13.2). Although the parasympathetic system originates in brainstem and sacral regions, it can provide parasympathetic innervation to organs in the thoracic and lumbar parts of the body, as just noted, by way of the vagus nerve (cranial nerve X). The sympathetic thoracolumbar system can influence organs in cranial and sacral regions by way of preganglionic sympathetic axons that travel to sympathetic postganglionic neurons in cervical and sacral regions of the sympathetic ganglion chain (see Fig. 13.3A, asterisk). Although blood vessels in all parts of the body receive sympathetic innervation, which most commonly produces vasoconstriction, most do not receive parasympathetic innervation (except those in glands and the external genitals).

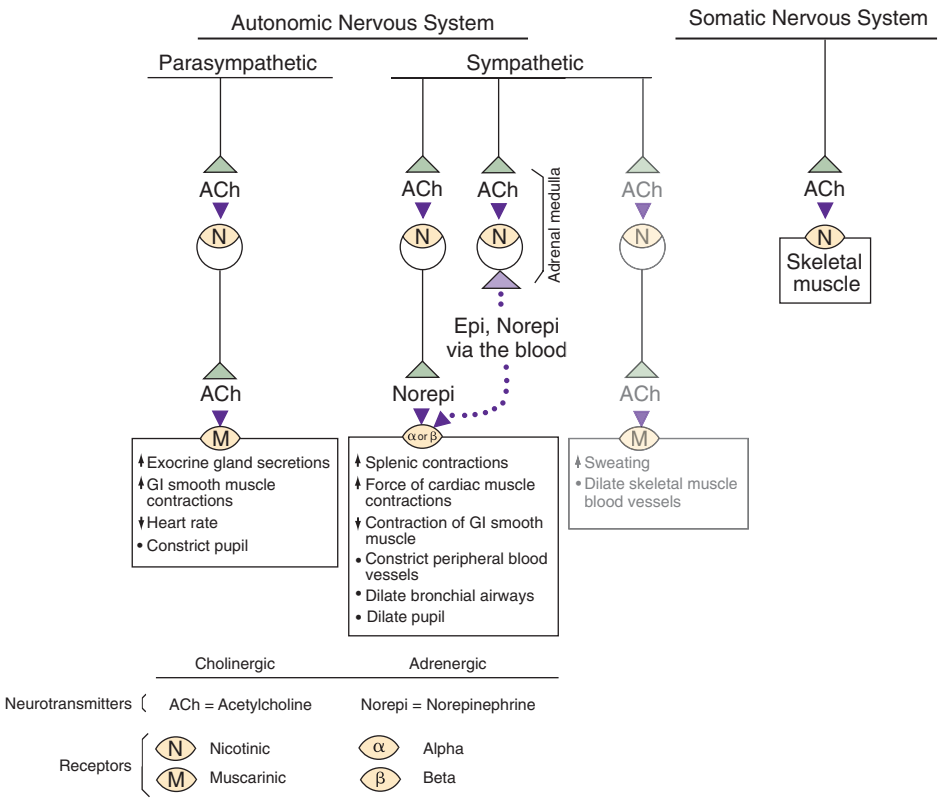
Most Sympathetic and Parasympathetic Neurons Secrete Either Acetylcholine or Norepinephrine as a Neurotransmitter

As described in Chapter 5, *acetylcholine* is the neurotransmitter at the somatic neuromuscular synapse. Acetylcholine is also released by the preganglionic neurons at all autonomic ganglia (Fig. 13.4). Parasympathetic postganglionic neurons release acetylcholine as well onto their target organs. Acetylcholine-releasing synapses are often called *cholinergic.* Most anatomically sympathetic postganglionic neurons secrete *norepinephrine* onto their targets. Norepinephrine-releasing synapses are often called *adrenergic.* However, in several species, anatomically sympathetic postganglionic neurons traveling to sweat glands secrete acetylcholine, as do some of the sympathetic postganglionic neurons to blood vessels in skeletal muscle, where they can produce vasodilation.

In the case of the adrenal medulla, incoming preganglionic axons release acetylcholine, but the neuroendocrine-like, postganglionic chromaffin cells primarily release epinephrine and some norepinephrine into the circulating blood. These chromaffin cells can be considered structural and functional analogs of sympathetic postganglionic neurons.

It is important that, when released, the neurotransmitter not linger in the synaptic cleft. The neurotransmitter must be either destroyed in the cleft or dissipated so that the postsynaptic membrane can recover its resting potential and be ready for the next synaptic transmission. Because some synapses can transmit impulses up to several hundred times per second, neurotransmitter destruction

Vet-ebooks.com



• **Fig. 13.4** Classification of autonomic and somatic motor neurons with regard to their transmitter or mediator released, their postsynaptic receptors, and their general influence on some effector organs. Acetylcholine (*ACh*), released from the presynaptic terminal, can stimulate a muscarinic (*M*) or a nicotinic (*N*) postsynaptic receptor, depending on the particular target. Norepinephrine (*Norepi*) released from presynaptic terminals, or norepinephrine and epinephrine (*Epi*) released together from the adrenal medulla, can stimulate α or β receptors (or both), again depending on the particular target tissue. The faded portion of the figure indicates the exception, rather than the rule, for sympathetic activation. *GI,* gastrointestinal.

must occur quickly. In the case of acetylcholine, *acetylcholinesterase* destroys the transmitter in the cleft. For norepinephrine, reuptake by the presynaptic neuron is the principal way in which its synaptic effect on the postsynaptic membrane is terminated. The hormonal actions of circulating epinephrine and norepinephrine released by the adrenal medulla, however, are primarily terminated by the enzyme *catechol-O-methyltransferase* (COMT), with a lesser contribution of the enzyme *monoamine oxidase* (MAO). These enzymes are widely distributed in the body, with the highest concentrations in the liver and kidney.

Acetylcholine and Norepinephrine Have Different Postsynaptic Receptors

The neurotransmitters secreted by the ANS typically stimulate their target organ by first binding with a postsynaptic receptor. These receptors are proteins in the cell membrane. When the transmitter binds with the postsynaptic receptor, the membrane's permeability to selected ions is often changed, and the postsynaptic membrane potential either increases or decreases, with a resulting change in the probability of action potentials in the postsynaptic cell.

Acetylcholine stimulates two different types of receptors (see Fig. 13.4). *Muscarinic* acetylcholine receptors are G-protein–coupled receptors (GPCRs; see Chapter 1) found on all the target cells stimulated by postganglionic parasympathetic neurons (and by the relatively small number of sympathetic postganglionic neurons that are cholinergic). Faster acting *nicotinic* receptors are ligand-gated ion channels (see Chapter 1) found at all synapses between autonomic preganglionic and postganglionic neurons, and at the somatic neuromuscular junction.

The classification of major types and subtypes of neurotransmitter receptors is usually based upon various combinations of the following: responses to agonist or antagonist drugs, distribution among various tissues and organs, signal transduction mechanism (e.g., G protein-coupled, ligand-gated). Muscarinic receptors were named because they are stimulated by *muscarine,* a toadstool poison. Muscarine does not stimulate nicotinic receptors. *Nicotine* stimulates the nicotinic receptors but not muscarinic receptors. Acetylcholine stimulates both, and different drugs block each receptor. For example, atropine blocks muscarinic receptors, whereas curare blocks nicotinic receptors. Although there are respective subtypes of nicotinic (e.g., N_m, N_n) and muscarinic (e.g., M_1-M_5) acetylcholine receptors, there are few therapeutic drugs that can distinguish among subtype members.

Adrenergic receptors are located at synapses between peripheral target tissues and sympathetic postganglionic neurons that release norepinephrine. However, these receptors can also be stimulated by the release of epinephrine and norepinephrine into the bloodstream from the adrenal medulla. There are two major types of adrenergic receptors, called alpha (α) and beta (β) receptors. The

Vet-ebooks.com

β receptors have been further subdivided into β_1 and β_2 receptors (and β_3, found mainly in fat cells) on the basis of the effect of adrenergic blocking and stimulating drugs. There are two classes of α receptors (α_1 and α_2) that can each be divided into additional subtypes. All adrenergic receptors are GPCRs, and the various subtypes (such as the muscarinic cholinergic receptor subtypes) have differential distributions among various target tissues. The relative representation of receptor subtypes within a target tissue determines the response of that tissue to sympathetic stimulation. There are many clinically useful drugs that can distinguish among the members within adrenergic receptor subtype groups.

Neurotransmitters Other Than Acetylcholine and Norepinephrine Play Some Role in Peripheral Autonomic Function

As more of a rule than an exception, individual neurons are capable of releasing more than one neurotransmitter. Multiple release often depends on how vigorously the neurons are activated by presynaptic stimulation. Therefore preganglionic and postganglionic sympathetic and parasympathetic neurons that release either acetylcholine or norepinephrine are respectively capable of releasing *co-transmitters* under certain circumstances. Most often these co-transmitters are peptides (e.g., vasoactive intestinal peptide, neuropeptide Y, luteinizing hormone-releasing hormone), but some purine (e.g., ATP) and atypical neurotransmitter (e.g., nitric oxide) co-release has been demonstrated as well. These molecules can be collectively referred to as *nonadrenergic, noncholinergic* (*NANC*) *transmitters* of the ANS. Often the postsynaptic response to release of a neurotransmitter is modified by the release of a co-transmitter from the same neuron. For example, acetylcholine released from parasympathetic postganglionic neurons can activate salivary glands, but co-release of vasoactive intestinal peptide from the same neurons can affect blood vessel diameter in the target region as well.

Acetylcholine and norepinephrine can also be found in the enteric nervous system: acetylcholine is released by excitatory enteric neurons of the gut (see Chapter 27), and postganglionic sympathetic neurons can release norepinephrine into enteric neuronal plexuses to induce inhibition. Like the sympathetic/parasympathetic systems, various enteric neurons also employ vasoactive intestinal peptide, neuropeptide Y, ATP, and nitric oxide. However, the variety of neurotransmitters, other than acetylcholine and norepinephrine, employed by neurons of the enteric nervous system is much more extensive than that found among the sympathetic and parasympathetic systems.

There Are General Differences in Sympathetic and Parasympathetic Function

Although the sympathetic and parasympathetic systems are both important for homeostasis, there are some important general differences in their function.

In physical and some emotional stress, the sympathetic system is capable of a massive, coordinated output with widespread effects on tissues and organs of the body. This causes an increase in heart rate and blood pressure, dilation of the pupil of the eye, an elevation in levels of blood glucose and free fatty acids, and an increased state of arousal. These widespread effects mobilize the body's resources for extra effort in responding to an emergency. Therefore the sympathetic system is sometimes referred to as the *fight or flight* system. The effect of sympathetic discharge not only is widespread but can also last longer than effects of parasympathetic discharge because of the prolonged circulation of epinephrine and norepinephrine. Indeed, the adrenal medulla's secretion of epinephrine and norepinephrine into the circulating blood provides prolonged adrenergic stimulation to the entire body, even to some tissues that do not have direct sympathetic postganglionic stimulation.

Under less stressful conditions, the sympathetic system plays an important role in homeostasis but with less universal control within the body. For example, sympathetic control of the skin for thermoregulation, or of the dilator smooth muscle in the iris for enlargement of the pupil in low ambient light, can respectively occur without extensive activation of other organs.

The parasympathetic system is characterized by a greater degree of independent control of tissues and organs, as well as a more precise control within a given tissue or organ, compared with the sympathetic system. In addition, unlike the sympathetic system that innervates virtually all parts of the body, the parasympathetic system does not innervate structures of the body wall and extremities. The parasympathetic system is generally concerned with the restorative aspects of daily living. For example, parasympathetic stimulation assists digestion and absorption of food by increasing gastric secretion, increasing intestinal motility, and relaxing the pyloric sphincter. For this reason, the parasympathetic nervous system is sometimes called the *anabolic* or *restorative* nervous system, as well as the *rest and digest* system.

Many organs of the body have both sympathetic and parasympathetic innervation, each with a reciprocal effect. For example, sympathetic stimulation increases heart rate, whereas parasympathetic stimulation decreases heart rate. Sympathetic stimulation enlarges pupillary diameter, whereas parasympathetic stimulation causes pupillary constriction. These sympathetic and parasympathetic systems work together, along with the enteric system, to exquisitely keep the body's internal environment stable.

Table 13.1 gives a more complete listing of the responses of various organs to adrenergic and cholinergic stimulation by the peripheral ANS.

Visceral Afferent (Sensory) Neurons Play an Important Role in Autonomic Nervous System Function

Autonomic reflexes regulate many of the body's visceral functions. As with reflex arcs in the somatic nervous system (see Chapter 7), autonomic reflex arcs also include a sensory side to the arc, including a visceral receptor; a sensory neuron, often called a *visceral afferent* neuron; and one or more synapses in the CNS. The ANS is commonly defined as the peripheral motor preganglionic and postganglionic neurons. Visceral afferent neurons are often not included in this definition, but because they are essential parts of the autonomic reflex arc, they will be described briefly below.

The peripheral portion of a visceral afferent neuron's axon travels toward the CNS along splanchnic, cranial, and pelvic nerves that carry the sympathetic or parasympathetic visceral efferents toward their peripheral targets. Like the somatic afferent neurons carrying touch information from the skin, the visceral afferent neurons have their cell bodies located in dorsal root or cranial nerve ganglia. The centrally directed portion of the axon synapses in the dorsal horn of the spinal cord or in a cranial nerve nucleus of the brain.

Vet-ebooks.com

TABLE 13.1 Responses of Effector Organs to Autonomic Nerve Impulses and Circulating Catecholamines

Effector Organ	Cholinergic Impulses: Response	Noradrenergic Impulses	
		Receptor	Response
Eye			
Radial muscle of iris	—	α_1	Contraction (mydriasis)
Sphincter muscle of iris	Contraction (miosis)	—	—
Ciliary muscle	Contraction for near vision	β_2	Relaxation for far vision
Heart			
Sinoatrial node	Decrease in heart rate	$\beta_1 > \beta_2$	Increase in heart rate
Atria	Decrease in contractility and shortened AP duration	$\beta_1 > \beta_2$	Increase in contractility and conduction velocity
Atrioventricular (AV) node	Decrease in conduction velocity; AV block	$\beta_1 > \beta_2$	Increase in conduction velocity
His-Purkinje conduction system	Little effect	$\beta_1 > \beta_2$	Increase in conduction velocity
Ventricles	Little effect	$\beta_1 > \beta_2$	Increase in contractility and conduction velocity
Arterioles			
Skeletal muscle, pulmonary, abdominal viscera	Dilation (sometimes skeletal muscle)	α_1	Constriction
		β_2	Dilation
Coronary, renal	—	α_1, α_2	Constriction
		β_2 (also β_1 renal)	Dilation
Skin and mucosal, cerebral, salivary gland	—	α_1, α_2 (cerebral α_1 only)	Constriction
Systemic veins	—	α_1, α_2	Constriction
		β_2	Dilation
Lung			
Bronchial muscle	Contraction	β_2	Relaxation
Bronchial glands	Stimulation	α_1	Decrease in secretion
		β_2	Increase in secretion
Stomach (Monogastric)			
Motility and tone	Increase	α_1, α_2, β_1, β_2	Decrease (usually)
Sphincters	Relaxation (usually)	α_1	Contraction (usually)
Secretion	Stimulation	α_2	Inhibition
Intestine			
Motility and tone	Increase	α_1, α_2, β_1, β_2	Decrease
Sphincters	Relaxation (usually)	α_1	Contraction
Secretion	Stimulation	α_2	Inhibition
Gallbladder and Ducts			
Gallbladder and ducts	Contraction	β_2	Relaxation
Urinary Bladder			
Detrusor	Contraction	β_2	Relaxation
Trigone and sphincter	Relaxation	α_1	Contraction
Ureter			
Motility and tone	Increase (?)	α_1	Increase

Vet-ebooks.com

TABLE 13.1 Responses of Effector Organs to Autonomic Nerve Impulses and Circulating Catecholamines—cont'd

Effector Organ	Cholinergic Impulses: Response	Noradrenergic Impulses: Receptor	Noradrenergic Impulses: Response
Reproductive System			
Uterus	Variable[a]	α_1, β_2	Variable
Male sex organs	Erection	α_1	Ejaculation
Skin			
Pilomotor muscles	—	α_1	Contraction
Sweat glands	Generalized secretion	α_1	Localized secretion[b]
Upper Abdominal Structures			
Spleen capsule	—	α_1	Contraction
		β_2	Relaxation
Adrenal medulla	Secretion of epinephrine and norepinephrine	—	—
Liver	—	α_1, β_2	Glycogenolysis and gluconeogenesis
Kidney	—	α_1	Decreased renin secretion
		β_1	Increased renin secretion
Pancreas			
Acini	Increased secretion	α	Decreased secretion
Islets	—	α_2	Decreased insulin and glucagon secretion
		β_2	Increased insulin and glucagon secretion
Other Glands			
Salivary glands	K^+ and H_2O secretion	α	K^+ and H_2O secretion
Lacrimal glands	Secretion	α	Secretion
Pineal gland	—	β	Increase in melatonin synthesis and secretion

[a]Depends on stage of estrous cycle, amount of circulating estrogen and progesterone, pregnancy, and other factors.
[b]On palms of human hands and in some other locations (adrenergic sweating).
Modified from Westfall TC, Westfall DP. Neurotransmission: the autonomic and somatic motor nervous systems. In: Brunton L, Hilal-Dandan R, Knollman B. *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. 13th ed. New York: McGraw-Hill; 2018.

Generally, visceral afferent neurons carrying nociceptive (pain-inducing) information travel in sympathetically characterized nerves (e.g., splanchnic), whereas those carrying non-nociceptive information run within parasympathetically characterized nerves (e.g., vagus, pelvic). Nociceptive stimuli from the viscera can result from strong dilation or contraction of an organ but are usually chemical in nature, resulting from inflammation or ischemia (restricted blood supply) of an organ. Normal stretch or movement of an organ or blood vessel, or changes in oxygen or carbon dioxide concentration in the blood, are examples of non-nociceptive visceral stimuli.

Non-nociceptive sensory signals from the viscera usually do not reach consciousness, but those that do tend to be diffuse and difficult to localize. Although nociceptive visceral signals often reach consciousness and are initially difficult to localize, the organ source of the pain is often "referred" to (feels like it's coming from) regions of the skin in the general vicinity of the organ. This *referred pain* is thought to result from the convergence of somatic (e.g., skin) afferents and visceral afferents within the dorsal horn of the same spinal cord segment. In human medicine, the locus of the skin pain can be a reliable clue as to the location of the affected organ.

The Autonomic Nervous System Participates in Many Homeostatic Reflexes

Autonomic reflexes are extremely common and are described in detail for each body system in later chapters. A few are described briefly here as examples.

Control of Blood Pressure

Among the body's major priorities is keeping a sufficient blood flow to the brain. Stretch receptors in the internal carotid artery and the aorta detect systemic blood pressure. When these receptors detect a drop in pressure, an increase in activity of sympathetic adrenergic neurons produces peripheral vasoconstriction and increased vascular resistance to increase blood pressure and restore sufficient blood flow to the brain. If blood pressure rises above normal limits in animals, sympathetic adrenergic vasoconstrictor nerves are inhibited and blood pressure falls back to within normal limits.

Vet-ebooks.com

Pupillary Light Reflex

When a flashlight is shined into an animal's eye, light stimulates photoreceptors in the retina (see Chapter 14). Sensory action potentials are then transmitted to the brainstem along the optic nerve, where, through several interneurons, parasympathetic cholinergic neurons stimulate the constrictor smooth muscle of the iris. This causes the pupillary diameter to become smaller.

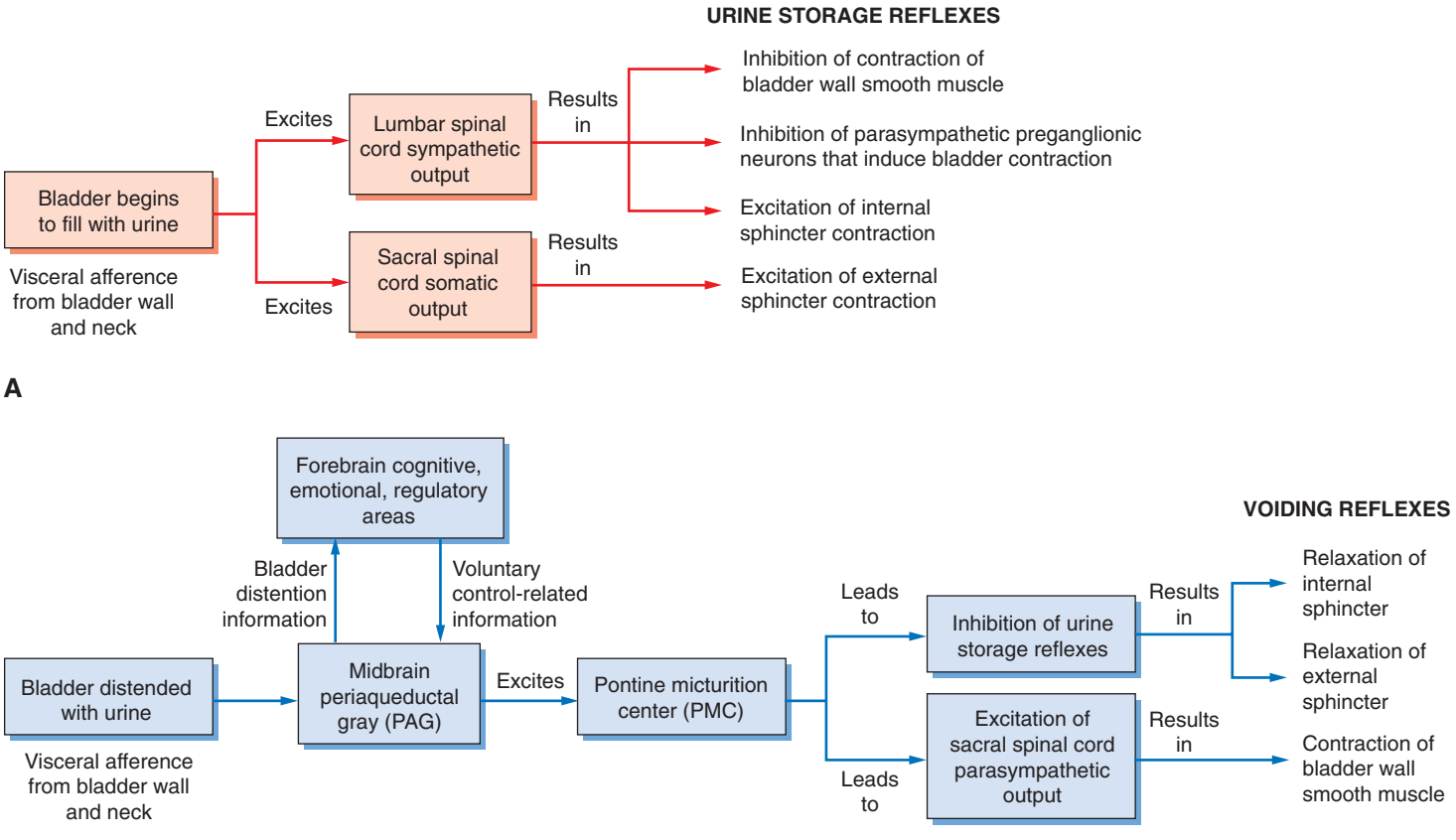
Micturition (Urination)

The normal regulation of *micturition* actually represents a complex interplay of autonomic reflex activity and skeletal muscle control, upon which an element of voluntary regulation is superimposed. Autonomic efferent control of bladder emptying is principally achieved by pelvic parasympathetic output that causes contraction of the bladder wall (detrusor muscle). The ability to store urine is facilitated by sympathetic efferent control originating in the lumbar splanchnic nerves (Fig. 13.5A). This sympathetic output will ultimately inhibit contraction of the detrusor muscle directly, as well as indirectly by inhibiting parasympathetic postganglionic neurons that stimulate its contraction. This sympathetic output also facilitates urine storage by contracting the smooth muscle internal sphincter located at the neck of the bladder. Superimposed upon the sympathetic facilitation of urine storage is the contraction of a skeletal muscle external sphincter (urethralis muscle) controlled by somatic motor neurons of the sacral spinal cord, which is also amenable to conscious voluntary regulation. The sympathetic facilitation of urine storage, together with contraction of the striated external sphincter, can be collectively referred to as the *urine storage reflexes,* which are primarily organized at the level of the spinal cord.

Visceral afferents to the lumbosacral spinal cord from the bladder wall and the urethra provide distention information about the extent of bladder filling. As the bladder begins to fill with urine, this visceral afferent input to the spinal cord activates the urine storage reflexes described previously (see Fig. 13.5A).

The distention information regarding bladder filling is also sent from the spinal cord to a region of the midbrain called the *periaqueductal gray* (PAG; see Fig. 13.5B). When bladder filling reaches a critical level, the PAG sends a signal to a region of the pons called the *pontine micturition center* (PMC), which, in turn, has projections to the spinal cord that coordinate the components of the *voiding reflexes.* The voiding reflexes include (1) inhibition of the sympathetically mediated storage reflexes, resulting in relaxation of the internal and external sphincters, and (2) excitation of the parasympathetic innervation of the bladder, resulting in bladder contraction.

Of course, animals do not immediately void urine when the bladder is full because survival or social conditions are not always favorable. There is an element of voluntary control over whether the brainstem PMC is activated to trigger the spinal cord voiding reflexes. The distention information sent to the PAG is passed from there to forebrain regions that can consciously perceive bladder



• **Fig. 13.5** Organization of micturition. (A) Neural events involved in the urine storage reflexes as the bladder begins to fill with urine. (B) Neural events involved in the voiding reflexes when the bladder becomes full and distended.

Vet-ebooks.com

fullness and to forebrain regions involved in cognitive (e.g., prefrontal cortex), regulatory (e.g., hypothalamus), and emotional (e.g., amygdala) processes (see Fig. 13.5B). These areas will communicate voluntary control-related information back to the PAG regarding the safety and social context of the situation. The PAG integrates this information, along with distention signals, to determine whether the PMC should be activated to initiate the voiding reflexes that produce emptying of the bladder.

Gastric secretion of digestive fluids in anticipation of food, and emptying of the rectum in response to filling, are but a few of the many other autonomic reflexes described in more detail throughout this book.

Preganglionic Neurons Are Influenced by Many Regions of the Brain

Much as the lower motor neuron of the somatic system is influenced by the upper motor neuron (see Chapter 9), the preganglionic autonomic neuron is influenced by CNS axons descending from the brainstem and forebrain. However, it should be noted that many of the CNS structures that affect autonomic function cannot be easily assigned a sympathetic or parasympathetic role, and many are involved in nonautonomic control in the body as well.

Many brainstem structures are known to influence autonomic preganglionic neurons to control particular visceral functions; we have just seen an example with respect to micturition (e.g., PMC). Another example is the vasopressor center of the medulla that increases peripheral vascular resistance and cardiac output. Often these "centers" are networked clusters of neurons as opposed to a discrete nucleus. Some brainstem structures are actually composed of the cell bodies of preganglionic autonomic neurons (e.g., motor nucleus of the vagus).

Most of the brainstem regions that influence autonomic preganglionic neurons receive input from the *hypothalamus* of the diencephalon, a structure critical for homeostasis. The hypothalamus coordinates the activity of these brainstem regions as an important means of regulating visceral function for purposes of homeostasis. The hypothalamus exerts its control over these brainstem regions using the guidance of cortical and other telencephalic inputs related to cognitive and emotional processes, and with the guidance of both somatic and visceral sensory information. In its critical role in homeostasis, the hypothalamus not only regulates autonomic function, it also regulates endocrine function, as well as certain somatic motor actions. By coordinating these three roles, the hypothalamus can produce some complex behaviors related to homeostasis, such as feeding.

Much of the visceral afferent information used by brain structures that influence autonomic preganglionic neurons synapses within the *solitary nucleus* of the medulla. The solitary nucleus then distributes this information, either directly or indirectly, to brainstem or forebrain areas, noted previously, that will influence the autonomic preganglionic neurons. Some of the solitary nucleus output synapses directly on preganglionic neurons of the brainstem, producing some of the simplest autonomic reflexes.

Although the organization of brain structures that affect the function of autonomic preganglionic neurons has been presented in a somewhat straightforward, hierarchal, rostral-to-caudal fashion, the relationship of these brain structures is really more complicated. For example, the hypothalamus can project directly to preganglionic autonomic neurons of the brainstem and spinal cord, there are telencephalic projections to autonomic-related brainstem structures that bypass the hypothalamus, and there are interconnections among different cortical regions involved in autonomic function. Therefore the central control of autonomic function represents more of a complex central autonomic network than a simple hierarchy of control. This complex system of upper motor neurons within the CNS helps coordinate autonomic reflexes and directly influences action potential frequency within preganglionic neurons. As more is learned about these central systems controlling the ANS, their specific role in such conditions as hypertension and various gastrointestinal diseases may become clearer.

There Are Interactions Between the Autonomic Nervous System and the Immune System

An area of expanding investigation is the relationship among central ANS-related structures (particularly the hypothalamus), peripheral ANS structures, and immune function. For example, some cytokines released by immune cells can bind to visceral afferents of the vagus nerve that send information to the aforementioned solitary nucleus of the medulla, which in turn sends axons to parts of the hypothalamus. Hypothalamic-regulated secretion of glucocorticoids from the adrenal cortex is known to modulate neuroinflammation. Furthermore, postganglionic sympathetic neurons have been found to innervate several immune related structures such as lymph nodes, spleen, and thymus, and activation of vagus nerve parasympathetic axons has been shown to modulate release of cytokines from immune cells. These peripheral ANS neurons are amenable to control by the hypothalamus. Interestingly, in light of the fact that sympathetic postganglionic axons release norepinephrine, and the adrenal medulla can release epinephrine and norepinephrine, it has also been shown that lymphocytes possess β-adrenergic receptors. Further research on ANS-immune interactions will have significant implications for both animal and human health.

CLINICAL CORRELATIONS

Horner's Syndrome

History

A 7-year-old male golden retriever is brought to your clinic for examination. The owner states that during the past 3 weeks, the dog has become progressively weaker in his left front leg and now cannot bear weight on that limb. The owner has also noticed that the dog's left upper eyelid seems to be droopy.

Clinical Examination

Physical examination abnormalities are limited to the nervous system. The dog is bright, alert, and responsive. Cranial nerve reflexes are within normal limits. The dog cannot bear weight on the left front leg, and the leg muscles are atrophied. No segmental reflexes (e.g., toe-pinch withdrawal) or intersegmental responses (e.g., proprioceptive placing) can be elicited in the left front leg. The left upper eyelid droops lower than the right upper lid, and

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

the left pupil is smaller than the right pupil. The left nictitating membrane (third eyelid) is prolapsed (slipped out of place) over part of the cornea, and the left eye seems more sunken into the orbit than the right eye.

Comment

This dog has a lesion of the left brachial plexus, probably a neoplasm. It has caused a lower motor neuron syndrome to the left front leg with atrophy, paralysis, and loss of reflexes. The tumor has damaged the preganglionic neurons of the left sympathetic nervous system as they leave the first two thoracic segments on their way toward the eye. Loss of the sympathetic innervation to the region of the eye causes the small pupil (miosis), drooping of the upper eyelid (ptosis), a sunken appearance in the eye (enophthalmos), and prolapse of the nictitating membrane. This constellation of clinical signs is called *Horner's syndrome.*

Sympathetic preganglionic neurons pass through the brachial plexus (where they were damaged in this dog) and ascend in the vagosympathetic trunk to synapse with the postganglionic neurons in the cranial cervical ganglia. The postganglionic cell axons then go to the region of the eye, where they innervate the dilator smooth muscle cells of the iris. When paralyzed, the constrictor fibers of the iris are unopposed, resulting in miosis. The sympathetic nervous system also innervates several smooth muscle fibers that lift the upper eyelid and help position the nictitating membrane and the eye within the socket. Because the preganglionic fibers are relatively exposed in the neck, they are often damaged in this condition.

Horner's syndrome can also result from damage to either the postganglionic neurons or the neurons that descend from the hypothalamus to the rostral thoracic cord to control the preganglionic neurons.

Treatment

Treatment involves removing the cause of sympathetic nerve damage, with variable effect upon neural symptoms.

Colic From Administration of Atropine to Treat Corneal Ulcer

History

A client calls you and says that the 14-year-old mare they are treating for a corneal ulcer is now agitated, looking at her sides, and trying to roll. The mare is not interested in eating or drinking. She was fine this morning when they treated her, but this afternoon she is looking agitated.

Clinical Examination

The mare has an increased heart rate, respiratory rate, and normal temperature. The eye that has the corneal ulcer is dilated (mydriatic). Her gastrointestinal borborygmi (growling sounds) are decreased in all quadrants. She is given an antiinflammatory medication (banamine, flunixin meglumine). Within a few minutes, she is not as agitated, and she is standing more comfortably. When a nasogastric tube is passed, a moderate amount of gas is present. On palpation per rectum, there is some increased gas present in her cecum, but no other abnormal findings. The mare appears to be more comfortable. You ask the client if there have been any changes in treatments, medications, and/or management. When they show you the medications for the corneal ulcer, you realize they had accidentally switched the frequency of administration: they were giving the triple antibiotic 2 times/day and the atropine 4 times/day, instead of the antibiotic 4 times/day and atropine 2 times/day, and they had done this for the last 3 to 4 days.

Comment

Although there are many different causes of colic, it is possible that the colic is due to the inadvertent overadministration of atropine. Atropine can be absorbed systemically from the ophthalmic administration. Atropine blocks the postsynaptic effects of acetylcholine at the level of the muscarinic receptors and is parasympatholytic (interrupting parasympathetic function), thereby reducing the antagonism of sympathetic effects at the end organs. The resulting relative increase in sympathetic tone can cause an increased heart rate, decreased gastrointestinal (GI) motility, and increased sphincter tone, as well as decreased GI secretions and urine retention.

Treatment

The horse is treated for colic with banamine, which inhibits cyclooxygenase, and is both an antiinflammatory and analgesic. Additionally, the horse is given xyalzine, which is an α_2 adrenergic agonist, with or without butorphanol. Xyalzine will decrease release of neurotransmitters from the neuron by binding to presynaptic α_2 receptors (negative feedback receptors). This decreases sympathetic outflow, including possibly causing bradycardia (decreased heart rate). Butorphanol, which is an opiate that is a κ (kappa) receptor agonist and weak μ (mu) receptor antagonist, is sometimes used as well in combination with xyalzine. In addition to antiinflammatory and possibly analgesics, the horse can also be given mineral oil and/or water via the nasogastric tube to correct possible dehydration. In some cases, if food bolus impaction or ingestion of an excessive or irritating amount of food has occurred, mineral oil may also be given.

Practice Questions

1. Choose the *incorrect* statement below:
- **a.** A ganglion is a collection of nerve cell bodies outside the CNS.
- **b.** Acetylcholine is a chemical transmitter at the parasympathetic postganglionic–to–target organ synapse.
- **c.** Sympathetic postganglionic neurons are usually longer than those of the parasympathetic system.
- **d.** The adrenal medulla secretes mostly norepinephrine and relatively little epinephrine.
- **e.** Muscarinic acetylcholine receptors are found on peripheral targets of parasympathetic postganglionic neurons.

2. A chemical neurotransmitter between preganglionic and postganglionic neurons of the parasympathetic component of the autonomic nervous system is:
- **a.** Norepinephrine
- **b.** Acetylcholine
- **c.** Epinephrine
- **d.** Serotonin
- **e.** γ-Aminobutyric acid

3. A neurotransmitter most often found at the synapse between sympathetic postganglionic neurons and their targets is:
- **a.** Norepinephrine
- **b.** Epinephrine
- **c.** Acetylcholine
- **d.** Dopamine
- **e.** γ-Aminobutyric acid

4. Which of the following is *true* regarding sympathetic preganglionic neurons?
- **a.** Their cell bodies are located in thoracic and lumbar regions of the spinal cord.
- **b.** Their axons synapse within the sympathetic ganglion chain.
- **c.** Their axons form splanchnic nerves.
- **d.** Their axons travel in the ventral root.
- **e.** All the above are true.

Vet-ebooks.com

5. Horner's syndrome is caused by the loss of:
- **a.** Sympathetic innervation to the eye.
- **b.** Parasympathetic postganglionic innervation to the eye.
- **c.** Peripheral muscarinic receptors.
- **d.** Vagus nerve fibers.
- **e.** Smooth muscle of the iris.

6. Which of the following is *true* regarding interactions between the autonomic nervous system and the immune system?
- **a.** Lymphocytes can possess β-adrenergic receptors.
- **b.** The hypothalamus can indirectly modulate neuroinflammation.
- **c.** Postganglionic sympathetic neurons innervate lymph nodes.
- **d.** Activating parasympathetic axons of the vagus nerve can modulate cytokine release from immune cells.
- **e.** All the above are true.

Bibliography

Benarroch EE. Central autonomic control. In: Robertson D, Biaggioni I, Burnstock G, et al, eds. *Primer on the Autonomic Nervous System*. 3rd ed. London, UK: Academic Press; 2012.

Boron WF. *Boulpaep EL: Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Evans HE, de Lahunta A. *Miller's Anatomy of the Dog*. 4th ed. Philadelphia: Saunders; 2013.

Fowler CJ, Griffiths D, de Groat WC. The neural control of micturition. *Nat Rev Neurosci*. 2008;9(6):453–466.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Matsukawa K, Shirai M, Murata J, et al. Sympathetic cholinergic vasodilation of skeletal muscle small arteries. *Jpn J Pharmacol*. 2002;88(1):14–18.

Nichols JG, Martin AR, Fuchs PA, Brown DA. *From Neuron to Brain*. 5th ed. Sunderland, Mass: Sinauer; 2011.

Papich M. *Saunders Handbook of Veterinary Drugs*. 4th ed. St Louis: Saunders; 2016.

Plumb D. *Veterinary Drug Handbook*. 8th ed. Hoboken, New Jersey: Wiley Blackwell; 2015.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Strain GM. Autonomic nervous system. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Westfall TC, Westfall DP. Neurotransmission: the autonomic and somatic motor nervous systems. In: Brunton L, Hilal-Dandan R, Knollman B, eds. *Goodman and Gilman's the Pharmacological Basis of Therapeutics*. 13th ed. New York: McGraw-Hill; 2018.

Vet-ebooks.com

14

The Visual System

BRADLEY G. KLEIN

KEY POINTS

1. The eye's anatomy is adapted to the eye's role as a visual receptor organ.
2. Through the process of accommodation, the lens changes shape to focus images from various distances onto the retina.
3. The vertebrate retina consists of five major cell types.
4. In some species, the fovea minimizes distortion of light compared with other areas of the retina.
5. Tissue layers behind the retina principally absorb light, or can reflect light, depending on the species' habits.
6. Photoreception and transduction of light occur in the rods and cones.
7. Visual image processing in the retina begins as the response of the photoreceptor to light is synaptically transmitted to the ganglion cells by the bipolar cells.
8. The electroretinogram records the electrical response of the retina to a flashing light.
9. Retinal ganglion cells transmit neural signals to the visual cortex by way of the lateral geniculate nucleus.
10. The diameter of the pupil is controlled by the autonomic nervous system.
11. The retina, optic nerve, and autonomic nerve supply controlling the pupil can be tested with a flashlight.
12. Aqueous humor determines intraocular pressure.

The eyes are complex sense organs that are basically an extension of the brain. They evolved from primitive light-sensing spots on the surface of invertebrates and in some species have developed many remarkable variations, providing special advantages in various ecological niches. Each eye has a layer of receptors, a lens system for focusing an image on these receptors, and a system of axons for transmitting action potentials to the brain. This chapter describes how these and other components of the eye work.

The Eye's Anatomy Is Adapted to the Eye's Role as a Visual Receptor Organ

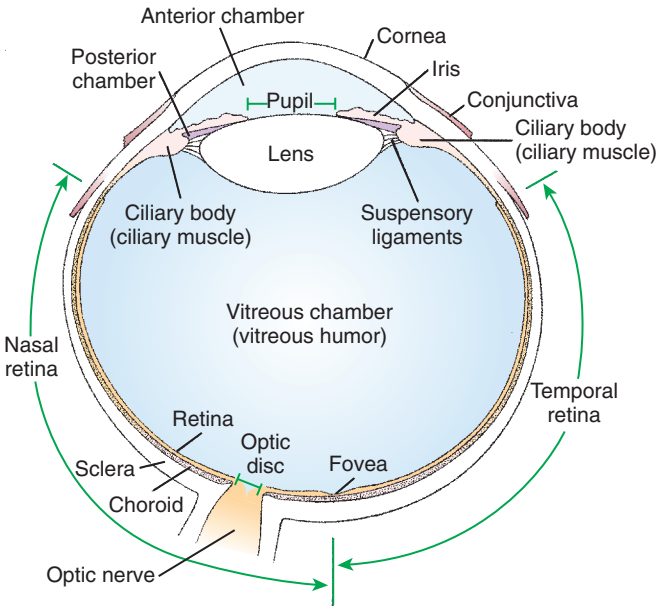
Fig. 14.1 shows the anatomy of the normal eye in the horizontal plane. The white, outer protective layer encasing most of the eyeball is called the *sclera.* It is modified anteriorly into a clear region called the *cornea* that consists of a specialized arrangement of collagen fibrils with an overlying, stratified, squamous epithelial layer. In the posterior two-thirds of the eye, the inner surface of the sclera is lined with a vascular and pigmented layer called the *choroid.* Interior to the choroid is the *retina,* the layer containing the photoreceptors.

As light passes through the transparent cornea, it undergoes some bending that will ultimately help to focus it on the retina. After passing through the cornea, the light enters a compartment called the *anterior chamber* (see Fig. 14.1). The anterior chamber and the *posterior chamber* are filled with a clear, water-like fluid called *aqueous humor* that supplies important nutrients to the cornea (as well as the lens). Separating the anterior and posterior chambers is a diaphragm of varying size called the *iris.* The iris is a pigmented structure containing dilator and constrictor smooth muscle fibers arranged to vary the diameter of the *pupil,* the hole in the iris through which light passes on its way to the retina. The size of the pupil regulates the amount of light entering the eye. Behind the iris is the lens. The lens is suspended in the eye by *suspensory ligaments* (known as *zonular fibers*), which attach to the lens and to the *ciliary body,* a muscular structure near the base of the iris. As discussed later, the lens provides variable focusing power, in contrast to the fixed cornea.

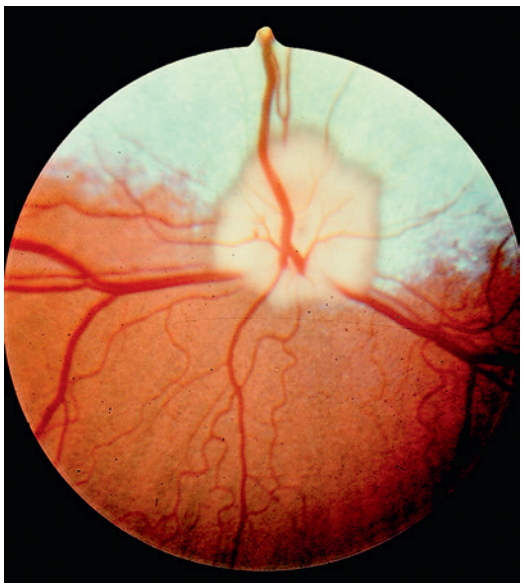
Behind the lens is a chamber filled with a gelatinous fluid called the *vitreous humor.* Because of the viscosity of this fluid, the pressure generated by the aqueous humor, and the fairly inelastic nature of the sclera and cornea, the globe of the eye is basically spherical. The vitreous humor also contains phagocytic cells that can clear ocular debris capable of obstructing the light path. Behind the vitreous humor is the neural retinal layer where light is transduced into the electrical activity of neurons. The retina is interrupted at a point where axons of the retina's ganglion cell layer, which travel across the inner surface of the retina, leave the eye on their way to the brain. This point, the *optic disc,* is a recognizable structure when the eye is examined with an ophthalmoscope (Fig. 14.2). The interruption of the light-processing retina at the optic disc produces a *blind spot,* another name for the optic disc. The retinal ganglion cell axons leaving the eye at the optic disc give rise to the *optic nerve* (cranial nerve 2), a cranial nerve so rich in axons that there are more axons in both optic nerves than in all the dorsal roots of the spinal cord.

Also visible through the ophthalmoscope, on the surface of the retina, are the *retinal blood vessels* (see Fig. 14.2). This network of arteries and veins enters the retina at the optic disc and provides much of the nutrition to the retina. Vessels of the choroid, which enter it after piercing the sclera near the optic disc, provide the remaining nutrition to the retina. Examination of retinal vessels often provides valuable clues to abnormalities elsewhere in the cardiovascular system.

Vet-ebooks.com



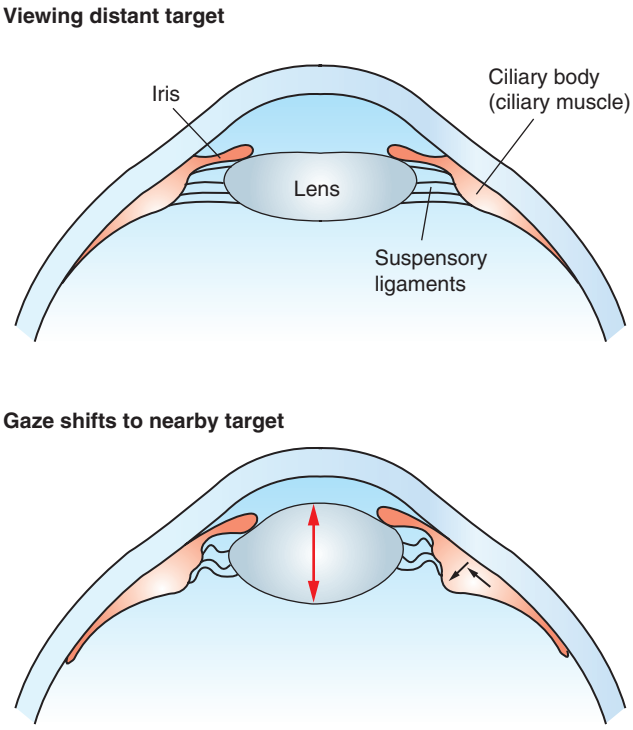
• **Fig. 14.1** Schematic diagram of a horizontal section through the right eye as viewed from above. (Redrawn from Walls GL. The vertebrate eye and its adaptive radiation. *Cranbrook Inst Sci.* 1942;19:1–785.)



• **Fig. 14.2** Ocular fundus (inner posterior surface) of the eye of a medium sized dog, as viewed with an ophthalmoscope, showing the optic disc, retinal blood vessels, and tapetum. The optic disc appears as a light, circular region with a darker border, located just off the center of the image. The reflective tapetum appears as the whiter-colored area surrounding the optic disc constituting the upper third of the image. (From De Lahunta A, Glass E. *Veterinary Anatomy and Clinical Neurology.* 3rd ed. Philadelphia: Saunders; 2009.)

The *lacrimal gland,* located near the lateral canthus (where upper and lower lids meet) of the eye, produces tears in response to parasympathetic nerve stimulation. Tears then flow over the cornea and are drained into the nose by the nasolacrimal duct. A regular flow of tears across the cornea is essential to the health of the cornea.

The eye is directed toward environmental sources of light by the six striated *extraocular muscles* that originate within the orbit



• **Fig. 14.3** The process of accommodation as seen in a horizontal section through the eye. As gaze shifts from a distant to a nearby target the ciliary muscle contracts, moving forward and inward (*small black arrows*), releasing tension on the suspensory ligaments. This reduced tension, and the inherent elasticity of the lens, allows the lens to widen anteroposteriorly (*red arrow*) toward its more natural, unstretched, spherical configuration.

and attach to the sclera. The muscles can rotate the eye around the dorsoventral, mediolateral, and anteroposterior axes. The third (oculomotor), fourth (trochlear), and sixth (abducens) cranial nerves contribute to this movement.

Through the Process of Accommodation, the Lens Changes Shape to Focus Images From Various Distances Onto the Retina

When a camera focuses the images of objects at various distances from the sensor, the distance between the lens and the sensor is changed. The eye, however, focuses images by changing the shape of the lens, not by changing the distance between the lens and the retina.

Fig. 14.3 shows the process of *accommodation,* whereby the lens adds extra focusing power by changing its shape. The lens of the eye is made up of an elastic *lens capsule* containing laminae of lens fibers arranged like the layers of an onion. Given this anatomy, if the eye's lens were taken out of the eye, the lens would assume a spherical shape, principally due to the elasticity of its capsule. When suspended in the relaxed eye, however, the suspensory ligaments pull on the equator of the lens, causing it to flatten in its anterior-posterior dimension. This flattened, less convex lens causes less refraction (bending) of light rays and allows the focus onto the retina of objects more than 20 feet away. To focus the image of objects closer to the eye, however, the lens must assume a more spherical, convex shape. This is accomplished by contraction of the ciliary muscle of the ciliary body. This contraction causes the ciliary muscle to move anterocentripetally (forward and inward),

Vet-ebooks.com

which results in a relaxation of tension on the suspensory ligaments. The result, given the inherent elasticity of the lens capsule, is a more spherical lens with more light refraction that focuses onto the retina the image of nearer objects. The more the ciliary muscle contracts, the more spherical the lens becomes.

In humans, as the lens ages, it becomes less elastic and tends to become less spherical, even when the ciliary muscles contract. This condition is known as *presbyopia.* Many people older than 40 years of age need reading glasses to help their less elastic lens focus on objects nearby. A related condition in dogs and cats, called *nuclear sclerosis,* may begin at around 7 years of age, but it does not appear to produce as significant a visual deficit as in human presbyopia.

The lens should be clear and free of opacities. In *cataracts,* however, the lens becomes more opaque, causing random refraction of light and blurring vision, often leading eventually to blindness.

The Vertebrate Retina Consists of Five Major Cell Types

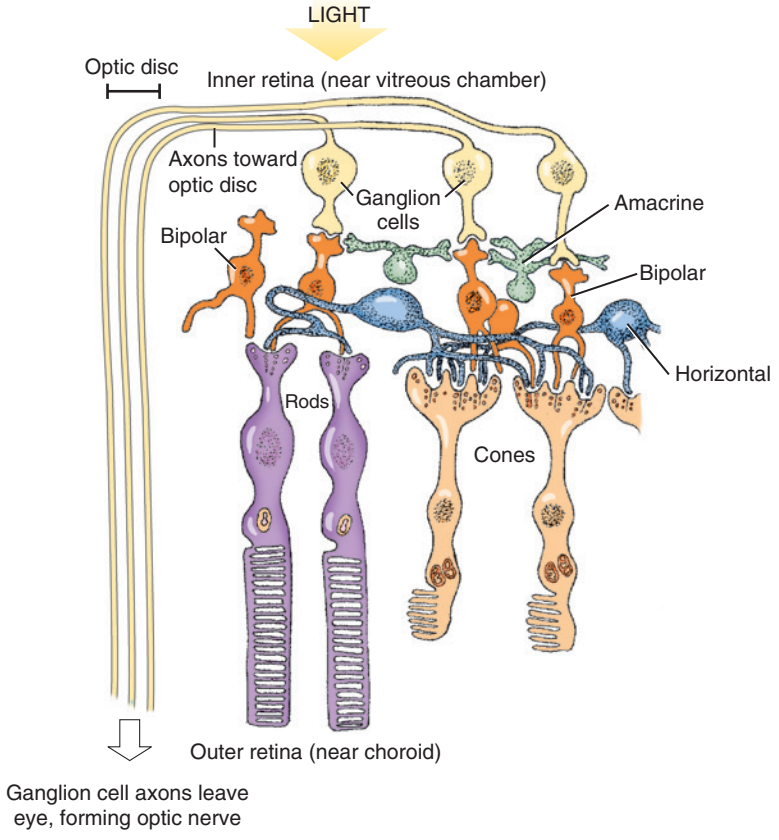
Given its elegant architecture, the retina, the neural portion of the eye, is capable of considerable processing of the visual image before the information is transmitted toward the brain. The vertebrate retina consists of five major cell types: photoreceptor cells, bipolar cells, horizontal cells, amacrine cells, and ganglion cells (Fig. 14.4). As in the cerebral and cerebellar cortices, the arrangement of these component cells is fairly consistent across the retina, suggesting a basic underlying processing mechanism and giving the retina a layered histology. However, local variations in the density of some cell types and in synaptic architecture account for particular functional variations within the retina.

There are two types of retinal photoreceptor cells: *rods* and *cones* (see Fig. 14.4). Both rods and cones make a direct synaptic connection with the interneurons called *bipolar cells,* which connect the receptors with the ganglion cells. Ganglion cell axons traverse the inner surface of the retina and converge at the optic disc to leave the eye as the optic nerve, sending action potentials to the brain.

Modifying the flow of information at the synapses among the photoreceptors, bipolar cells, and ganglion cells are two interneuron cell types: the horizontal cells and the amacrine cells (see Fig. 14.4). The *horizontal cells* mediate lateral interactions among the photoreceptors and bipolar cells. The *amacrine cells* mediate lateral interactions among the bipolar cells and the ganglion cells.

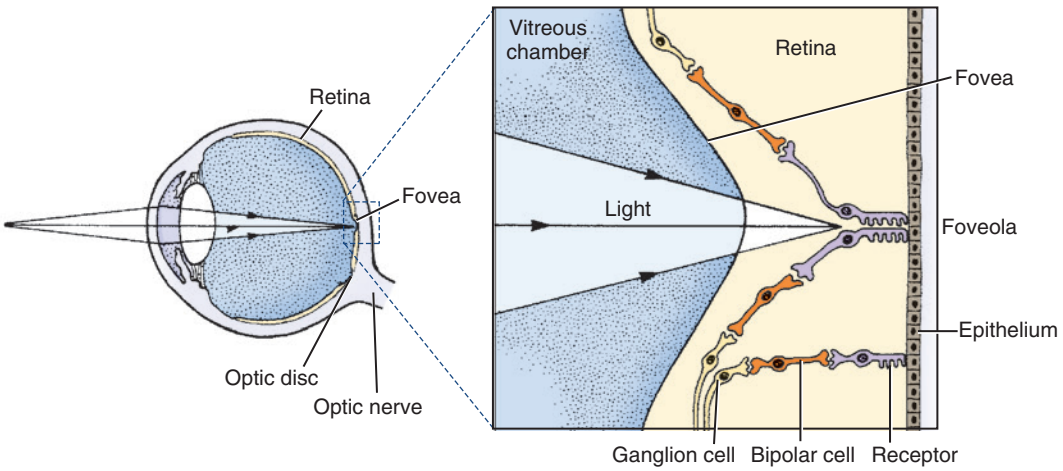
In Some Species, the Fovea Minimizes Distortion of Light Compared With Other Areas of the Retina

The retinal ganglion cells are located in the inner retina (closer to the vitreous humor), whereas the photoreceptor cells (rods and cones) are located in the outer retina (closer to the choroid; see



• **Fig. 14.4** Schematic of a patch of vertebrate retina, bordering the optic disc, that shows the five major retinal cell types: photoreceptor cells (either rods or cones), bipolar cells, horizontal cells, amacrine cells, and ganglion cells. The optic disc is formed as retinal ganglion cell axons, which run across the inner surface of the retina, leave the eye to form the optic nerve. (Modified from Kandel ER, Schwartz JH, eds. *Principles of Neural Science.* 2nd ed. New York: Elsevier Science & Technology; 1985.)

Vet-ebooks.com



• **Fig. 14.5** In most of the retina, light must first pass through inner, and then outer, layers of neurons and their processes before it reaches the photoreceptors. In the center of the fovea (found in many primates), in a region called the *foveola,* these neural elements are shifted aside; therefore light has an almost unobstructed pathway to the photoreceptors in this region. An enlarged drawing of the foveal region is shown on the *right.* (From Kandel ER, Schwartz JH, Jessell TM, eds. *Principles of Neural Science.* 4th ed. New York: McGraw-Hill; 2000.)

Fig. 14.4). Therefore throughout most of the retina, light rays travel through ganglion cells, bipolar cells, amacrine cells, and horizontal cells before reaching the photoreceptors. Although these inner neurons are unmyelinated (the ganglion cell axons become myelinated on leaving the eye) and therefore relatively transparent, they still cause some distortion of light rays.

The *fovea,* an area that demarcates the central retina in many primates, is designed to minimize this distortion. This sloping pit is formed as the neural tissue near the inner surface of the central retina is pushed aside, permitting light rays to have a less obstructed path to the outer retina. Distortion is least in the center of the fovea, in an area called the *foveola,* where light rays have almost unobstructed access to the photoreceptors (Fig. 14.5). This is functionally significant because it allows light to have a less distorted path to the region of the retina associated with the highest visual acuity (ability to discern detail). The optic disc is just nasal to the fovea.

In many mammalian veterinary species, although the central retina is also the area of highest visual acuity, no distinct fovea is formed. However, in these species, as well as primates, major blood vessels that can potentially interfere with the light path avoid passing across the central retina. Furthermore, in the veterinary species, the region of high visual acuity has more of an ovoid, horizontal orientation. The locus of highest acuity is dorsal and temporal to the optic disc with a shorter extension temporally and a longer, but increasingly thinner, extension nasally (Fig. 14.6). It is believed that this horizontal orientation may provide better acuity for a wider field of view when scanning the horizon.

Tissue Layers Behind the Retina Principally Absorb Light, or Can Reflect Light, Depending on the Species' Habits

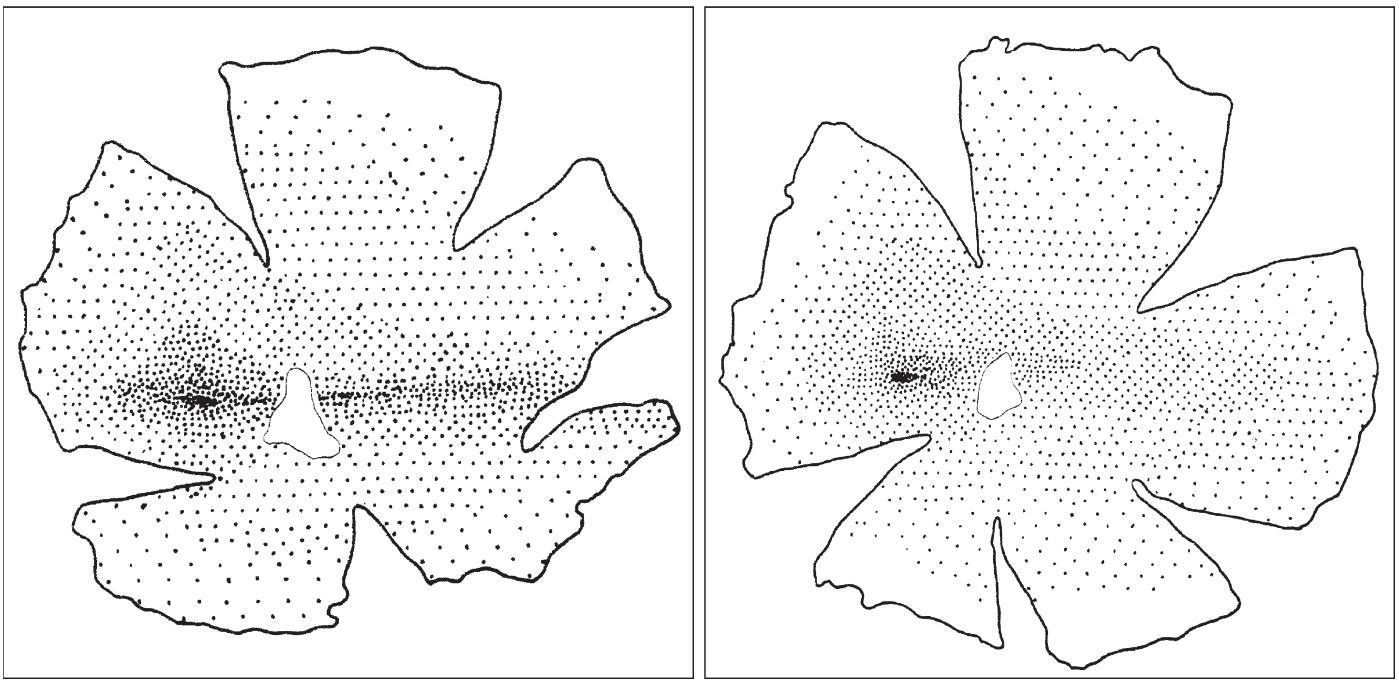
In animals that rely heavily on acute, daylight vision, there is a dark melanin pigment in the epithelial layer between the photoreceptors and the choroid. This pigment absorbs light that has passed by the photoreceptors without stimulating them. If such light were reflected back into the retina, the sharpness of the visual image would be blurred. In nocturnal animals and most domestic mammals, however, there is a patch of reflective material within the choroid called the *tapetum* (see Fig. 14.2). When the tapetum is present, the region of the epithelial layer that overlies it does not contain the dark, light-absorbing pigment noted above. This arrangement facilitates the reflection of nonabsorbed light back onto the retina allowing it to make optimal use of what light it receives, but at the expense of visual acuity. Reflection of light off the tapetum causes the familiar "night shine" from nocturnal animals' eyes.

Photoreception and Transduction of Light Occur in the Rods and Cones

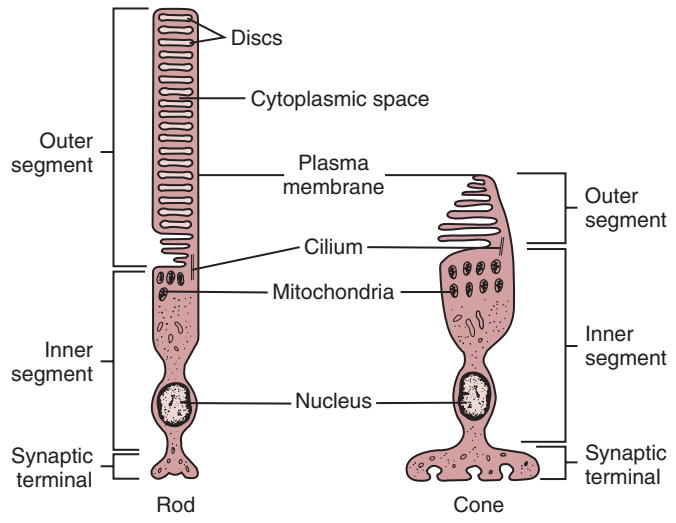
The anatomical structures of the rod and cone photoreceptors are similar, but there are some important differences. Because they are neurons, the rods and cones are primary receptors. Both cell types are divided into three parts: a synaptic terminal, an inner segment, and an outer segment (Fig. 14.7). The photoreceptor synaptic terminal synapses with the bipolar cells. The inner segment includes the nucleus, mitochondria, and other cytoplasmic structures. The inner and outer segments are connected by a microtubule-containing cilium. The outer portions are specialized for photoreception. They contain an elaborate array of stacked membranous discs whose membranes contain visual photopigment. *Visual photopigment* is a light-sensitive molecular complex that initiates a biochemical chain of events, transducing light into neural electrical signals.

The discs in the outer segments of the photoreceptors are regularly being formed near the cilium, phagocytized by the pigmented epithelium, and their photopigment cycled back to newly formed discs. Loss of this normal turnover in the outer segment may be important in several retinal diseases (e.g., retinitis pigmentosa). Photopigment of the discs is made up of a protein, called *opsin,* and *retinal,* an aldehyde of vitamin A. The light-sensitive retinal molecule

Vet-ebooks.com



• **Fig. 14.6** Representation of the visual streak region in flattened portions of the right retinas of two different breeds. In both images, temporal is to the reader's left and nasal is to the right. The streak region has a horizontally oriented, ovoid shape with a more rounded head temporally and a progressively narrowing tail nasally. Dot intensity indicates retinal ganglion cell density, which is a reflection of visual acuity. The irregularly shaped white hole near the center of each image is the optic disc. Left image is from a German shepherd (more "wolf-like") and right image is from a beagle. (Modified with permission from "The topography of ganglion cells in the dog and wolf retina" by Leo Peichl, *J Comp Neurol* 1992;324:603-620.)



• **Fig. 14.7** The two types of photoreceptors, rods and cones, have characteristic structures. Both rod cells and cone cells are differentiated into inner and outer segments connected by a cilium. The inner segments of both cell types contain the nucleus and most of the cell's biosynthetic machinery and are continuous with the synaptic terminals. The membranous discs in the outer segment contain the light-transducing photopigment. The discs in the outer segments of rod cells are separated from the plasma membrane, whereas the discs of cone cells are not. (Modified from O'Brien DF. The chemistry of vision. *Science.* 1982;218[4576]:961–966; from Kandel ER, Schwartz JH, Jessell TM, eds. Principles of Neural Science. 4th ed. New York: McGraw-Hill; 2000.)

is bound to the opsin, which is a member of the G-protein–coupled membrane receptor family (see Chapter 1). When light is absorbed by a rod or cone, the retinal is transformed in a way that activates the opsin, producing a change in the concentration of an intracellular second messenger, which leads to a change in the membrane potential of the photoreceptor cell. The configurational change in retinal is the only light-sensitive event in vision.

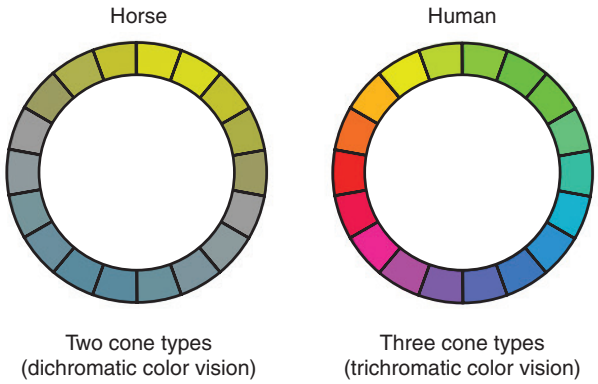
Unlike most sensory receptor cell membranes that depolarize with stimulation, photoreceptors hyperpolarize when struck by light. In rods, the visual photopigment is called *rhodopsin.* In the dark, many gated sodium/calcium ion channels remain open, allowing leakage of Na^+ and Ca^{2+} ions into the rod, which keeps the membrane in a depolarized state. When photons of light strike rhodopsin, the resulting change in second-messenger concentration produces a closing of many of the gated Na^+/Ca^{2+} channels. The result is a hyperpolarization of the receptor cell membrane and a decrease in transmitter released at the synapse with the bipolar cell. Photoreception in cones works in a similar manner except that the photopigment opsin is different from that of rhodopsin. Depending on the species, different populations of cones may have different opsins. Because different opsins differ in their ability to absorb particular wavelengths of light, the number of these different cone populations in a species is related to the ability of that species to discriminate different colors.

Differences in the functional properties of rods and cones, differences in their retinal distribution, and differences in the synaptic organization of other retinal neurons to which they pass their information result in a "rod retinal system" and a "cone retinal

Vet-ebooks.com

TABLE 14.1 Functional Differences Between Rod and Cone Systems

Rod System	Cone System
Most sensitive to light	Less sensitive to light
Night vision (low light)	Day vision (normal indoor and daylight)
Low acuity	High acuity (good at discerning image detail)
Achromatic	Color vision
Peripheral retina	Central retina



• **Fig. 14.8** Color wheels representing the difference in color perception between the horse, with two populations of cones (dichromatic color vision), and human, with three populations of cones (trichromatic color vision). (Modified from Carrol J, Murphy CJ, Neitz M, et al. Photopigment basis for dichromatic color vision in the horse. *J Vis Sci.* 2001;1:80.)

system" that have different functional attributes (Table 14.1). Because individual rods are more sensitive to light than individual cones, and because several rods feed their synaptic input to a single bipolar cell (convergence), the rod system is the night-vision system concerned with detecting the presence of light. Because of the convergence of many rods on a bipolar cell, however, the rod system is not optimal for discerning the detail of a visual image (poor visual acuity). Also, because the rods are more highly concentrated in the peripheral retina than the cones, this region is better at detecting the presence of light than discerning its details. Finally, because all rods contain the same visual photopigment, the rod system cannot discriminate different colors.

As noted, a cone is less sensitive to light than a rod. Also, one or only a few cones feed their synaptic input to a single bipolar cell. Therefore the cone system needs higher levels of illumination to function compared with the rod system. In fact, the cone system actually represents the day-vision system because the rod system does not function well in high levels of illumination. However, due to a lack of convergence onto bipolar cells, the cone system is good at discerning image details. Because the cones have a particularly high density in the region of the fovea, or central retina, this area is the part of the retina with the highest visual acuity. This makes sense considering the fovea, or central retina, is where light falls when an animal looks directly at an object of interest. Again, different populations of cones can contain different photopigments, each differentially sensitive to a range of light frequencies. Thus the cone system is also the color-vision system. Generally, the more cone photopigments present in the retina, each within a different population of cones, the better the ability of the species to discriminate different wavelengths of light, or to "see" color (Fig. 14.8). In primates, there are three different populations of cones, each with a different photopigment. According to the *Young-Helmholtz theory,* the brain assigns color to a given wavelength of light by comparing the relative activation of the different cone populations. Humans who lack one or more of these cone types, as a result of faulty genetic transmission, cannot see colors within particular ranges of the visible spectrum of light. The most common types of such "color blindness" are linked to the X chromosome.

The extent to which various veterinary species perceive color is still controversial. It is believed that mammalian ancestors may have had four different types of cones, whereas early nocturnal mammals may have had only two types of cones, having traded some of their color discrimination ability for rod light sensitivity. Most current mammals, including dogs, cattle, swine, and horses, have only two types of cones. As noted, primates have three types of cones. It is thought that their successful exploitation of color-rich arboreal environments may have induced selection pressure for additional color discrimination ability, above that imparted by two populations of cones. Modern birds appear to have four populations of cones, one of which is sensitive to light in the ultraviolet (UV) range of the spectrum. Behavioral experiments that show their ability to discriminate UV light suggest that birds have a richer color perception than primates. Lizards, turtles, and some fish also possess UV-sensitive cones. Only primates are known to have the color vision with which humans are familiar. Although there is some evidence that cats may also have three types of cones, behavioral studies suggest all three types may not be used for color discrimination with cats exhibiting vision akin to human red-green color blindness.

Visual Image Processing in the Retina Begins as the Response of the Photoreceptor to Light Is Synaptically Transmitted to the Ganglion Cells by the Bipolar Cells

The hyperpolarizing response of the rods and cones to light synaptically influences bipolar cells; in turn, the bipolar cell influences action potential frequencies in the ganglion cell axons on their way to the brain. As alluded to earlier, this transmission of information from the outer to the inner retina can be modulated by the horizontal cells and the amacrine cells (see Fig. 14.4). The horizontal cells are in communication with photoreceptor-bipolar cell synapses, and with each other, to allow lateral communication among different photoreceptor-bipolar cell synapses. Amacrine cells can serve a somewhat analogous function with respect to bipolar-ganglion cell synaptic interactions. An interesting consequence of this architecture is that light that hits a specific point on the retina can influence the activity and/or responsiveness to light of cells in an adjacent patch of retina.

Considerable processing of the visual image occurs in the retina, and it is particularly good at detecting luminance changes of small spots of light. The initial stages of color discrimination, contrast

Vet-ebooks.com

detection and enhancement, and directional sensitivity also occur in the retina. The synaptic interactions among photoreceptors, bipolar cells, and horizontal cells play a significant role in contrast enhancement and directional sensitivity. The interactions of amacrine cells with bipolar cells and ganglion cells appear to play a role in processes such as the ability to detect *changes* in the rate of alternation between light and dark. A more detailed description of the synaptic and membrane changes in the chain of transmission within the retina, accounting for these types of retinal image processing, is beyond the scope of this book. To learn more about the many interesting and unusual phenomena occurring in the retina, the reader should refer to the Bibliography.

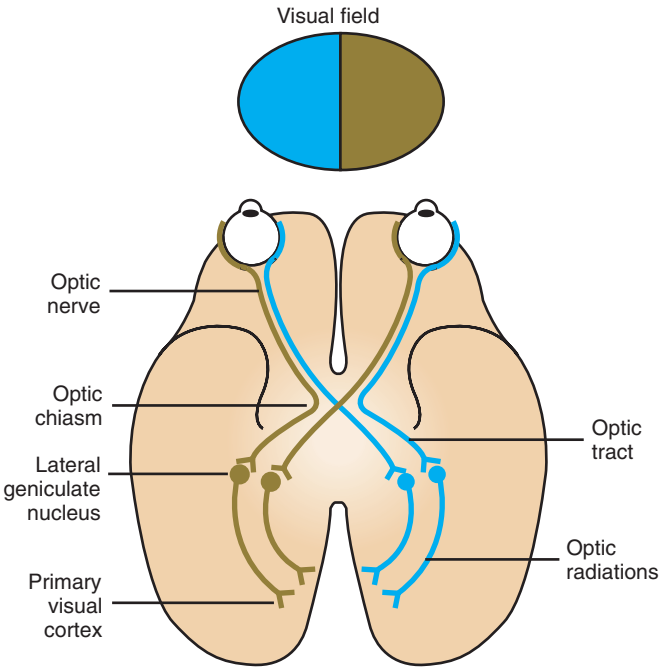
The Electroretinogram Records the Electrical Response of the Retina to a Flashing Light

The *electroretinogram* (ERG) is a clinical electrophysiological recording from the cornea and skin near the eye. It records the electrical response of the retina to a light flashed into the eye. It has three waves: the *A wave,* corresponding primarily to the activation of visual photopigment and photoreceptors; the *B wave,* caused primarily by the response of retinal bipolar cells; and a slower *C wave,* thought to originate in the pigment epithelium. Wave latencies and amplitudes with respect to flash onset, or between waves, can be used to assess the integrity of retinal function. The ERG is a fairly simple, general electrodiagnostic tool for assessing some forms of retinal dysfunction or degeneration. It can be particularly useful if a visual deficit is identified by neurological examination but ophthalmoscopic results appear normal, or to assess potential retinal dysfunction when cataract prohibits proper ophthalmoscopic examination of the retina.

Retinal Ganglion Cells Transmit Neural Signals to the Visual Cortex by Way of the Lateral Geniculate Nucleus

Sets of retinal ganglion cell axons leaving the eye participate in three important visual pathways: the *retino-geniculo-striate pathway,* the *retino-tectal pathway,* and the *retino-hypothalamic pathway.* The retino-geniculo-striate pathway is principally involved in conscious visual perception of form, color, motion, orientation, and depth. The retino-tectal pathway plays an important role in pupillary reflexes and reflex orientation of the eye to visual targets. The retino-hypothalamic pathway plays a role in the regulation of physiological rhythms by light-dark cycles (e.g., seasonal changes in day length).

For the retinal image, originating from light in the visual field, to reach consciousness, the information must be transferred to the visual cortex. Fig. 14.9 shows the retino-geniculo-striate pathway by which the axons of the retinal ganglion cells project to the lateral geniculate nucleus of the thalamus and by which cells in the *lateral geniculate nucleus* project their axons to the primary visual cortex of the occipital lobe. Note that ganglion cell axons from the temporal retina (closest to the ear; see Fig. 14.1) travel along the *optic nerve* to the optic chiasm and then project ipsilaterally to the lateral geniculate nucleus on the same side of the brain. Ganglion cell axons from the nasal retina (closest to the nose) come to the *optic chiasm* and cross to the contralateral lateral



• **Fig. 14.9** Retino-geniculo-striate pathway to cerebral cortex. Axons of the retinal ganglion cells project to the lateral geniculate nucleus of the thalamus, and axons from cells of the lateral geniculate nucleus project to the ipsilateral primary visual cortex of the occipital lobe. The right environmental visual field projects to the left half of each retina, and the left environmental visual field projects to the right half of each retina. The right half of each retina sends information to the right lateral geniculate, and from there to the right primary visual cortex. The left half of each retina sends information to the left lateral geniculate, and from there to the left primary visual cortex. Therefore information from the left visual field arrives in the right primary visual cortex, and information from the right visual field arrives in the left primary visual cortex. The labeled components of each pathway are present on both sides of the brain.

geniculate nucleus. Retinal ganglion cell axons between the region of the optic chiasm and the lateral geniculate nuclei are referred to as the *optic tracts.* Cells in each lateral geniculate nucleus then send axons to the ipsilateral *primary visual cortex* of the occipital lobe by way of the *optic radiations.* This arrangement results in the right half of each retina sending information to the right primary visual cortex and the left half of each retina sending information to the left primary visual cortex. Light originating in the left half of the environmental visual field generally falls on the right half of the retina of each eye, whereas light originating in the right half of the environmental visual field falls on the left half of the retina of each eye. Given the anatomical map of the retino-geniculo-striate pathway in Fig. 14.9, information from an image arising in the left field of vision would be received in the right primary visual cortex, and image information arising from the right field of vision would be received in the left primary visual cortex.

The features of the visual image to which lateral geniculate nucleus cells respond are surprisingly similar to those features processed by the retina (see earlier discussion). However, more complex feature extraction can begin to be seen in the primary visual cortex. For example, whereas the retina processes information

Vet-ebooks.com

about small spots of light and the lateral geniculate nucleus processes information about larger spots of light composed of those smaller spots, the primary visual cortex can process information about bars or edges of light composed of those larger spots. In addition, the primary visual cortex also processes information about the orientation of those bars or edges and about whether the information comes from one eye or both eyes. Information reaching the *primary visual cortex,* or *V1* (also called *striate cortex*), is passed to other occipital lobe areas of the visual cortex (e.g., V2, V3) for still more complex processing, then to visual association areas of the parietal and temporal lobes for the highest forms of visual information processing and integration with other sensory modalities. There appears to be parallel processing of retinal information within the retino-geniculo-striate pathway and its higher-order processing regions. One major stream of information in this system apparently is concerned with conscious object identification (e.g., form, color), whereas a parallel stream of information relates to conscious awareness of where an object is located (e.g., motion, position). The three major projection pathways of retinal ganglion cell axons—retino-geniculo-striate, retino-tectal, and retino-hypothalamic—represent another example of the parallel processing of retinal information.

The *visual evoked potential* (VEP) is an electrodiagnostic technique that can aid in identification of abnormalities in the post-retinal visual pathway. Recordings to transient photic stimuli (e.g., light flashes) are made from the skin over the region of primary visual cortex. Because the signal is usually small, multiple recordings are usually averaged so that the random background "noise" of the EEG (see Chapter 16) activity drops out and the small evoked signal remains. Analogous to the brainstem auditory evoked response (BAER; see Chapter 16), the multiple peaks of the evoked signal represent the activity of different structures in the pathway from retina to the region of visual cortex. Changes in the latencies and amplitudes of the VEP waveform peaks can be used to evaluate suspected abnormalities in the pathway.

The Diameter of the Pupil Is Controlled by the Autonomic Nervous System

The iris of the eye contains two sets of smooth muscle fibers. One set, arranged in a circular pattern around the pupil, causes the pupil to become smaller (*constrict*) when the fibers contract. These constrictor fibers are innervated by postganglionic parasympathetic neurons whose cell bodies lie in the ciliary ganglion, just behind the eye, and which secrete acetylcholine as the neurotransmitter to the muscle. These parasympathetic postganglionic neurons are activated by parasympathetic preganglionic neurons whose axons travel in the *oculomotor cranial nerve* (cranial nerve 3) and whose cell bodies reside in the *Edinger-Westphal nucleus* of the midbrain.

The other smooth muscle fibers of the iris are arranged radially from the pupil, resembling spokes of a wheel. When these radial smooth muscle fibers contract, they cause the pupil to become larger (*dilate*). These dilator fibers are innervated by the sympathetic nervous system. The sympathetic preganglionic neurons begin in the first three or four thoracic segments and course cranially in the vagosympathetic trunk to synapse in the superior cervical ganglion of the neck. The sympathetic postganglionic axons originate in cells of this ganglion and course to the region of the eye, where they innervate the dilator fibers of the iris, in addition to a muscle that helps lift the upper eyelid and a muscle that helps keep the "third eyelid" in place at the medial canthus of the eye. The sympathetic postganglionic axons also innervate sweat glands and vascular smooth muscle to the face.

The Retina, Optic Nerve, and Autonomic Nerve Supply Controlling the Pupil Can Be Tested With a Flashlight

When a light is shone into the eye, the pupil of that eye constricts. This action is called the *direct pupillary light reflex* (Fig. 14.10). The light triggers the photoreception mechanism, leading to ganglion cell action potentials transmitted along the optic nerve. Some of the ganglion cell axons of the retino-tectal pathway synapse in the *pretectal nuclei* of the brain (near the diencephalic/mesencephalic border). The pretectal neurons then synapse on the parasympathetic preganglionic neurons of the Edinger-Westphal nucleus (in the mesencephalon), whose axons travel in the oculomotor nerve to synapse on parasympathetic postganglionic neurons in the ciliary ganglion (of the orbit). Stimulation of these postganglionic neurons causes constriction of the pupil by stimulating the constrictor smooth muscle fibers of the iris. A normal direct pupillary light reflex tests the integrity of the retina, the ipsilateral second and third cranial nerves, a limited region of the brainstem, and the iris. Given that a significant proportion of optic nerve axons cross the midline at the optic chiasm, and that many axons of pretectal neurons also cross the midline (see Fig. 14.10B), when a light is shone into one eye, not only does the pupil on the same side constrict (direct pupillary light reflex), but the contralateral pupil also constricts. This action is called the *indirect* or *consensual pupillary light reflex.* It also requires the integrity of the contralateral oculomotor (third) cranial nerve.

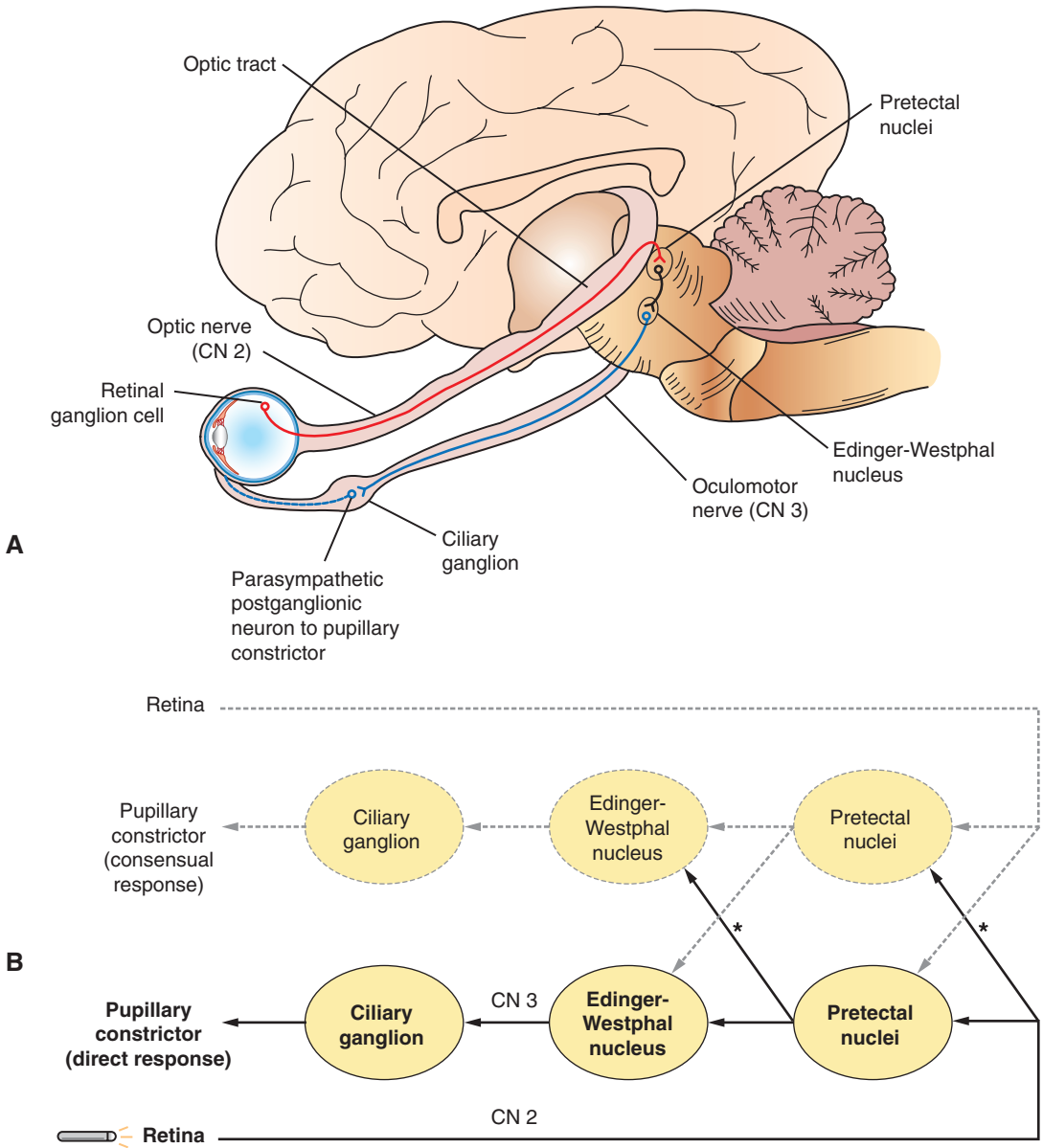
Aqueous Humor Determines Intraocular Pressure

Aqueous humor is a clear liquid found in the anterior and posterior chambers of the eye. Its rate of production and absorption is sufficiently high to replace the entire chambers' volume several times a day.

Aqueous humor is produced by the epithelium covering the *ciliary processes,* a system of finger-like projections on the ciliary body of the posterior chamber. Aqueous humor is thought to be formed by the active transport of sodium, chloride, and bicarbonate ions into the posterior chamber. This establishes an osmotic gradient, causing water to flow passively into the posterior chamber. Aqueous humor flows from the posterior to the anterior chamber through the pupil. Flow is caused by a pressure gradient established by the active process of formation in the posterior chamber.

Aqueous humor is then absorbed into the venous system at the angle between the cornea and the iris. This absorption is driven by a pressure gradient and is assisted, in many species, by a system of trabeculae and canals. If this absorption into the venous system is obstructed, intraocular pressure increases because the production of aqueous humor continues. This pathological increase in intraocular pressure is called *glaucoma.* As intraocular pressure exceeds intravascular pressure in the blood supply to the retina, blindness results.

Vet-ebooks.com



• **Fig. 14.10** The pupillary light reflex. (A) The structures and synapses involved in the direct pupillary light reflex represented on one side of the animal. (B) Schematic diagram showing connections that cross the midline (*) responsible for the consensual (indirect) pupillary light reflex. Retinal axons that cross the midline do so at the optic chiasm. Although bilaterally symmetrical, structures and connections originating on one side of the animal have been de-emphasized for ease of interpretation by the reader. *CN,* Cranial nerve.

Vet-ebooks.com

CLINICAL CORRELATIONS

Homonymous Hemianopia

History

You examine a 10-year-old male German shepherd whose owner reports that the dog has recently begun to bump into objects with the left side of his face and has had two seizures. The seizures were characterized by turning of the head to the left and stiffening of the left front leg.

Clinical Examination

Physical examination abnormalities are limited to the nervous system. When presented with a maze of unfamiliar objects in the examination room, the dog collides with objects as if he does not see from the left side. He seems somewhat weak in the left front leg. The dog is otherwise bright, alert, and responsive. Cranial nerve and spinal segmental reflexes are within normal limits, as are intersegmental, proprioceptive placing responses for the right front and right rear legs. However, the proprioceptive placing responses for the left front and left rear legs are quite prolonged.

Comment

This dog's history and neurological examination abnormalities are common in dogs with brain tumors. This dog has a tumor (neoplasm) arising from the meninges over the right posterior cerebral cortex. It is in this posterior (occipital) cortex that the visual image is interpreted from the visual field of the left side (see Fig. 14.9). Lesion-induced functional damage to the right occipital cortex results in loss of vision from the left visual field (*homonymous hemianopia*). Also in the right cerebral cortex, the conscious proprioception response for the left legs is interpreted. This dog's seizures feature turning the head to the left and displaying transient rigidity of the left front leg, because the seizure activity arose from the cerebral cortex at the site of the tumor and spread to the right motor cortex but remained limited to the cerebral cortex on the right side. Because the pyramidal system's corticospinal tract controlling the muscles of the left neck and left front leg arise in the right motor cortex (see Chapter 10), seizure activity causes the transient head turning and left leg stiffness.

Treatment

This dog had a meningioma of the right posterior cerebral cortex. Surgical removal was not attempted in this case.

Uveitis in a Horse

History

A client calls you out to look at their 8-year-old Appaloosa mare's right eye for squinting, pain, and discharge of a few days duration. The horse does not have any other clinical signs other than the eye.

Clinical Examination

The ophthalmic examination demonstrates a miotic (constricted) pupil, corneal edema (giving it a hazy appearance), and flare (clouding of aqueous humor), of the right eye. You use fluorescein dye to stain the eye, and there is a 3 × 4-mm superficial corneal ulcer present. After dilating the eye with tropicamide, you see the optic disc looks edematous (swollen) at the edges. Your examination on the left eye reveals no abnormalities.

Comment

This mare likely has uveitis with an associated corneal ulcer. The uvea is comprised of the choroid, iris, and ciliary body. It is possible that she developed the uveitis, due to a *Leptospira* species (a bacterial spirochete) or another cause, and then because the eye was painful, she scratched the cornea to cause the ulcer. Alternatively, the mare could have developed the ulcer by rubbing her face or from a scratch from a foreign body (i.e., dirt, twigs), and then she developed uveitis secondary to the ulcer.

With uveitis, there is a breakdown of the normal blood-aqueous barrier, allowing protein, cells, and fibrin to enter the aqueous humor, clouding it. Corneal endothelial cells are also affected, which can lead to reduced fluid removal from the corneal tissue, resulting in corneal edema. The lacrimal glands, under parasympathetic stimulation, release tears. Because of the inflammation, the ciliary body and pupillary sphincter can constrict or spasm, which can affect the pupil and may be painful.

Treatment

With the corneal ulcer present, steroids are usually not used topically, because there would be a greater chance for infection. The horse is treated with topical atropine, which is an anticholinergic, to help dilate the pupil and relieve ciliary spasm, thus decreasing pain. Horses are also treated with topical antibiotics for possible infection and also with systemic antiinflammatory medications, such as flunixin meglumine. In some cases, topical antiinflammatory medications are also needed. The horse is treated for a few days after the ulcer resolves to maximize the chances of tear production normalizing. If the optic disc is still edematous and the mare is still in pain, based on squinting and/or miotic pupil, she could be switched over to topical steroids to provide better control of inflammation.

Practice Questions

1. A patient whose left pupil diameter is smaller than normal, whose left upper eyelid droops, and whose left "third eyelid" is abnormally displaced likely has a lesion of which of the following structures?
- **a.** Left oculomotor nerve
- **b.** Left vagosympathetic nerve trunk
- **c.** Right oculomotor nerve
- **d.** Right vagosympathetic nerve trunk
- **e.** Left optic tract

2. Which of the following is *not* characteristic of the cone system?
- **a.** Many photoreceptors converging on single bipolar cell
- **b.** Good visual acuity
- **c.** Multiple photopigments
- **d.** Day vision
- **e.** Color vision
- **f.** All the above are characteristic of the cone system

3. In the veterinarian's waiting room, your friend, a member of the soccer team, is trying without much success to explain the cause for the team's recent losing trend. A variety of implausible explanations are proposed until she mentions that she is progressively losing vision, but only in her left visual field, and has frequent headaches. You recommend that she see a neurologist because she likely has a lesion in her:
- **a.** Left optic tract
- **b.** Right optic nerve
- **c.** Right optic tract
- **d.** Optic chiasm
- **e.** Left optic radiations

Vet-ebooks.com

4. Which of the following regarding the lens is *false*?
   a. It lies behind the iris.
   b. It plays an important role in focusing a visual image on the retina.
   c. It contains layered fibers in an "onion-like" arrangement.
   d. Its shape does not normally change.
   e. An increase in its opacity is called a *cataract.*

5. You examine a patient's pupillary light reflexes. Shining a light into the left eye produces both a positive direct and a positive consensual pupillary response. However, shining the light into the right eye produces neither a direct nor an indirect pupillary response. This patient's pathological lesion is located in which one of the following structures?
   a. Left optic nerve
   b. Left oculomotor nerve
   c. Right optic nerve
   d. Right oculomotor nerve
   e. Left primary visual cortex

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

Carrol J, Murphy CJ, Neitz M, et al. Photopigment basis for dichromatic color vision in the horse. *J Vis*. 2001;1:80.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Goldsmith TH. What birds see. *Sci Am*. 2006;295(1):68–75.

Goldstein BE. *Sensation and Perception*. 8th ed. Pacific Grove, Calif: Wadsworth; 2009.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Maggs D, Miller P, Ofri R. *Slatter's Fundamentals of Veterinary Ophthalmology*. 6th ed. St. Louis: Elsevier/Saunders; 2017.

Miller PE, Murphy CJ. Vision in dogs. *J Am Vet Med Assoc*. 1995;207(12):1623–1634.

Mowat FM, Petersen-Jones SM, Williamson H, et al. Topographical characterization of cone photoreceptors and the area centralis of the canine retina. *Mol Vis*. 2008;14:2518–2527.

Nicholls JG, Martin AR, Fuchs PA, Brown DA. *From Neuron to Brain*. 5th ed. Sunderland, Mass: Sinauer; 2012.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Sheppard AL, Davies LN. In vivo analysis of ciliary muscle morphologic changes with accommodation and axial ametropia. *Invest Ophthalmol Vis Sci*. 2010;51(12):6882–6889.

Vet-ebooks.com

15

Cerebrospinal Fluid and the Blood-Brain Barrier

BRADLEY G. KLEIN

KEY POINTS

1. Cerebrospinal fluid has many functions.
2. Most cerebrospinal fluid is formed at the choroid plexus of the ventricles.
3. Cerebrospinal fluid flows down a pressure gradient through the ventricular system into the subarachnoid space.
4. Cerebrospinal fluid is part of a brain "glymphatic" system that helps clear waste products.
5. Cerebrospinal fluid is absorbed into the venous system.
6. Hydrocephalus is an increased volume of cerebrospinal fluid in the skull.
7. Permeability barriers exist between blood and brain.

Cerebrospinal fluid (CSF) is a clear fluid present in the ventricles (core cavities) of the brain, in the central canal that runs through the core of the spinal cord, and in the subarachnoid space that surrounds the entire outer surface of the brain and spinal cord (Fig. 15.1). The CSF contains almost no blood cells, little protein, and differs from plasma with respect to the concentration of various ions. Its rates of formation, flow, and absorption are sufficiently high to cause its replacement several times daily. Sampling its pressure, cell count, and levels of various biochemical constituents is a common diagnostic procedure for central nervous system (CNS) pathology called a *spinal tap.* Injecting radiopaque dyes into the CSF of the subarachnoid space is the basis of a common neuroradiographic technique called *myelography,* often used in conjunction with computerized imaging, which can assess the integrity of the spinal canal. Obstruction of the flow of CSF produces a condition called *hydrocephalus.* An understanding of the formation, flow, and absorption of CSF is essential for understanding these diagnostic procedures and the pathophysiology of hydrocephalus.

The *blood-brain barrier* (BBB) refers to the selective nature of CNS blood vessels with respect to the materials that can move across their walls, compared with blood vessels in other parts of the body. Understanding the BBB helps clarify how the brain is protected from potentially harmful neuroactive chemicals in the blood, and why it is difficult to deliver certain drugs effectively to the brain.

Cerebrospinal Fluid Has Many Functions

To work properly, the CNS needs protection not only from physical injury but also from significant variation in the local environment of its neurons. A buildup of toxins or a significant change in ionic concentration in this neuronal microenvironment could result in pathological changes in neuronal physiology.

One of the most important functions of CSF is to cushion the brain, protecting it against blows to the head. Because the specific gravities of the brain and CSF are similar, the brain floats in the fluid. Thus the force of a blow to the head is buffered by the CSF instead of being transferred directly to brain tissue.

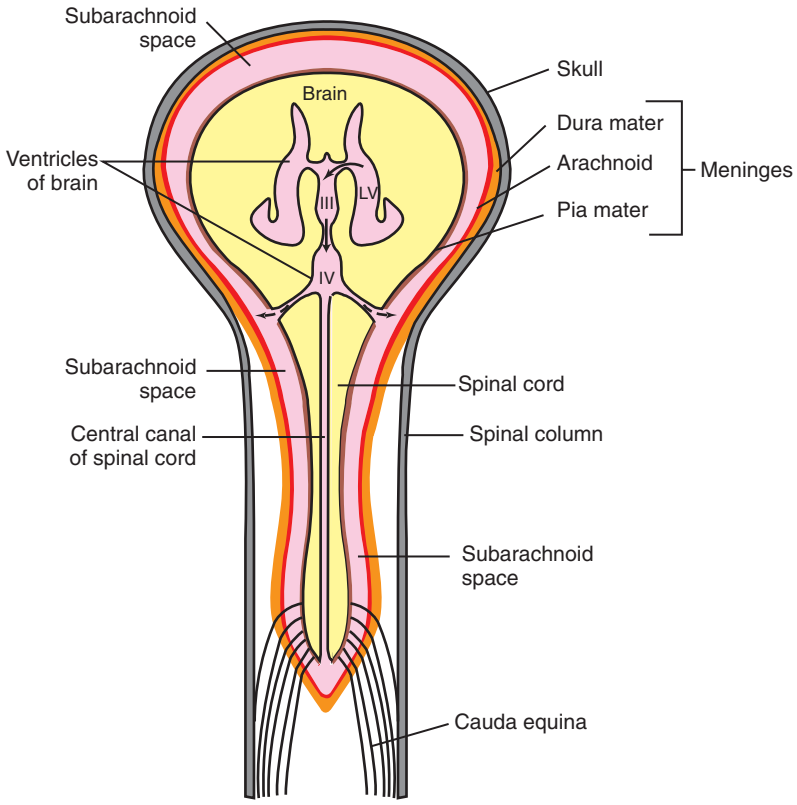
Because the composition of the CSF is tightly controlled and in equilibrium with the extracellular fluid of the brain and spinal cord, the CSF also helps maintain a consistent extracellular microenvironment for the neurons and glia of the CNS. This diffusional equilibrium between the CSF and the extracellular fluid, in conjunction with the flow and multiple daily turnover of the CSF, also makes the CSF an effective waste control system that can remove potentially harmful cellular metabolites. Evidence indicates that these properties may also allow the CSF to function as a brain distribution system for some polypeptide hormones, growth factors, cytokines, and immune cells secreted into the CSF, some of which are important in brain development.

Most Cerebrospinal Fluid Is Formed at the Choroid Plexus of the Ventricles

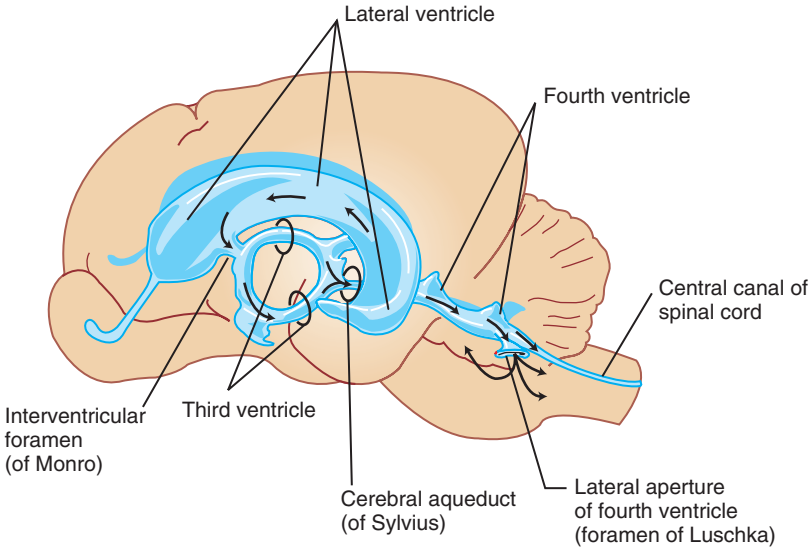
The *ventricles* are a series of interconnected cavities in the core of the brain that have an ependymal cell lining and are filled with CSF (Fig. 15.2). The *lateral ventricles* are respectively located in the two cerebral hemispheres, the *third ventricle* is found at the midline of the diencephalon, and the *fourth ventricle* is located between the cerebellum and the dorsal surface of the hindbrain (pons and medulla) (Fig. 15.3).

A choroid plexus located in each of the four ventricles forms the majority of CSF. These are small, cauliflower-like growths of clustered villi that form a portion of the floor or roof of each ventricle (see Fig. 15.3). The plexuses consist of tufts of capillaries covered by a layer of epithelial cells. These epithelial cells, unlike the cells lining the rest of the ventricle, form a selective, tight-junction barrier to the secretions of the leaky capillaries and to other surrounding fluids (e.g., CSF, extracellular fluid). Membrane

Vet-ebooks.com



• **Fig. 15.1** Schematic diagram of the relationships among the central nervous system, ventricles, cerebrospinal fluid (CSF), and meninges. The CSF is colored pink. *dashed curved arrow,* Lateral aperture of the fourth ventricle; *III,* third ventricle; *IV,* fourth ventricle; *LV,* lateral ventricle; *solid curved arrow,* interventricular foramen; *solid straight arrow,* cerebral aqueduct. (Modified from Behan M. Organization of the nervous system. In: Reece WO, ed. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.)

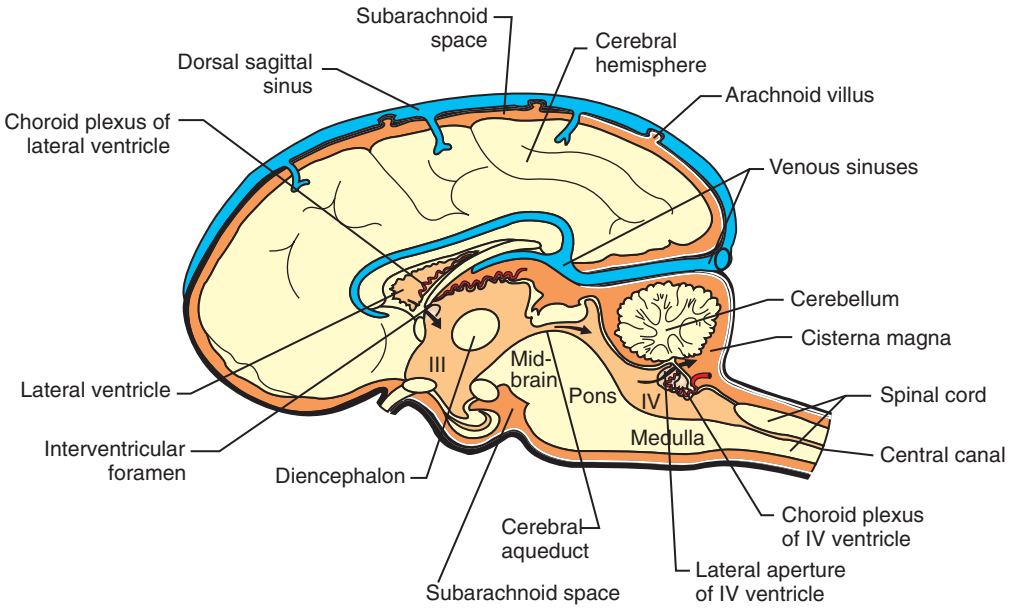


• **Fig. 15.2** Lateral view of the ventricular cavities and their approximate spatial position within the brain. *Arrows* represent the flow of cerebrospinal fluid through and ultimately out from the ventricles at the lateral apertures. (From De Lahunta A, Glass E. *Veterinary Neurology and Clinical Neurology.* 3rd ed. St. Louis: Saunders; 2009.)

transporters and selective channels regulate the passage of ions and molecules across the epithelial cell barrier, effectively controlling the composition of the CSF synthesized in the ventricle. The active transport of sodium ions (Na^+) contributes to a net movement of sodium chloride (NaCl) into the ventricles. This osmotic gradient regulates the water content of the CSF as water follows the NaCl passively into the ventricle. There is evidence that some potentially harmful metabolic waste products deposited in the CSF can actually be absorbed and removed by the choroid plexus.

It is important to note that CSF is formed at an almost constant rate, independent of either CSF pressure or blood pressure. Therefore if CSF pressure or general intracranial pressure were to rise as a result of an obstruction to CSF flow or a space-occupying mass, CSF formation would continue.

Vet-ebooks.com



• **Fig. 15.3** Midsagittal section of the brain showing portions of the ventricles and subarachnoid space, the choroid plexuses that produce cerebrospinal fluid (CSF), and the dorsal sagittal sinus into which much of the CSF is absorbed. The cisterna magna is a common location for the sampling of CSF. The CSF within the ventricles is colored *light tan,* and the CSF that has emptied from the ventricles into the subarachnoid space (through the lateral apertures) is colored a *darker tan.* The larger portion of the lateral ventricle lies within the hemisphere, hidden from view. *III,* Third ventricle; *IV,* fourth ventricle. (Modified from Fletcher TF. Spinal cord and meninges. In: Evans HE, ed. *Miller's Anatomy of the Dog.* 3rd ed. Philadelphia: Saunders; 1993.)

Cerebrospinal Fluid Flows Down a Pressure Gradient Through the Ventricular System Into the Subarachnoid Space

CSF flows, by bulk, down a pressure gradient from its site of formation at the choroid plexuses through the ventricular system and subarachnoid space into the venous system. Fluid formed in the lateral ventricles passes into the third ventricle through the interventricular foramina (foramina of Monro) (see Figs. 15.1 to 15.3). The fluid mixes with fluid formed in the third ventricle and from there it passes through the *cerebral aqueduct* (aqueduct of Sylvius) of the midbrain into the fourth ventricle. Fluid in the fourth ventricle passes into the subarachnoid space through two lateral apertures or foramina of Luschka. Some mammals have a third, medially located passageway from fourth ventricle to *subarachnoid space* (foramen of Magendie).

Recall that the brain and spinal cord are encased in bone (the skull and spinal canal, respectively) and are covered by a series of three membranes called the *meninges* (see Chapter 3). From outer to inner, these membranes are the dura, arachnoid, and pia (see Fig. 15.1). The subarachnoid space lies between the arachnoid and pia, and when the CSF exits the brain through the apertures (foramina) of the fourth ventricle, the CSF fills the subarachnoid space, spreading out over the entire outer surface of the brain and spinal cord (and even into the surface along the outer walls of penetrating blood vessels). Thus the entire CNS is essentially floating in a fluid-filled, membranous bag. As the CSF circulates up over the dorsal convexity of the brain, it is absorbed into the venous system near the midline.

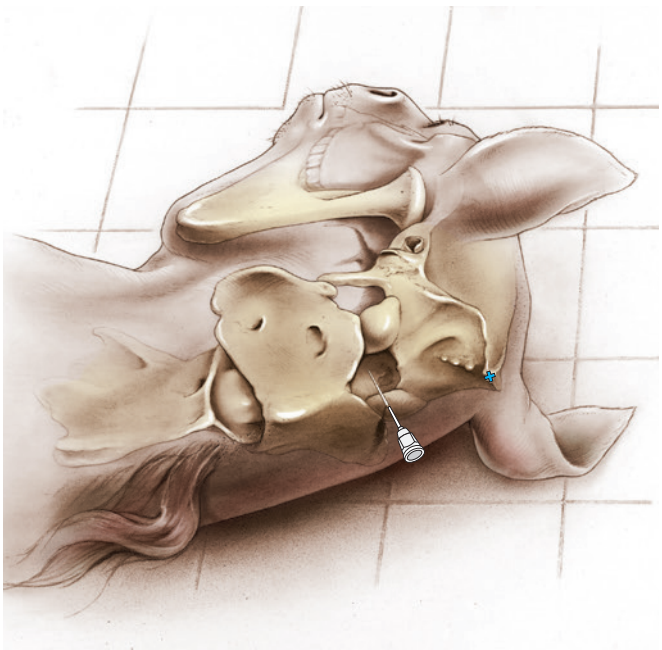
The pressure, cell count, and chemical constituents of CSF can be sampled by placing a styletted spinal needle into the subarachnoid space. Anatomically, the most convenient place to perform this varies with species. In humans, it is usually performed in the lumbar spinal column because the human spinal cord tapers to a cone (conus medullaris) near the first lumbar vertebra (humans have five lumbar vertebrae) as the dura and arachnoid continue down to around the second sacral vertebra. This provides a relatively large subarachnoid space (lumbar cistern) in the human midlumbar spinal column from which to sample. In most veterinary species, however, the conus medullaris extends to about the sixth or seventh lumbar vertebra, leaving only a small subarachnoid space in the spinal column. Therefore most veterinary spinal taps are performed by sampling from the subarachnoid space region accessed between the skull and the first cervical vertebra (atlas) in anesthetized animals (Fig. 15.4). In this location the subarachnoid space, formed as the arachnoid stretches from the caudal cerebellar surface to the dorsal surface of the medulla, is called the *cisterna magna* ("big cistern"; also called cerebellomedullary cistern) and is much deeper than other portions of the subarachnoid space (see Fig. 15.3). Spinal taps provide valuable information about such neuropathological lesions as intracranial space-occupying masses and inflammation.

Normal CSF is clear and translucent. Turbidity indicates increased cellularity, and a pink tint suggests the presence of blood. A common cause of increased CSF cell count is inflammation of the CNS. Neutrophils can be indicative of bacterial infection. Subarachnoid hemorrhage may be responsible for blood in the CSF. Increased CSF protein, in the absence of increased nucleated cell count, can result from conditions such as neurodegeneration or neoplasia. CSF can be cultured if bacterial infection is suspected.

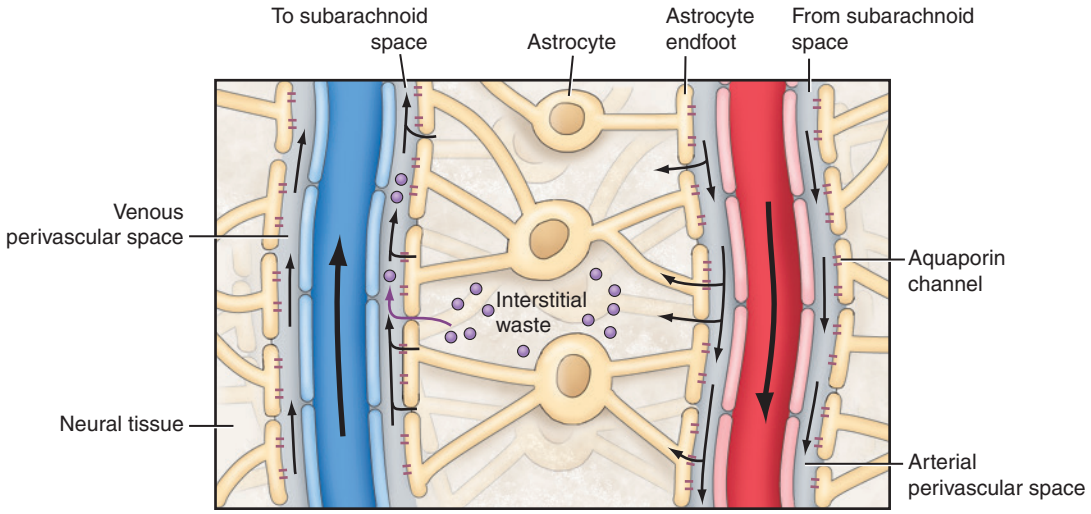
Vet-ebooks.com

Cerebrospinal Fluid Is Part of a Brain "Glymphatic" System That Helps Clear Waste Products

Most organs of the body possess a series of lymphatic vessels that help to clear products of cellular metabolism from the *interstitial fluid* (fluid in spaces between tissue cells) to venous blood. The brain and spinal cord represent the only organs that do not contain lymphatic vessels. However, it has recently been shown that the CSF is part of what may be considered a pseudolymphatic system that helps in the removal of interstitial waste products from the brain, most notably those of high molecular weight. After CSF enters the subarachnoid space, brain arterial pulsatility helps to drive CSF along perivascular arterial spaces that lie adjacent to arterial vessels as they dive from the subarachnoid space into the brain (Fig. 15.5). The perivascular spaces, which lie around the outside of brain venous vessels as well, are like donut-shaped tunnels running along the length of the vessel between the vessel and the end-feet of astrocytes that border these vessels along their length. The arrangement is somewhat analogous to an arm running through the sleeve of a long sleeve shirt with the arm representing the blood vessel, the sleeve representing the astrocytic end-feet, and the gap between arm and sleeve representing the perivascular space. Brain cells lie outside the sleeve. It is believed that convective forces drive the periarterial CSF out into the brain tissue where it interfaces with, and carries metabolic waste products from, the brain interstitial fluid into the venous perivascular spaces. The waste-carrying venous perivascular fluid can then empty into subarachnoid spaces. The flow of CSF out from the arterial perivascular spaces, and the subsequent movement of waste-carrying brain interstitial fluid into the venous perivascular spaces, is believed to be facilitated by aquaporin water channels on the astrocytic end-feet that circumscribe the perivascular spaces. This CSF-associated, lymphatic-like clearance of waste products from the brain tissue, combined with the putative role for the aquaporins of astrocytic end-feet in its operation, has led to this process being labeled the glial-associated lymphatic, or "glymphatic," system. There is some evidence that *glymphatic system* function decreases with age, perhaps contributing to the buildup of Alzheimer's disease-associated Aβ protein in the brain.

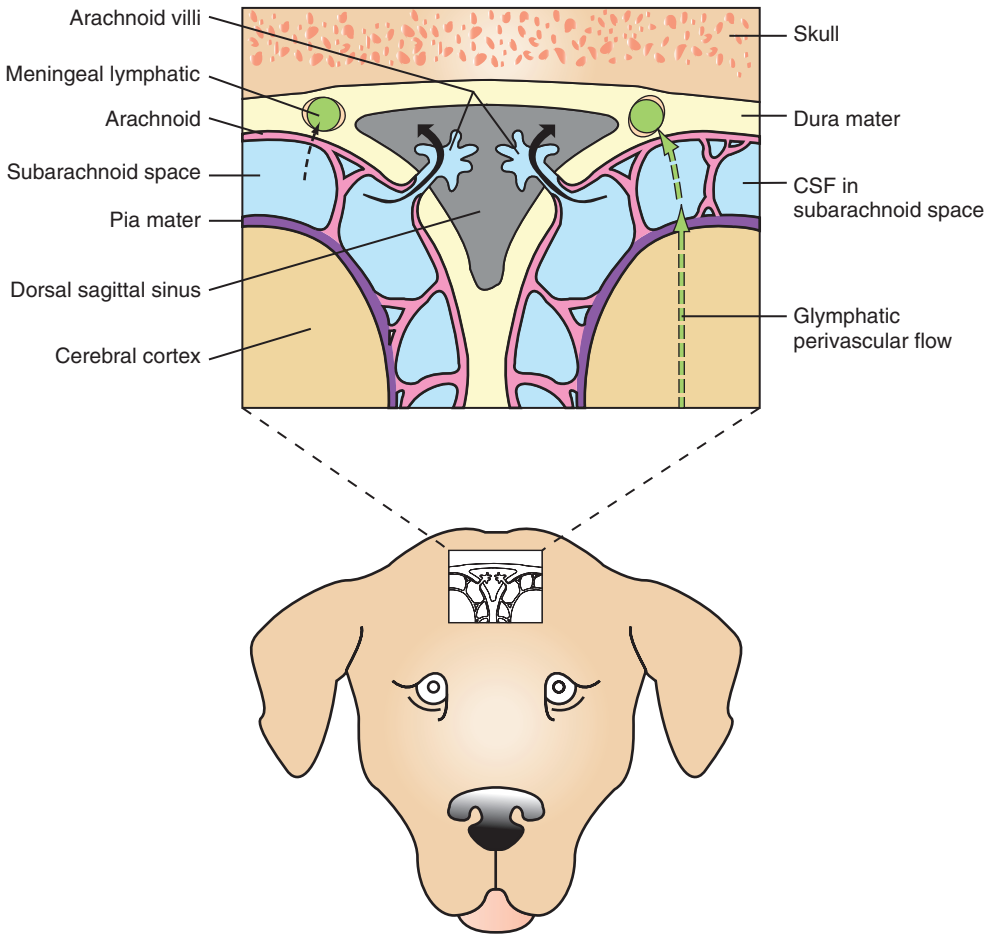


• **Fig. 15.4** Region of approach, between the skull and first cervical vertebra (atlas), for collection of a cerebrospinal fluid (CSF) sample from the cisterna magna (cerebellomedullary cistern) in a recumbent horse. (From De Lahunta A, Glass E. *Veterinary Anatomy and Clinical Neurology.* 3rd ed. Philadelphia: Saunders; 2008.)



• **Fig. 15.5** Principal elements of the glymphatic system. Efflux of periarterial cerebrospinal fluid (CSF), originating from the subarachnoid space, moves into the brain where it interfaces with, and carries waste material from, the brain interstitial fluid. This flow then enters the venous perivascular space where it is carried back to subarachnoid spaces. Aquaporin water channels on astrocytic end-feet facilitate the exchange of fluid between brain and the perivascular spaces. (Modified from Louveau A, Plog BA, Antila S, et al. Understanding the functions and relationships of the glymphatic system and meningeal lymphatics. *J Clin Invest.* 2017;127(9):3210-3219.)

Vet-ebooks.com



• **Fig. 15.6** Transverse (coronal) section through the dorsal midline of the brain showing the absorption of CSF into the dorsal sagittal sinus through arachnoid villi. The cerebrospinal fluid (CSF) is colored *light blue.* Meningeal lymphatic vessels are also shown along with the putative drainage of CSF (*dashed black arrow*), and of glymphatic flow (*dashed green arrows*), into these intracranial vessels. The small window on the dog's head shows the approximate dorsoventral position of the dorsal sagittal sinus. (Modified from Oliver JE, Lorenz MD. *Handbook of Veterinary Neurology.* 2nd ed. Philadelphia: Saunders; 1993.)

Cerebrospinal Fluid Is Absorbed Into the Venous System

CSF is absorbed into the venous system, a large proportion into dura-lined venous sinuses within the skull. The *dorsal sagittal sinus* lies between the dorsal surfaces of the cerebral hemispheres (see Fig. 15.3), and the *transverse sinuses* lie between the cerebral hemispheres and the cerebellum. In primates, the dorsal sagittal sinus plays the more significant role, whereas in quadrupeds, it is the transverse sinuses. Fluid is absorbed from the subarachnoid space into the dural sinuses through *arachnoid villi* (Fig. 15.6, and see Fig. 15.3). These are small, finger-like projections of the arachnoid membrane that poke through the walls of the sinus. Absorption appears to be pressure-dependent and is unidirectional; CSF can flow from subarachnoid space to venous sinus, but venous blood cannot normally move from the sinus back into the subarachnoid space. The movement of CSF into the venous sinus is sometimes referred to as a "bulk flow" because all constituents of the fluid, including waste products and other foreign matter (e.g., red blood cells), move into the sinus. Materials cross the cells of the arachnoid villi by both vesicular transport and by the formation and movement of giant, fluid-filled vacuoles. CSF can also cross from subarachoid space to sinus through intercellular spaces between cells of the villi. CSF is produced at a rate of about 1 mL/h in cats, about 3 mL/h in dogs, and about 20 mL/h in humans. The entire volume of CSF is replaced several times a day, approximately six times per day in species such as sheep and goat.

In addition to the absorption of CSF into the dural venous sinuses, it has also been shown that CSF can exit the cranial vault in the perineural subarachnoid spaces that extend around cranial nerves, ultimately draining into the deep cervical lymph nodes of the periphery. The major contribution to this route is along the olfactory nerves. However, a substantial amount of recent research has revealed the existence of an intracranial network of lymphatic vessels that run within the dura mater adjacent to venous sinuses, the meningeal vascular supply, and cranial nerves (see Fig. 15.6). These meningeal lymphatic vessels represent a direct connection between the intracranial space and the peripheral lymphatic vasculature, principally the deep cervical lymph nodes. Furthermore, it is believed that this intracranial lymphatic network may represent an important drainage route for the glymphatic system, although it is not yet clear how the perivenous return from this system gains access to the intracranial lymphatic vessels. The relationship between the glymphatic system and the intracranial lymphatic vessel network

Vet-ebooks.com

promises to be a fruitful area of investigation for neural-immune interactions, mechanisms of CNS disorders, and potential therapeutic approaches.

Hydrocephalus Is an Increased Volume of Cerebrospinal Fluid in the Skull

In normal animals, CSF pressure is significantly regulated by its absorption at the arachnoid villi because the rate of absorption can respond to changes in CSF pressure, whereas its formation is fairly constant and independent of changes in pressure. Therefore any obstruction of CSF absorption into the venous sinus causes CSF pressure to rise almost immediately. In some pathological conditions such as brain tumors or meningitis, CSF pressure can increase dramatically.

Hydrocephalus is defined as an increased CSF volume in the skull often associated with an increased ventricular volume and increased intracranial pressure. In theory, hydrocephalus could be caused by too much fluid production at the choroid plexuses, obstruction to its flow through the ventricular system or subarachnoid space, or impaired absorption at the arachnoid villi. In practice, overproduction seems rare, whereas obstruction to flow seems more common, particularly at such vulnerable sites as the narrow cerebral aqueduct (connecting the third and fourth ventricles) and the exits from the fourth ventricle. Such blockages in the ventricular system produce a *noncommunicating hydrocephalus* that results in a buildup of freshly produced CSF in portions of the ventricular system behind the blockage. This causes the ventricular regions inside the brain to expand at the expense of the surrounding brain tissue, and intracranial pressure rises (Fig. 15.7).

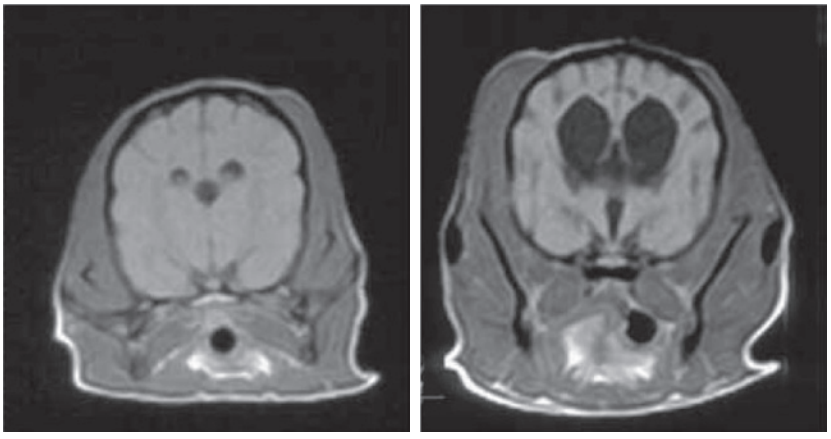
Impairment of absorption (a type of *communicating hydrocephalus*) can be secondary to meningitis or hemorrhage, presumably as a result of cellular debris that obstructs the transfer of CSF from subarachnoid space to venous sinus at the arachnoid villi. This can increase CSF volume in the subarachnoid space, which increases pressure on the outside surface of the brain and increases intracranial pressure.

The pathogenesis of many cases of hydrocephalus is not known. A common form of treatment in humans is surgical implantation of a tube that shunts CSF into the atria of the heart or into the peritoneal cavity, thus relieving episodes of increased CSF pressure and preventing further brain damage. Ventriculoperitoneal shunts have also been used in veterinary species. The carbonic anhydrase inhibitor acetazolamide or the diuretic furosemide has been used to decrease CSF production.

Permeability Barriers Exist Between Blood and Brain

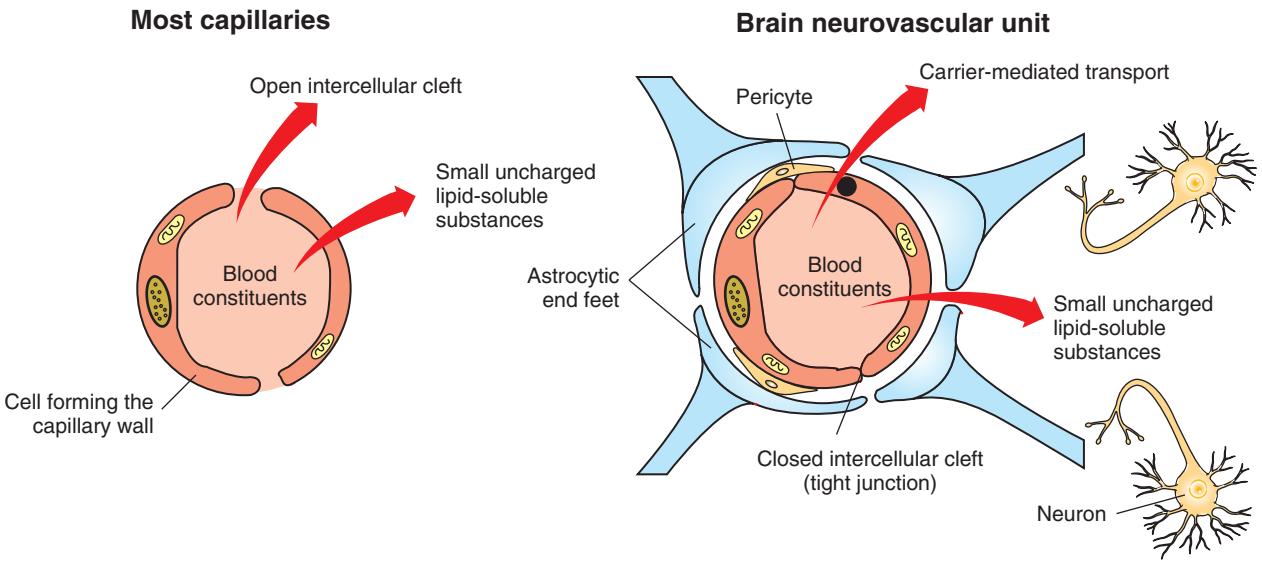
Many dyes, when injected into the blood, stain other tissues of the body but not the brain. This suggests that the brain's blood vessels have the ability to restrict certain substances from accessing brain tissue. This physiological property of CNS blood vessels is referred to as the *blood-brain barrier* (BBB). The BBB contributes to a stable environment for the neurons and glial cells of the CNS. Such protection from direct exposure to the blood supply is necessary because the composition of blood can significantly vary with factors such as diet, exercise, metabolic activity, illness, age, and exposure to environmental toxins. Many of the varying blood-borne nutrients, metabolites, and toxins are neuroactive and capable of affecting membrane receptors, transporters, or ion channels. In the absence of a BBB, these substances could result in unregulated and undesirable changes in neural activity and behavior.

In most capillaries, water-soluble compounds leave through open clefts between capillary endothelial cells, and exchange is relatively unrestricted. In brain capillaries, however, tight junctions block passage through intercellular clefts, and exchange of blood solutes is highly selective (Fig. 15.8). As a general rule, molecules that are small, uncharged, lipid-soluble, and unbound to plasma proteins (e.g., O_2, CO_2, ethanol, nicotine) can easily pass across the capillary endothelium of the BBB. Some molecules that do not fit this profile (e.g., glucose, some amino acids) are able to pass through the BBB by specific, carrier-mediated transport mechanisms. Brain capillaries have a greater number of mitochondria, which reflects the operation of such transporters. Certain degradative enzymes expressed by brain capillary endothelium (e.g., monoamine oxidase) provide a further restriction on substances that can pass the BBB.



• **Fig. 15.7** Magnetic resonance imaging (MRI) showing an example of the outcome of a noncommunicating hydrocephalus in a Maltese dog (*right*). The large dark ovals near the top of the image on the right are the greatly expanded lateral ventricles. An image from a normal dog is shown on the left for comparison. (From Nam JW, Choi CB, Woo DC, et al. Evaluation of hydrocephalic ventricular alterations in Maltese dogs using low field MRI. *Intern J Appl Res Vet Med.* 2011;9(1):58–67.)

Vet-ebooks.com



• **Fig. 15.8** Blood-brain barrier (BBB). Unlike most capillaries of the body, cells of brain capillary walls are joined by tight junctions that restrict the passage of material between the cells. These brain capillaries are part of a complex cellular microsystem, called the *neurovascular unit.* The unit also includes surrounding pericytes, glial astrocytic end-feet, and neurons. Materials leaving brain capillaries must pass through the cells forming the capillary wall. Substances that are not small, uncharged, and lipophilic must be carried across the cells by selective transport mechanisms. Interactions between the elements of the neurovascular unit are thought to contribute to development and maintenance of the tight-junction organization.

The capillary endothelial cells that characterize the BBB reside within a complex cellular microsystem that also includes surrounding pericytes, glial astrocytic end-feet, and neurons (see Fig. 15.8). This system is called the *neurovascular unit.* The interaction of these elements of the neurovascular unit, although not completely understood, appears important for the development, structural maintenance, and function of the specialized capillary endothelial BBB, which facilitates molecular homeostasis of the brain.

The integrity of the BBB can be compromised by certain pathological states such as ischemic stroke and traumatic brain injury. Reduced integrity of BBB tight junctions is a hallmark of neuroinflammatory disease. Interestingly, even inflammation of peripheral nerves appears to compromise BBB tight junctions. Diseases such as diabetes, as well as multiple sclerosis and Alzheimer's disease in humans, appear to be associated with decreased BBB integrity, although the causative nature is not clear in these cases.

Unfortunately, for many patients, the protection normally afforded to the CNS by the BBB often prevents many antibiotics and other drugs from reaching the brain, particularly drugs with low lipid solubility or drugs bound to plasma proteins. Compounding this problem is the fact that some drugs that do have the properties that would normally permit passive movement through the BBB capillary endothelium are actively moved from the endothelium back out into the blood by carrier-mediated transport that seems to have an affinity for some of these drugs. Attempts to circumvent this problem have focused on temporary disruption of the BBB, direct delivery to the brain, "hitching a ride" on particular BBB membrane transporters, and increasing the lipid solubility of drugs.

In some parts of the brain known as the *circumventricular organs,* which include the hypothalamus, the brain capillaries do not form tight junctions, and the BBB is apparently not effective. This is significant because these brain regions are involved in functions such as the control of serum osmolality and glucose levels, hormonal communication, feeding, drinking, and vomiting, and therefore they need to detect the levels of many serum solutes.

CLINICAL CORRELATIONS

Increased Intracranial Pressure

History

You examine a 9-year-old female boxer dog. The owner states that recently the dog has seemed drowsier than usual and had what you recognize to be a generalized tonic-clonic seizure the preceding night.

Clinical Examination

Physical examination of the dog reveals a hard, nodular mass of the mammary gland. Other deficits are referable to the nervous system and are characterized by apparent drowsiness and confusion, and by a deficit of the proprioceptive positioning reaction of the right front and right rear legs. Lateral radiographs of the chest reveal metastatic, neoplastic lesions in the lungs. The CSF pressure, as measured with a manometer through a needle placed in the cisterna magna, is 310 mm CSF. (The normal CSF pressure in dogs is less than 180 mm CSF.)

Comment

This is a typical case of a dog with a neoplasm of the mammary gland that has spread first to the lungs, which contain the first capillary bed filter encountered by tumor cells as they invade the venous system, and then to

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

the brain. As the tumor mass increases within the fixed encasement of the cranial vault, CSF and other fluid volumes are displaced. Some loss of myelin may compensate temporarily for the expanding intracranial mass, but eventually the expanding tumor, encased in the skull, causes an increase in intracranial pressure, which is reflected in an increased CSF pressure in the cisterna magna. When measuring this pressure, the dog is anesthetized, and a styletted spinal needle is placed in the cisterna magna. The stylet is removed, and a rigid glass or plastic tube (manometer) is attached using a right-angle, three-way valve. CSF rises up the manometer to a height proportional to intracranial pressure. Its height is measured off the millimeter graduations marked on the tube.

The proprioceptive deficits of the right front and right rear legs result from a focal, asymmetrical lesion of the left cerebral cortex. The seizure also resulted from this mass. With the mammary mass, metastatic lesions in the lungs, asymmetrical neurological signs, seizures, and elevated CSF pressure, it is reasonable to conclude that this dog has an intracranial neoplasm that probably spread from the mammary gland to the lungs and the brain. Magnetic resonance imaging (MRI) is warranted to further define the tumor in the brain.

Treatment

If this dog had a focal tumor within the CNS, other treatments would be possible. However, extensive treatment in this case would be futile because of the metastatic lesions. Palliative care (affording relief but not cure), steroids, and analgesics will make the animal more comfortable.

Seizures in a Foal

History

A 2-day-old Arabian colt from an unobserved foaling displays lethargy and inability to rise, and he started to go into seizure in the last hour. The foal stood, but it took longer than normal. The foal has been nursing, but the mare does not appear to have much milk; this is the mare's first foal. The foal seemed less active than normal foals and became more lethargic as the day progressed. Finally, it seemed as though he would not stand, and he had a seizure while traveling to the clinic.

Clinical Examination

The foal has a fever, and pulse and respiration are increased. Mucous membranes are darker red than normal, the membranes are dry (dehydration), and the capillary refill time is prolonged (poor perfusion). Auscultation (listening for sounds within the body) reveals harsh lung sounds and crackles. The umbilicus (navel) is thickened and wet. There is petechiation (petechial hemorrhages) inside the ears and in the sclera. There appears to be signs of uveitis (see Chapter 14) in the eyes. The foal is not responsive to manipulations and thrashes, although not in seizure, while being examined. The foal also lacks a suckle reflex.

Comment

Although many possibilities exist for the cause of seizures in this foal, the two most likely reasons are low glucose level (hypoglycemia) and infection (meningitis). The blood glucose level is high in this foal, so meningitis (inflammation of the meninges) with septicemia is most likely. This foal probably has become septic (blood-borne infection) based on the history and clinical signs, which support that he likely did not receive enough colostrum and may not be receiving enough milk from the mare. Without adequate colostrum and nutrition, the foal's immune system is more susceptible to infection. With the thickened umbilicus, harsh lung sounds, fever, petechiation, uveitis, and seizures, the signs are consistent with septicemia, which is manifesting in different regions of the body. The umbilicus could be infected (omphalophlebitis), the lung sounds are consistent with infection (pneumonia), and seizures are consistent with meningitis.

Complete blood count, chemistry, blood gases, and blood culture are warranted to determine the overall status of the foal. In many cases, these tests will be sufficient to make a diagnosis and determine treatment. In some cases, to make a definitive diagnosis, a CSF tap is warranted. An atlanto-occipital (A/O) CSF tap (which samples from the cisterna magna) is best because it is closest to the site of the lesion, compared with a lumbosacral tap. The tap also allows a culture to be submitted so that the foal can be treated with the most efficacious antibiotics. When performing the CSF tap, the foal can be sedated with diazepam (Valium). The CSF is submitted for protein, glucose, cytology, and culture. Typically, the protein level is high with meningitis, and cytology shows an increased number of leukocytes (neutrophils). There is always the potential for a false-negative result with the culture.

Treatment

Prognosis for a septic foal with seizures is poor, with many factors to consider. With regard to the meningitis, treatment consists of antibiotics, antiinflammatory agents, and anticonvulsants as needed. Seizures can cause hypoxia to the affected area, which can result in permanent damage. Besides meningitis, other problems include the umbilical infection, respiratory infection, and uveitis. With septicemia, other organs often become infected (joints, gastrointestinal tract, renal system). Additional concerns are potential renal insult caused by dehydration or complications associated with some antibiotics. Supportive care is another consideration. Managing a recumbent foal is challenging, not only for the reasons already listed, but also because of other factors including additional infection, aspiration, and nutritional support.

Practice Questions

1. Obstruction of cerebrospinal fluid (CSF) flow at the cerebral aqueduct (aqueduct of Sylvius) would lead to dilation (enlargement) of the:
- **a.** Lateral ventricles
- **b.** Fourth ventricle
- **c.** Central canal of the spinal cord
- **d.** Subarachnoid space
- **e.** Conus medullaris

2. CSF is principally formed at the:
- **a.** Arachnoid villi
- **b.** Aqueduct of Sylvius
- **c.** Choroid plexuses
- **d.** Subarachnoid space
- **e.** Dorsal sagittal sinus

3. You are performing a spinal tap on an anesthetized horse and measuring CSF pressure. Cellular debris has obstructed the arachnoid villi following meningitis. What would you expect regarding CSF pressure?
- **a.** Pressure would be higher than normal
- **b.** Pressure would be lower than normal
- **c.** Pressure would be normal

4. For many veterinary species, diagnostic sampling of CSF is often performed by placing a sampling needle in the:
- **a.** Lateral ventricles
- **b.** Dorsal sagittal sinus
- **c.** Third ventricle
- **d.** Cerebral aqueduct of Sylvius
- **e.** Cisterna magna

Vet-ebooks.com

5. Which *two* of the following are *false* regarding the blood-brain barrier (BBB)?
- **a.** The BBB is very effective at the circumventricular organs of the brain.
- **b.** Astrocytic end-feet are part of the neurovascular unit.
- **c.** Brain capillaries generally have a high number of endothelial tight junctions.
- **d.** Many dyes injected into the blood can typically penetrate most tissues of the body but usually not the brain.
- **e.** Small, uncharged, lipid-soluble molecules do not generally pass through the BBB.

6. Which one of the following statements regarding the glymphatic system is *false*?
- **a.** Venous perivascular spaces are a major component.
- **b.** It plays an important role in clearance of interstitial fluid waste products from the brain.
- **c.** Brain arterial pulsatility is a driving force for the system.
- **d.** Movement of water through aquaporin water channels on microglial end-feet is an important functional component.
- **e.** Intracranial lymphatic vessels may represent an important drainage route for the system.

Bibliography

Abbott NJ, Rönnbäck L, Hansson E. Astrocyte-endothelial interactions at the blood-brain barrier. *Nat Rev Neurosci*. 2006;7(1):41–53.

Benarroch EE. Choroid plexus–CSF system: recent developments and clinical correlations. *Neurology*. 2016;86:286–296.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Di Terlizzi R, Platt SR. The function, composition and analysis of cerebrospinal fluid in companion animals: part I—function and composition. *Vet J*. 2006;172(3):422–431.

Di Terlizzi R, Platt SR. The function, composition and analysis of cerebrospinal fluid in companion animals: part II—analysis. *Vet J*. 2009;180(1):15–32.

Fletcher TF. Spinal cord and meninges. In: Evans HE, de Lahunta A, eds. *Miller's Anatomy of the Dog*. 4th ed. Philadelphia: Saunders; 2013.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Hawkins BT, Davis TP. The blood-brain barrier/neurovascular unit in health and disease. *Pharmacol Rev*. 2005;57(2):173–185.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2013.

Louveau A, Plog BA, Antila S, et al. Understanding the functions and relationships of the glymphatic system and meningeal lymphatics. *J Clin Invest*. 2017;127(9):3210–3219.

Mollanji R, Papaiconomou C, Boulton M, et al. Comparison of cerebrospinal fluid transport in fetal and adult sheep. *Am J Physiol Regul Integr Comp Physiol*. 2001;281(4):R1215–R1223.

Paolinelli R, Corada M, Orsenigo F, Dejanaa E. The molecular basis of the blood-brain barrier differentiation and maintenance. Is it still a mystery? *Pharmacol Res*. 2011;63(3):165–171.

Plog BA, Nedergaard M. The glymphatic system in CNS health and disease: past, present and future. *Annu Rev Pathol*. 2018;13:379–394.

Pollay M. The function and structure of the cerebrospinal fluid outflow system. *Cerebrospinal Fluid Res*. 2010;7:9.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Rossmeisl JH Jr. Cerebrospinal fluid collection, analysis, and myelography. In: Ettinger SJ, Feldman EC, Cote E, eds. *Textbook of Veterinary Internal Medicine*. 8th ed. St. Louis: Elsevier; 2017.

Thomas WB. Hydrocephalus in dogs and cats. *Vet Clin North Am Small Anim Pract*. 2010;40(1):143–159.

Vet-ebooks.com

16

The Electroencephalogram and Sensory-Evoked Potentials

BRADLEY G. KLEIN
EDITED BY JOHN H. ROSSMEISL, JR.

KEY POINTS

- All areas of the cerebral cortex share common histological features.
- The electroencephalogram (EEG) is a common clinical tool.
- The collective behavior of cortical neurons can be studied noninvasively through the use of macroelectrodes on the scalp.
- Stimulation of sensory pathways can be recorded as evoked potentials.

When many excitable cells are present in a living tissue, their electrical behavior can be detected by macroelectrodes placed on the body at a distance from these cells. Several clinically important electrophysiological diagnostic procedures are based on this concept.

Underlying these procedures is a theory called *volume conduction.* This theory describes the spread of ionic currents within the extracellular fluid from a group of neurons or muscle cells to more distant points in the body, such as the skin. These ionic currents can be measured from the skin. Their waveforms are characteristic of the tissues from which they arise. The best-known electrophysiological recording is the electrocardiogram from heart muscle (see Chapter 20). The electromyogram from skeletal muscle (see Chapter 6) and electroretinogram (see Chapter 14) are other examples.

This chapter introduces two additional clinical electrophysiological tools: the *electroencephalogram (EEG)* and *sensory-evoked potentials,* particularly brainstem auditory-evoked responses (BAERs). These tools represent two general types of clinical electrophysiological recordings. The first is a record of the spontaneous activity of tissue. The second is a record of potentials that are artificially evoked by electrical or magnetic stimulation of tissue or by activation of sensory receptor organs. Before discussing the EEG and sensory-evoked potentials, it is necessary to understand more about the histology and electrophysiology of the cerebral cortex.

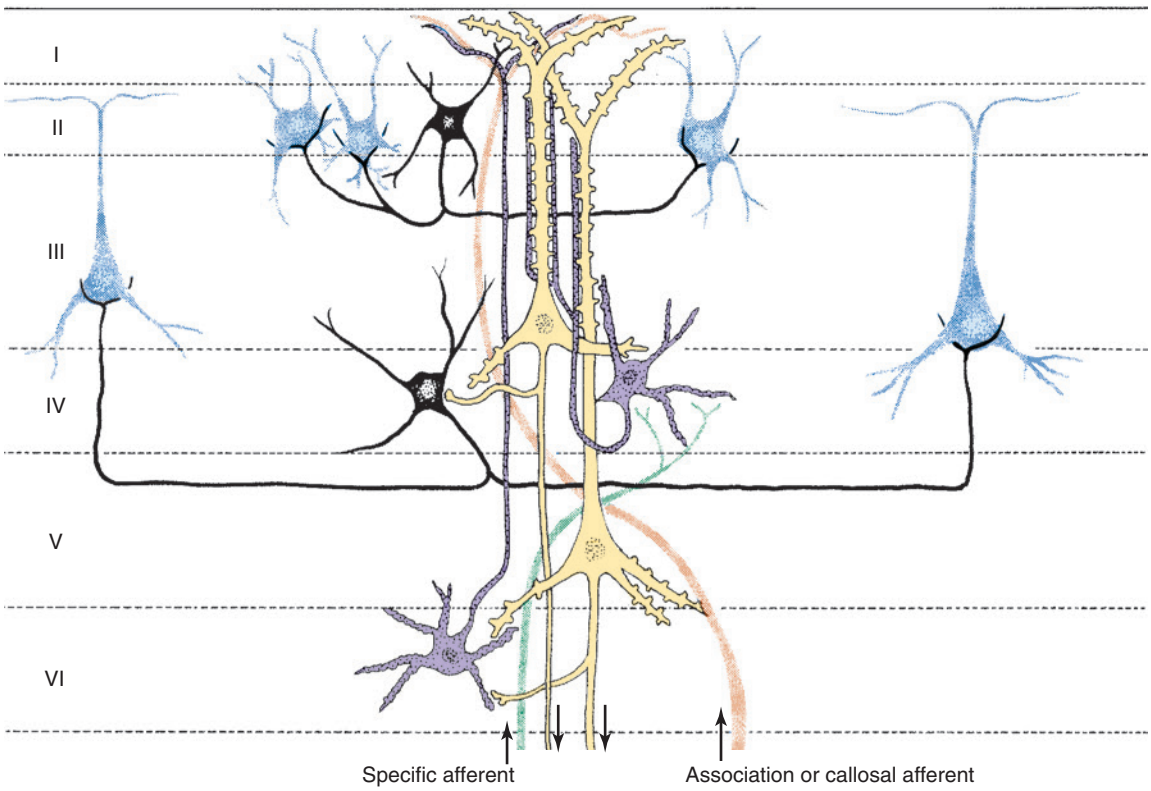
All Areas of the Cerebral Cortex Share Common Histological Features

Different regions of the cerebral cortex have different functions. For example, the motor cortices (see Chapter 10) project to the brainstem and spinal cord to initiate skilled, learned, conscious movement. The occipital cortex processes visual information received from the retina of the eye (see Chapter 14). The temporal cortex processes similar information from the ear (see Chapter 17). Even though different cortical regions have different functions, they have an underlying histological similarity. Therefore cortical synaptic processing of information shares common features across regions, but differences in the origin of input signals and the destination of output signals contribute significantly to functional differences among regions. However, cerebral cortical cells can also work collectively over vast regions of the brain in such normal states as sleep and wakefulness and in such disease states as coma and seizures.

The cerebral cortex contains several different cell types, but most belong to two major classes: pyramidal cells and stellate cells (Fig. 16.1). These cells are arranged in six layers (I–VI). The *pyramidal cells,* so named because their cell bodies are pyramid-shaped, have dendrites projecting up toward the pial surface of the cortex, usually reaching and branching within layer I. These cells also have basal dendrites that extend horizontally from the cell body. Pyramidal cells are projection neurons, with their axons leaving their cortical region of origin and projecting to other parts of the central nervous system (CNS) or to a different region of cerebral cortex. Pyramidal cells are generally excitatory at their axon's synapse. *Stellate cells,* so named because most have a starlike appearance, are local-circuit interneurons within the cortex, and can be either excitatory or inhibitory. The majority of subcortical information arrives at the cortex through a massive input from the thalamic nuclei, most of which is targeted to layer IV. Input from some portions of the thalamus, as well as from other regions of the cerebral cortex, have a more diffuse termination across cortical layers. Information arriving from cortical afferents is processed by local cortical circuitry, and pyramidal cells then carry the processed information to other CNS regions.

As with other regions of the brain, neurons in the cerebral cortex are intricately anatomically and functionally associated with glial cells. Three types of glial cells are present in the cortex: *astrocytes, oligodendrocytes,* and *microglia.* They do not develop action potentials, but as noted in Chapter 3, they may indirectly monitor neuronal electrical activity and modulate the effectiveness of neural communication. Glia also take up excess potassium ions, neurotransmitter, and toxins from the extracellular space and play a role in immune function. In addition, they help guide the course of developing neurons, play a role in synapse formation,

Vet-ebooks.com



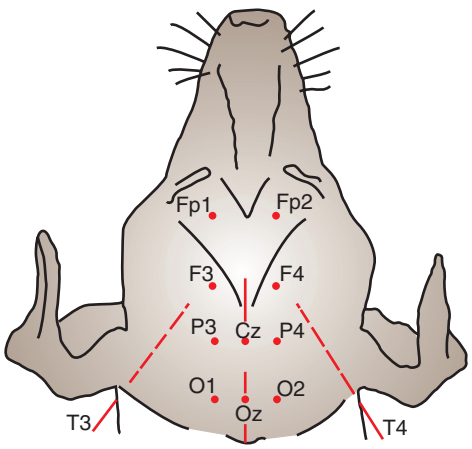
• **Fig. 16.1** The principal neuron types and their interconnections have a basic similarity in the various regions of the cerebral cortex. Note that the two large pyramidal cells *(yellow)* in layers III and V receive multiple synaptic contacts from the star-shaped interneuron (stellate cell, *violet*) in layer IV. Basket cell *(black)* inhibition is directed to the somata of cortical neurons. Major input to the cortex derives from specific thalamic nuclei (*specific afferent*) and is directed mostly to layer IV; association and callosal inputs (*association or callosal afferent*) have a more widespread termination pattern among the cortical layers. (From Kandel ER, Schwartz JH, editors: *Principles of neural science,* ed 2, New York, 1985, Elsevier Science & Technology.)

and stabilize the position of neurons, thus the origin of the term glia ("glue").

The Electroencephalogram Is a Common Clinical Tool

It has been known since the 1930s that a fluctuating electrical voltage reflecting brain activity could be recorded from macroelectrodes on the scalp (Fig. 16.2). Such a recording is known as an EEG. The frequency of the waveform recorded varies inversely with its amplitude. Both frequency and amplitude change with changes in levels of arousal (Fig. 16.3). An alert animal has a fairly high-frequency, fairly low-amplitude EEG, whereas a more relaxed animal has a slower-frequency, higher-amplitude EEG. A sleeping animal usually begins sleep by exhibiting a slow-wave, high-amplitude EEG. Paradoxically, there are periods of high-frequency, low-amplitude EEG during the sleep cycle. Four frequency ranges have been given names: *alpha* (8–13 Hz), *beta* (13–30 Hz), *delta* (0.5–4 Hz), and *theta* (4–7 Hz).

This technique has been applied clinically since the early 1960s. Abnormal EEG activity has been associated empirically with several brain diseases. In human neurology, EEGs have been used to classify the epilepsies, to localize lesions, and to help define "brain death." EEGs have not been as widely used in veterinary medicine but still have clinical utility in veterinary neurology, especially for the



• **Fig. 16.2** Points of scalp electrode attachment (electrode montage) for recording the electroencephalogram (EEG). Recordings are typically made from multiple paired combinations of subdermal electrodes (dots). Due to the large masticatory muscles present in dogs, recording from the temporal loci (T3 and T4) requires use of longer insulated monopolar needle electrodes that are passed through the muscles to the skull (dashed lines). (From Pellegrino FC, Sica REP: Canine electroencephalographic recording technique: findings in normal and epileptic dogs, *Clin Neurophysiol* 115:477–487, 2004.)

Vet-ebooks.com

diagnosis of seizure disorders (Fig. 16.4). The development of ambulatory EEG microsystems is an evolving area of research in veterinary medicine. When implanted into the patient, these miniaturized devices are capable of continuously recording EEG activity for months. Prototypes of this technology have been shown to be effective and well tolerated in dogs with epilepsy. Newer generations of these instruments are intended to be capable of wireless transcranial recording and transmission of EEG data using web-based platforms. Ultimately, the goal of these devices is to produce dynamic and real-time biofeedback circuits that would allow for an intervention to occur before the actual clinical onset of a seizure. When impending seizure activity is detected in the EEG of the patient through the use of predictive computer-generated algorithms, an alert from the unit is sent to a mobile device carried by the animal's caretaker that signals that a specific medical intervention needs to be administered. Alternatively, the impending seizure activity could automatically activate another implanted device, such as a parenteral drug reservoir, that would deliver a treatment to abolish the seizure.

We now discuss where these scalp recordings originate and how they relate to brain function.

The Collective Behavior of Cortical Neurons Can Be Studied Noninvasively Through the Use of Macroelectrodes on the Scalp

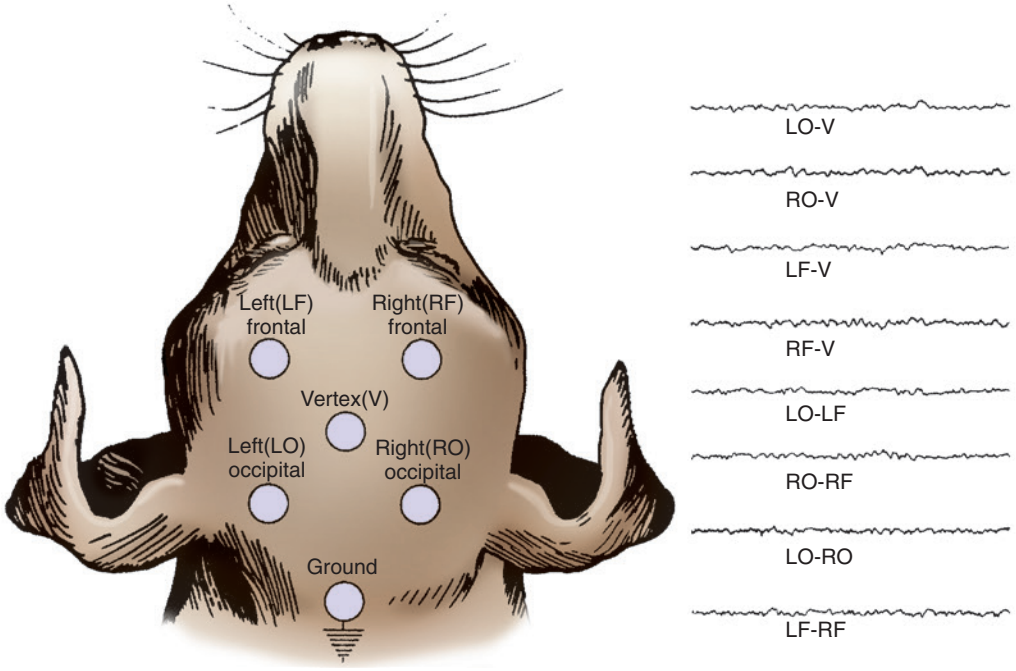
The scalp EEG records a fluctuating voltage resulting from changes in postsynaptic potentials in thousands of neurons under the electrode. Each change in voltage has a polarity.

By convention, changes in voltage measured by extracellular electrodes, such as those on the scalp, have a standard direction of deflection when observed on the recorder device. When the voltage change is in a positive direction, the recorded deflection is "down"; when in a negative direction, the deflection is "up" (Fig. 16.5). The polarity of the voltage change at the scalp depends on the nature and location of the postsynaptic potential change. If an *excitatory postsynaptic potential* (EPSP) occurs in a deep cortical layer, positive ions (e.g., Na^+) enter the cell there, leaving the extracellular fluid at that location relatively negative. Through principles of volume conduction, this leaves the extracellular fluid near the cortical surface positive with respect to the deeper, negatively charged region of extracellular fluid (see Fig. 16.5; for simplicity, only one cell is indicated). This results in a positive voltage change being recorded at the scalp macroelectrode near the cortical surface. Based on the same principles, if the EPSP occurs near the cortical surface (see Fig. 16.5), the voltage recorded from the scalp is negative. The polarity of these changes would be reversed for *inhibitory postsynaptic potentials* (IPSPs).

Voltage changes recorded from the scalp are the result of the summated extracellular voltage changes caused by the postsynaptic potentials of a large number of active cortical neurons, primarily pyramidal cells, because the voltage change from any one neuron is too small to record. Action potentials contribute little to the EEG with scalp electrodes.

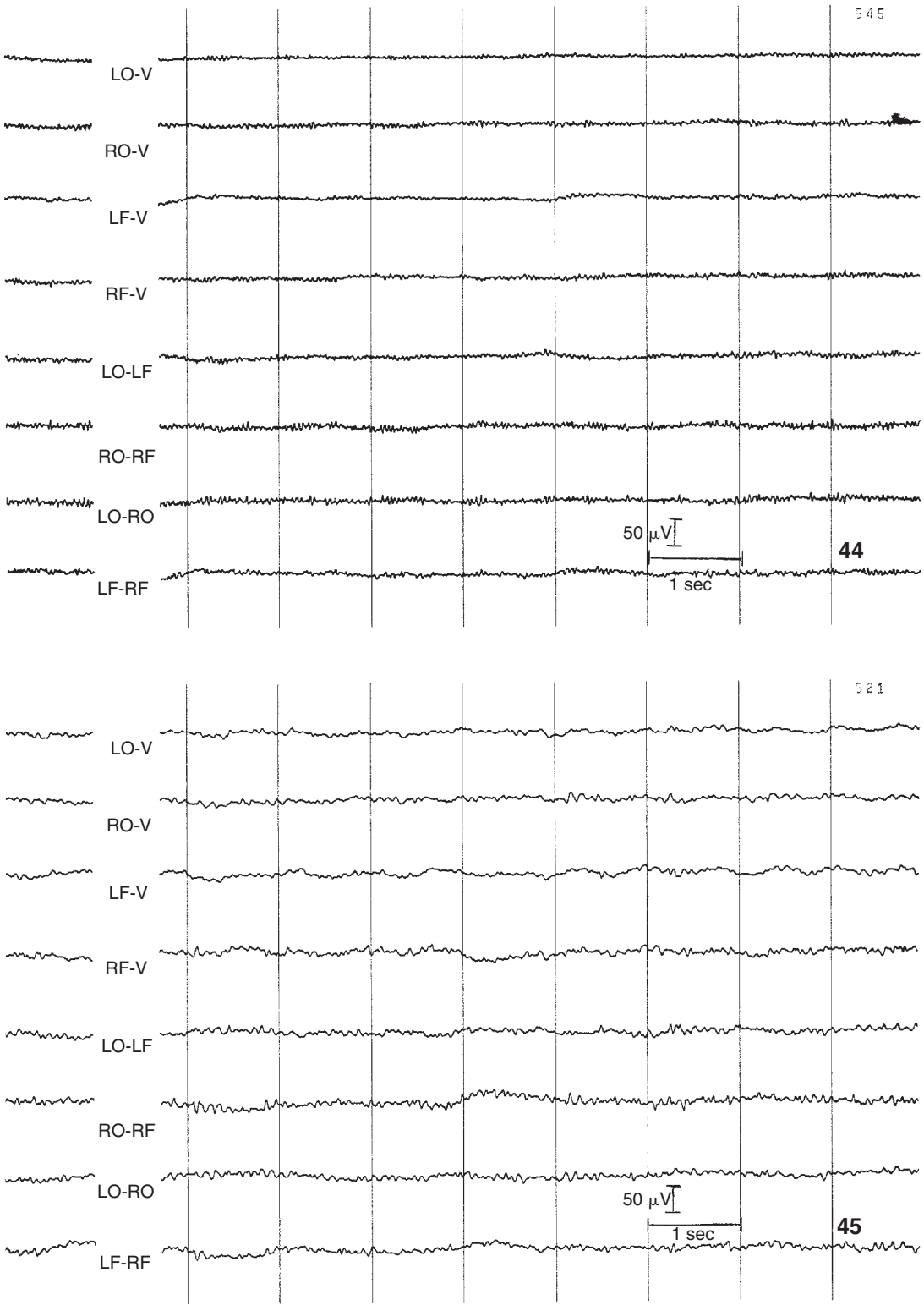
The *amplitude* (height) of voltage fluctuations in the scalp-recorded EEG is a function of how many cortical cells are changing their postsynaptic potentials in the same direction at the same time. Because a high-amplitude voltage change would result from a large number of neurons firing synchronously, a high-amplitude, slow-frequency EEG is said to be a *synchronized EEG.* When neurons are firing more or less at random, a low-amplitude, high-frequency EEG results, said to be a *desynchronized EEG.*

The *frequency* with which EEG voltage changes occur is largely determined by the *reticular activating system.* As noted in Chapter 10, ascending projections of the reticular formation play an



• **Fig. 16.3** EEG recorded from various combinations of lead points using an older, simpler electrode configuration (the Redding configuration). It shows the difference in frequency and amplitude between an alert animal *(44),* a relaxed animal *(45),* and an animal in light sleep *(80).* Note the decrease in frequency and increase in amplitude in the progression from alert to relaxed to light sleep. (From Oliver JE, Hoerlein BF, Mayhew IG, editors: *Veterinary neurology,* Philadelphia, 1987, Saunders.)

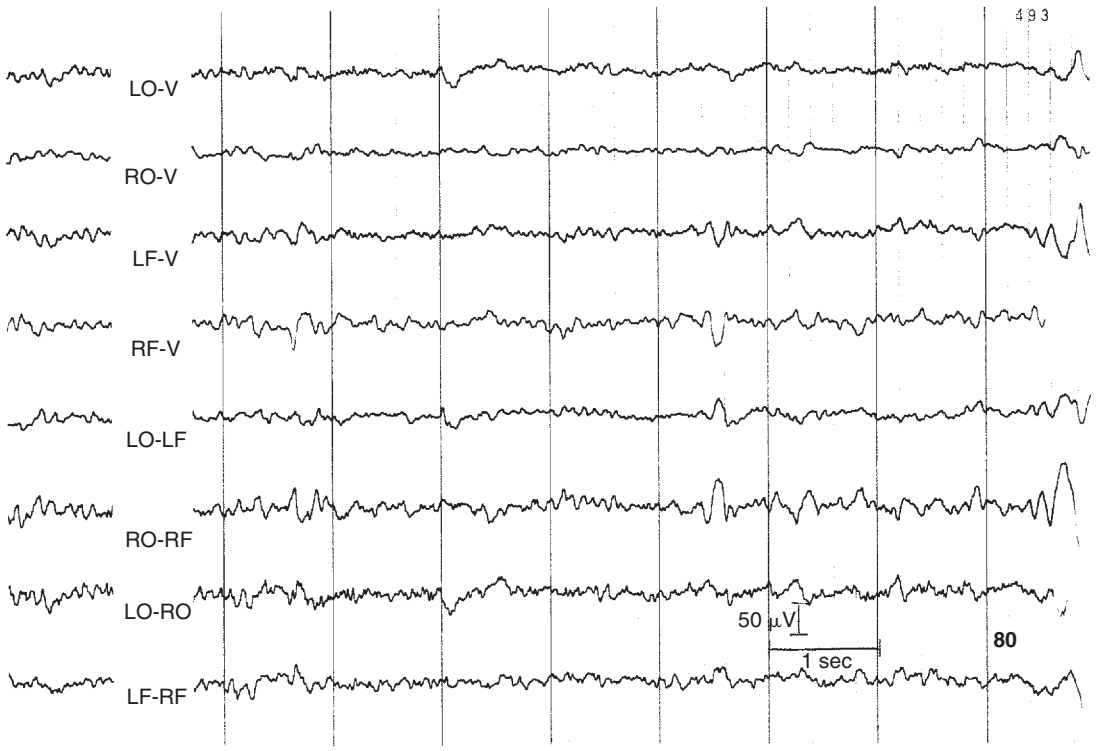
Vet-ebooks.com



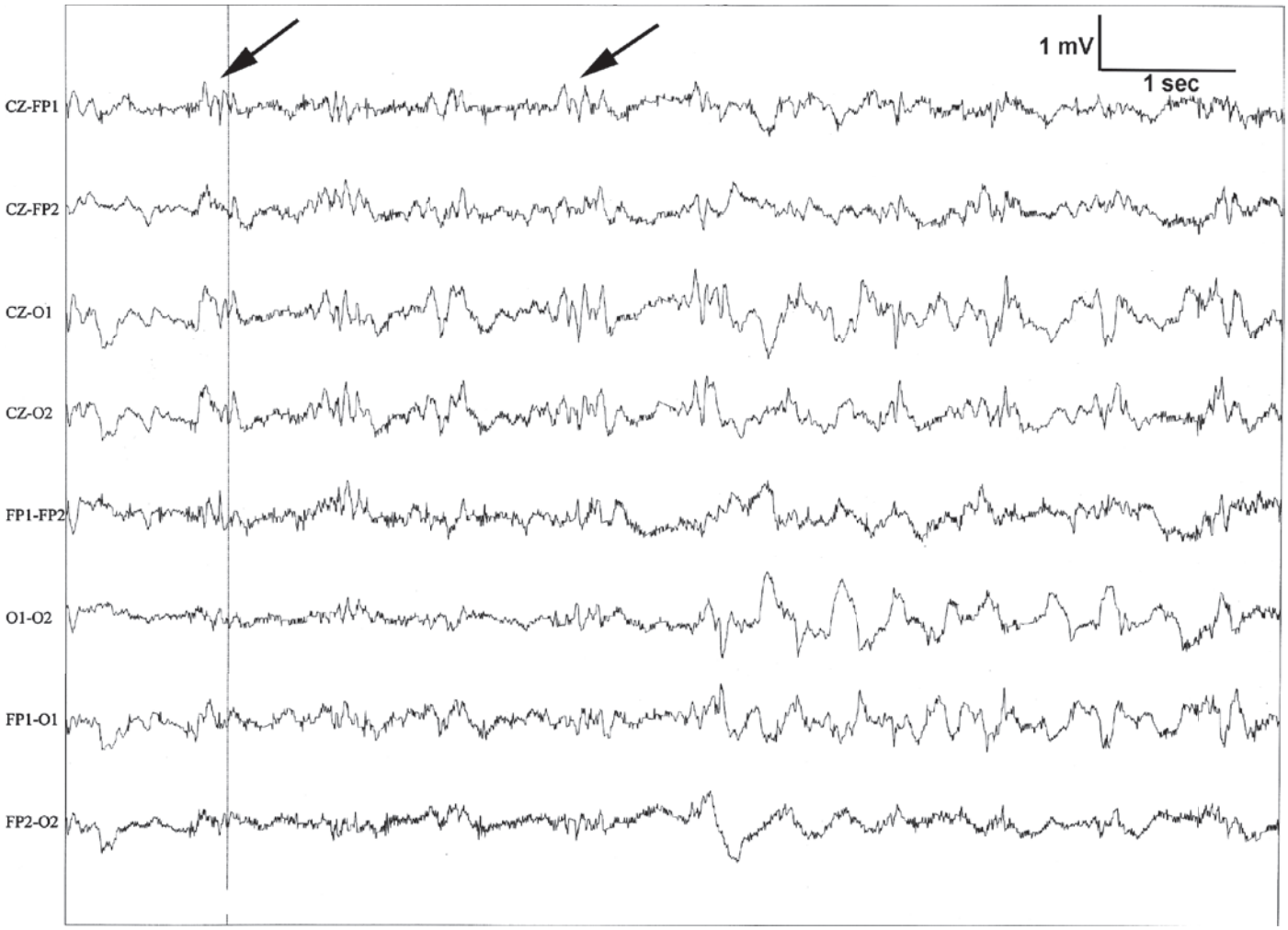
• **Fig. 16.3, cont'd**

Continued

Vet-ebooks.com

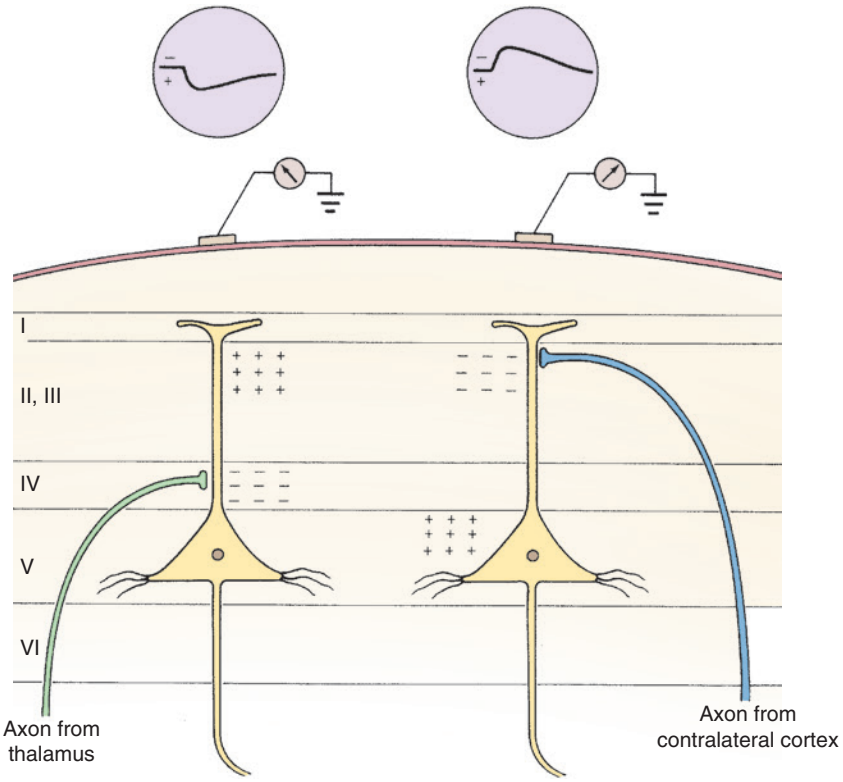


• **Fig. 16.3, cont'd**



• **Fig. 16.4** EEG recorded from a dog using the electrode montage depicted in Fig. 16.2. This tracing demonstrates generalized spike and wave discharges (arrows), consistent with epileptic seizure activity. (Image courtesy of Dr. John H. Rossmeisl, Jr., Department of Small Animal Clinical Sciences, Virginia-Maryland College of Veterinary Medicine, Virginia Tech.)

Vet-ebooks.com



• **Fig. 16.5** Scalp recordings and underlying synaptic mechanisms. **Left,** Potential recorded from a scalp electrode after activation of thalamic inputs. The terminals of thalamocortical neurons make excitatory connections on cortical neurons predominantly in layer IV. Thus the site of inward current flow (sink) in layer IV leaves the extracellular fluid at that location relatively negative and the extracellular fluid near the cortical surface relatively positive. Because the recording electrode is located on the scalp, near the cortical surface, it records a positive potential. By convention, a positive extracellularly recorded potential is, unlike intracellular recordings, a downward deflection. **Right,** Potential recorded from an excitatory afferent originating in the contralateral cortex. The axon has traversed the corpus callosum and terminates in a superficial cortical layer. A negative potential (upward deflection) is recorded because the electrode is closer to the site of inward current flow, which leaves the extracellular fluid near the cortical surface relatively negative. (From Kandel ER, Schwartz JH, editors: *Principles of neural science,* ed 2, New York, 1985, Elsevier Science & Technology.)

important role in modulating consciousness, arousal, and attention. Many of these projections synapse primarily in the thalamus, hypothalamus, or directly in the cerebral cortex in a diffuse fashion. Diffuse cortical projections from portions of the thalamus (intralaminar nuclei) and hypothalamus (lateral hypothalamus) along with diffuse, direct cortical projections from the reticular formation, likely regulate consciousness and arousal. Neurons that project to the cortex from specific sensory relay nuclei of the thalamus and receive reticular formation input probably influence attention. The term reticular activating system collectively refers to these ascending reticular formation neurons and the neurons that relay their activity to the cortex, both of which affect consciousness, arousal, and attention.

Stimulation of Sensory Pathways Can Be Recorded as Evoked Potentials

Large areas of the brain and the spinal cord are not reflected in the EEG. Other clinical electrophysiological recordings can help examine the function of these areas.

Synaptic activity in a sensory pathway can be recorded from the scalp by a computerized technique that averages out the more random background EEG activity, allowing clearer detection of the average electrical response to multiple stimulations of a sensory system. Such signals are called *sensory-evoked potentials.*

Because scalp macroelectrodes can more easily record the EEG electrical signals generated from the closer cerebral cortical cells, these higher-voltage signals must be eliminated; otherwise, they would mask the sensory-evoked potentials. Because the background EEG signals are relatively random, a computer can average them together and functionally erase them from the recording while at the same time averaging the nonrandom sensory-evoked potential signals recorded from multiple stimulations of a sensory pathway. In this way, scalp macroelectrodes can be used to record electrical events generated in brain locations distant from the recording electrode. For this reason, these sensory-evoked potentials are often called *far-field potentials.*

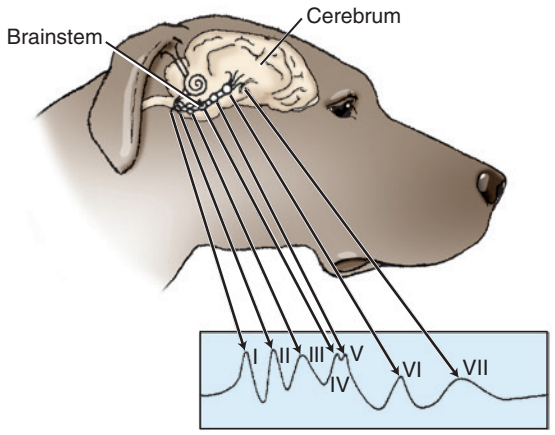
One such sensory-evoked potential is the BAER. This clinical electrophysiological procedure, in which the electrode placement is configured to include early activity from the caudal brainstem, records the electrical events for 10 msec after a click stimulus to the ear (Fig. 16.6). Usually, seven waves are recorded that are thought to be generated by neural activity in components of the auditory pathway from the auditory nerve through the auditory radiations that leave the medial geniculate nucleus of the thalamus (Fig. 16.7). Recordings longer than 10 msec are sometimes taken, and these are referred to as *middle latency recordings.* These later waves reflect cortical response to auditory stimulation. BAER is used in animals and humans to assess brainstem function in general, and auditory function in particular. Other sensory-evoked potentials can be recorded from the visual system, the somatosensory system, and other sensory modalities.

BAER is also often used with simultaneous EEG recording in the confirmation of brain death. A flat EEG, a very crude indicator of brain death, combined with a viable BAER can suggest that the functional deficit may not be irreversible.

Vet-ebooks.com



• **Fig. 16.6** Canine instrumented for a BAER examination using tubal ear inserts *(thick red and blue tubes)* and a vertex-mastoid skin electrode montage *(thin wires).* (Image courtesy of Dr. John H. Rossmeisl, Jr., Department of Small Animal Clinical Sciences, Virginia-Maryland College of Veterinary Medicine, Virginia Tech.)



• **Fig. 16.7** Brainstem auditory-evoked response (BAER): idealized diagram of waveforms recorded by signal averaging. The putative neural generators responsible for the sequential auditory waves are grouped as follows: Wave I reflects the cochlea, spiral ganglia, and cranial nerve VIII. Wave II reflects the cochlear nuclei. Wave III reflects the nucleus of the trapezoid body. Waves IV and V reflect the lateral lemniscus and lemniscal nuclei, and caudal colliculus, respectively (these two waves are frequently combined to form one wave). Wave VI reflects the medial geniculate body. Wave VII reflects auditory radiations. Positive is upward. (From Oliver JE, Hoerlein BF, Mayhew IG, editors: *Veterinary neurology,* Philadelphia, 1987, Saunders.)

CLINICAL CORRELATIONS

Brain Tumor

History

You examine a 10-year-old Boston terrier. The owner states that, during the past 3 weeks, the dog has had seizures of increasing frequency, characterized by turning the head to the right, rigidity of the right front and right hind legs, collapsing to the ground, and urination. More recently, the dog has seemed weak, drowsy, and confused. He tends to walk in circles and seems weak on the right front leg.

Clinical Examination

Important physical examination deficits are referable to the nervous system. The dog seems weak, drowsy, confused, and unsteady in gait. He tends to walk in counterclockwise circles. Cranial and spinal segmental reflexes are within normal limits. Proprioceptive positioning reaction is abnormal in the right front leg and normal in the other three legs (see Chapters 7 and 10). An EEG reveals that the dominant frequency is slower and the amplitude higher over the left parietal cortex than over the rest of the brain. Occasional bursts of electrical spiking activity can also be seen from the area of the left parietal cortex. A magnetic resonance imaging (MRI) scan of the brain is warranted to determine the presence and nature of a tumor or other focal pathological process suspected from the EEG patterns. MRI provides the best imaging of intracranial lesions to determine whether this is a primary tumor (originating from the brain tissue) or a secondary tumor (originating from other tissue, e.g., hemangiosarcoma, lymphosarcoma). On MRI, primary tumors usually appear as solitary masses in the brain, whereas secondary tumors are frequently multifocal.

Comment

This is an old dog, with a recent history of progressive, asymmetrical brain disease. The history suggests a focal intracranial lesion, perhaps a brain tumor. A focal lesion is further confirmed by the EEG and brain imaging. Brain tumors within the cerebral hemispheres often cause focal slowing of the EEG frequency with increased amplitude. This is called a *slow-wave focus.* The tumor itself is electrically silent, but its effects on the surrounding cerebral cortex are slowing, and the intermittent bursts of electrical spikes represent seizure activity within the cortex. Between clinical seizures, these spikes can still be seen with the EEG, but they do not spread widely enough within the cortex to cause a clinical seizure. During a clinical seizure, this abnormal electrical activity spreads more widely throughout the brain tissue, causing the various motor and other events of the seizure. Why such spikes only occasionally spread to incorporate more distant parts of the brain to cause seizures, and why seizures stop, is still unknown.

Treatment

Many forms of seizure disorders can be managed successfully by removing the underlying cause, or the frequency of the seizures can be reduced with antiepileptic medication. In this dog, the cause is likely a brain tumor. Depending on the nature of the tumor, surgery and radiation therapy may be possible and may extend the dog's life. However, the prognosis is likely poor. Antiepileptic and steroid medication may improve the quality of the dog's remaining life.

Outbreak of Neurological Disease in Goats

History

A dairy goat producer calls you to say that, over the last few days, she has noticed several goats that are stumbling and wandering around in their pens. A few have head tilts and don't seem as though they are eating as well. Some are mildly affected, but others are more severely affected, and they all appear to be getting worse. The farm just got new silage in the last 3 weeks.

Clinical Examination

You examine the worst goat. She has a fever, increased heart rate and respiratory rate, and appears to be head pressing (standing with head against a wall or other object for no apparent reason) with a head tilt to the right. She is not interested in the food she is offered. Aside from apparent vestibular involvement suggested by the head issues, other cranial nerves appear normal. She has conscious proprioceptive (CP) deficits (ataxia) in all four legs. Based on the history of change in silage, and with the number of animals affected, *Listeria monocytogenes* would be one of the top differentials. Cerebrospinal fluid (CSF) analysis with increased white blood cells and protein also supports a diagnosis of *Listeria.* Culture of CSF could also be performed.

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

Comment

The neurological form of *Listeria* can cause disease affecting brainstem, cerebrum, or spinal cord. Typical changes include mononuclear cell inflammation and microabscesses within the CNS. Infection of the caudal brainstem (pons and medulla) can affect cranial nerves V and VII, VIII, IX, and X, and involvement of the reticular activating system can result in a depressed level of consciousness or altered mentation (mental activity). A head tilt is indicative of infection of cranial nerve VIII. CP deficits are due to infection and interference with ascending sensory pathways. BAER could potentially be performed here to identify where some of the changes are present, although EEG would be of little utility for the focal caudal brainstem involvement suggested here. However, because of expenses, money is invested in treatment instead of the additional diagnostic assay.

Treatment

Goats are often treated with oxytetracycline or procaine penicillin. Other affected goats should be treated, and the rest of the herd should be monitored. The silage should be examined as a possible source of *Listeria* and/or throwing away the silage and getting a new source should be considered.

Practice Questions

1. Which of the following regarding EEG is *false?*
- **a.** Its measurement relies on volume conduction.
- **b.** It predominantly measures postsynaptic potentials in the cerebral cortex.
- **c.** It is commonly used to measure the activity of a small number of neurons.
- **d.** It is a measure of the spontaneous activity of brain tissue.
- **e.** Both a and d.

2. A lesion in which of the following brain structures would be least likely to have a significant effect on the EEG?
- **a.** Cerebral cortex
- **b.** Thalamus
- **c.** Hypothalamus
- **d.** Cerebellum
- **e.** Reticular activating system

3. Which of the following statements is *true?*
- **a.** A slow-frequency, high-amplitude EEG is said to be "desynchronized."
- **b.** The EEG alone is used to confirm brain death.
- **c.** There are some periods of high-frequency, low-amplitude EEG during sleep.
- **d.** The EEG is usually measured in response to the stimulation of a sensory receptor organ.

4. The BAER requires averaging out the random background EEG activity before it can be observed.
- **a.** True
- **b.** False

5. A brain tumor may cause focal slowing of the EEG from the brain tissue immediately surrounding the tumor.
- **a.** True
- **b.** False

Bibliography

Bagley RS. *Fundamentals of Veterinary Clinical Neurology*. Ames, Iowa: Blackwell Publishing; 2005.

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Blades Golubovic S, Rossmeisl JH Jr. Status epilepticus in dogs and cats, part 1: etiopathogenesis, epidemiology, and diagnosis. *J Vet Emerg Crit Care (San Antonio)*. 2017;27(3):278–287.

Dewey CW, da Costa RC, Ducote JM. Neurodiagnostics. In: Dewey CW, da Costa RC, eds. *A Practical Guide to Canine and Feline Neurology*. 3rd ed. Oxford, UK: Wiley-Blackwell; 2015.

Haines DE, ed. *Fundamental Neuroscience for Basic and Clinical Applications*. 3rd ed. Philadelphia: Churchill Livingstone; 2006.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Kandel ER, Schwartz JH, Jessell TM, et al, eds. *Principles of Neural Science*. 5th ed. New York: McGraw-Hill; 2012.

Lorenz MD, Coates JR, Kent M. *Handbook of Veterinary Neurology*. 5th ed. Philadelphia: Saunders; 2010.

Pellegrino FC, Sica REP. Canine electroencephalographic recording technique: findings in normal and epileptic dogs. *Clin Neurophysiol*. 2004;115:477–487.

Poma R, Chambers H, da Costa RC, et al. MRI measurement of the canine auditory pathways and relationship with brainstem auditory evoked responses. *Vet Comp Orthop Traumatol*. 2008;21(3):238–242.

Rossmeisl JH, Pancotto TE. Intracranial neoplasia and secondary pathological effects. In: Platt S, Garosi L, eds. *Small Animal Neurological Emergencies*. London, UK: Manson Publishing Ltd; 2012.

Smith B. *Large Animal Internal Medicine*. 5th ed. St. Louis: Mosby Elsevier; 2015.

Strain GM. Consciousness and higher cortical function. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Williams DC, Aleman T, Holliday TA, et al. Qualitative and quantitative characteristics of the electroencephalogram in normal horses during spontaneous drowsiness and sleep. *J Vet Intern Med*. 2008;22(3):630–638.

Vet-ebooks.com

17 Hearing

BRADLEY G. KLEIN

KEY POINTS

1. Sound waves are alternating phases of condensation and rarefaction (pressure waves) of molecules in the external environment.
2. Outer and middle ears funnel sound waves to the cochlea.
3. The cochlea is located in the inner ear.
4. The cochlea transduces sound waves to action potentials of the eighth cranial nerve.
5. Deciphering of sound wave frequency begins in the cochlea.
6. Action potentials from the cochlea are transmitted up through the brainstem to the cerebral cortex.
7. Deafness results from an interruption in the hearing process.

Our lives are enriched by music and conversation, and altered by the sounds of danger. Many mammalian species have a particularly acute sense of hearing. Hearing depends on the remarkable properties of hair cell receptors in the cochlea that mediate transduction of sound into action potentials that are then sent to the brain. Fortunately, the auditory system is not usually a site of pathological lesions in veterinary medicine, except for occasional congenital defects and exposure to ototoxic drugs. Nevertheless, hearing is sufficiently important to warrant a brief discussion of its physiology.

Sound Waves Are Alternating Phases of Condensation and Rarefaction (Pressure Waves) of Molecules in the External Environment

Sound waves are longitudinal vibrations of molecules in the external environment characterized by alternating phases of condensation and rarefaction (increases and decreases in pressure). These alternating changes in pressure produce the sensation of sound after they strike the tympanic membrane and are subsequently transduced into neural signals that ultimately reach the cerebral cortex. Sound waves reaching the tympanic membrane can be expressed as changes in sound pressure as a function of time (Fig. 17.1).

In general, the subjective *loudness* of the sound is correlated with the *amplitude* of a sound wave; the subjective *pitch* is correlated with the *frequency* of the wave. The amplitude of a sound is usually quantified according to the logarithmic *decibel scale,* which expresses the energy of the sound relative to the energy of a standard reference sound. This standard sound, representing 0 decibels (dB), is the threshold for human hearing where the sound of the background movement of air molecules can almost be heard. Normal conversation is about 60 dB, and the loudest tolerable sound for humans is about 120 dB, about 1 million times the threshold amplitude. The loudest reported dog barking has been recorded as 108 dB. Sound frequency, the number of pressure oscillation cycles per unit time, is usually expressed in units called hertz (Hz), where 1 Hz equals 1 cycle per second.

Outer and Middle Ears Funnel Sound Waves to the Cochlea

The outer ear, composed of the fleshy part (pinna) and the ear canal, funnels sound waves to the *tympanic membrane,* or eardrum (Fig. 17.2). Some animals can move the pinna to more effectively collect sound waves, and the natural shape of the pinna can act to selectively filter out or accentuate certain sound frequencies. In some species, the contours of the pinna are particularly helpful for localizing sound in the vertical plane. The eardrum is a membrane between the outer and the middle ear. The middle ear is an air-filled cavity in the temporal bone connected to the nasopharynx by the auditory (eustachian) tube. Three tiny bones—the malleus, incus, and stapes—collectively called the *ossicles,* are connected to each other and are located in the middle ear. The malleus is connected to the eardrum, and the stapes is connected to the oval window, a membranous separation between the middle and the inner ear. The ossicles transfer vibration of the eardrum to the oval window in a manner that avoids a significant loss of energy as the sound wave is transferred from the air-filled outer ear to the fluid-filled inner ear. Two small skeletal muscles are also located in the middle ear, with one attached to the malleus and one attached to the stapes. Their contraction reduces the transfer of vibration between the eardrum and the oval window. This attenuation reflex can function to protect the inner ear from very loud sounds.

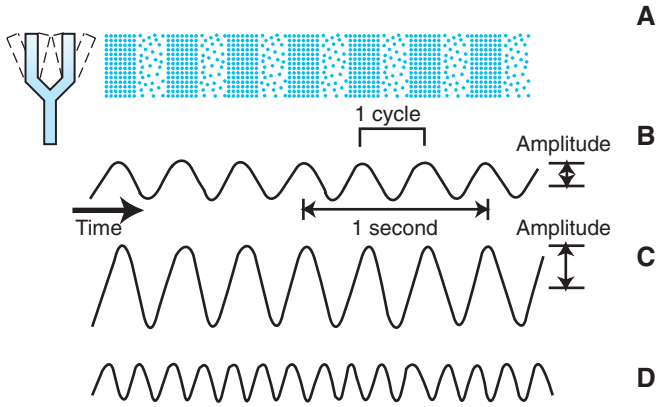
The Cochlea Is Located in the Inner Ear

The inner ear (labyrinth) contains the receptor organs of two sensory systems: (1) the vestibular system, which detects acceleration and static tilt of the head (see Chapter 11), and (2) the auditory system, which detects and analyzes sound. The inner ear consists of the *bony labyrinth* and, within the bony labyrinth, the *membranous labyrinth.* The bony labyrinth is a series of tunnels within the petrous temporal bone. Inside these tunnels, surrounded by a fluid called *perilymph,* is the membranous labyrinth. The membranous labyrinth follows the contour of the bony labyrinth and contains *endolymph.* The vestibular and auditory portions of the inner ear are contiguous,

Vet-ebooks.com

and the "membranous tunnel within a bony tunnel" design is an anatomical feature of both parts. The auditory portion of this inner ear complex is called the *cochlea* (see Figs. 11.1 and 17.2E).

The cochlear portion of the labyrinth is coiled like a snail shell. If we could mentally uncoil this arrangement to a linear form and then take a transverse section through it, perpendicular to the long axis (like cutting a salami and then looking at the cut end), we would see two membranes, the basilar and Reissner's, dividing the cochlea into three chambers, or *scalae* (Fig. 17.3). The dorsally located *scala vestibuli* and ventrally located *scala tympani* contain perilymph. The flexible middle scala, or *scala media* (cochlear duct), is formed by the membranous portion of the labyrinth and contains endolymph. The *basilar membrane* is the floor of the scala media, and atop this membrane lies the hair cell receptor organ for hearing called the *organ of Corti.* An anchored, gel-coated ridge called the *tectorial membrane* lies just atop the hair cells of the organ of Corti.



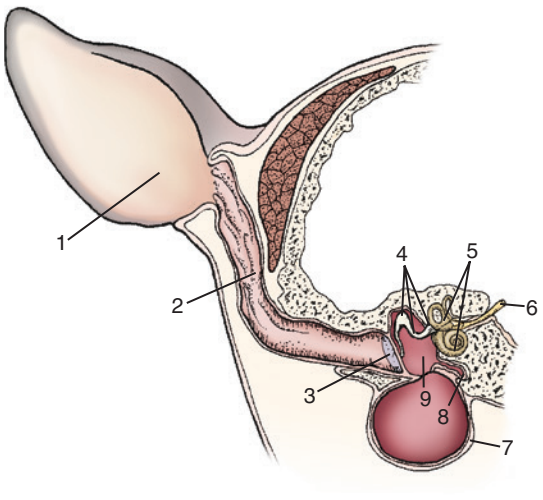
• **Fig. 17.1** Characteristics of sound waves. (A) Cyclical expansion and contraction of the tuning fork produces a cyclical compression and rarefaction of air molecules and a cyclical change in air pressure. (B) Cyclical change in air pressure corresponding to a pure tone. The number of cycles per second is the frequency of the tone and is expressed in hertz (Hz). The frequency of the tone in *B* is 3 Hz. The amplitude of the wave reflects the magnitude of the pressure increase and is usually expressed in decibels (dB). (C) Tone with a greater amplitude is perceived as louder than *B.* (D) Tone with a greater frequency is perceived as having a higher pitch than *B* and *C.*

The morphological organization just described is virtually the same all along the length of the cochlea, except that the scala vestibuli and scala tympani connect with each other at the distal end (farthest from the oval window).

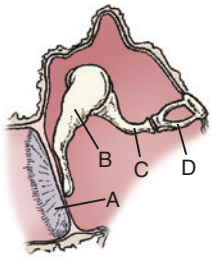
The Cochlea Transduces Sound Waves to Action Potentials of the Eighth Cranial Nerve

The organ of Corti mediates the transduction of sound waves into action potentials. The hair cell receptors of the organ of Corti are similar in structure and function to the hair cells that form the vestibular sensory organs. The hair cells synapse on sensory neurons that form the cochlear portion of the eighth cranial (vestibulocochlear) nerve, which projects to the brainstem's *cochlear nuclei.* The cell bodies of these sensory neurons reside in the *spiral ganglion* (see Fig. 17.3). Sound-induced bending of the hair cell cilia changes the frequency of action potentials on the eighth nerve fibers.

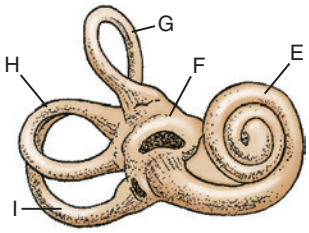
Sound waves in the external environment are collected by the outer ear and cause vibrations of the tympanic membrane. These vibrations are transmitted through the middle ear by movement of the ossicles and result in similar vibrations of the oval window of the cochlea. As the oval window vibrates, sound energy is transferred down through the perilymph of the scala vestibuli and down through the endolymph of the scala media to the basilar membrane. This energy produces a series of traveling waves that begin near the base of the basilar membrane (closest to the oval window) and move along its length. The situation is analogous to whipping the free end of a rope that is stationary at the opposite end. A diagram of this transmission is shown in Fig. 17.4. The movement of the traveling wave causes portions of the flexible basilar membrane to move up and down. Because the organ of Corti sits atop the basilar membrane, this up-and-down motion causes the hair cell cilia to be sheared back and forth against the anchored, overlying tectorial membrane (Fig. 17.5). This, in turn, changes the release of transmitter from the hair cells onto the eighth nerve neurons, which in turn alters the action potential firing rate of these neurons. It is at this point that the organ of Corti has transduced the sound wave energy into neural activity. As the amplitude of an environmental sound wave increases



Schematic section through the left ear



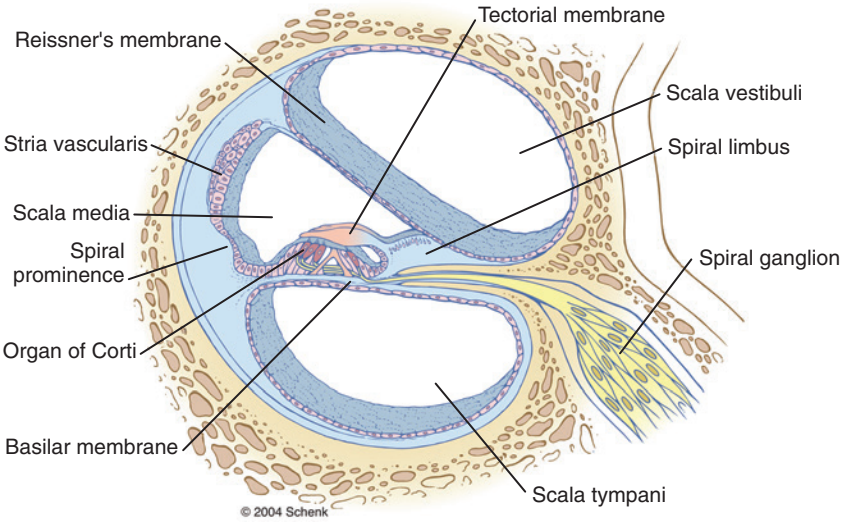
The auditory ossicles



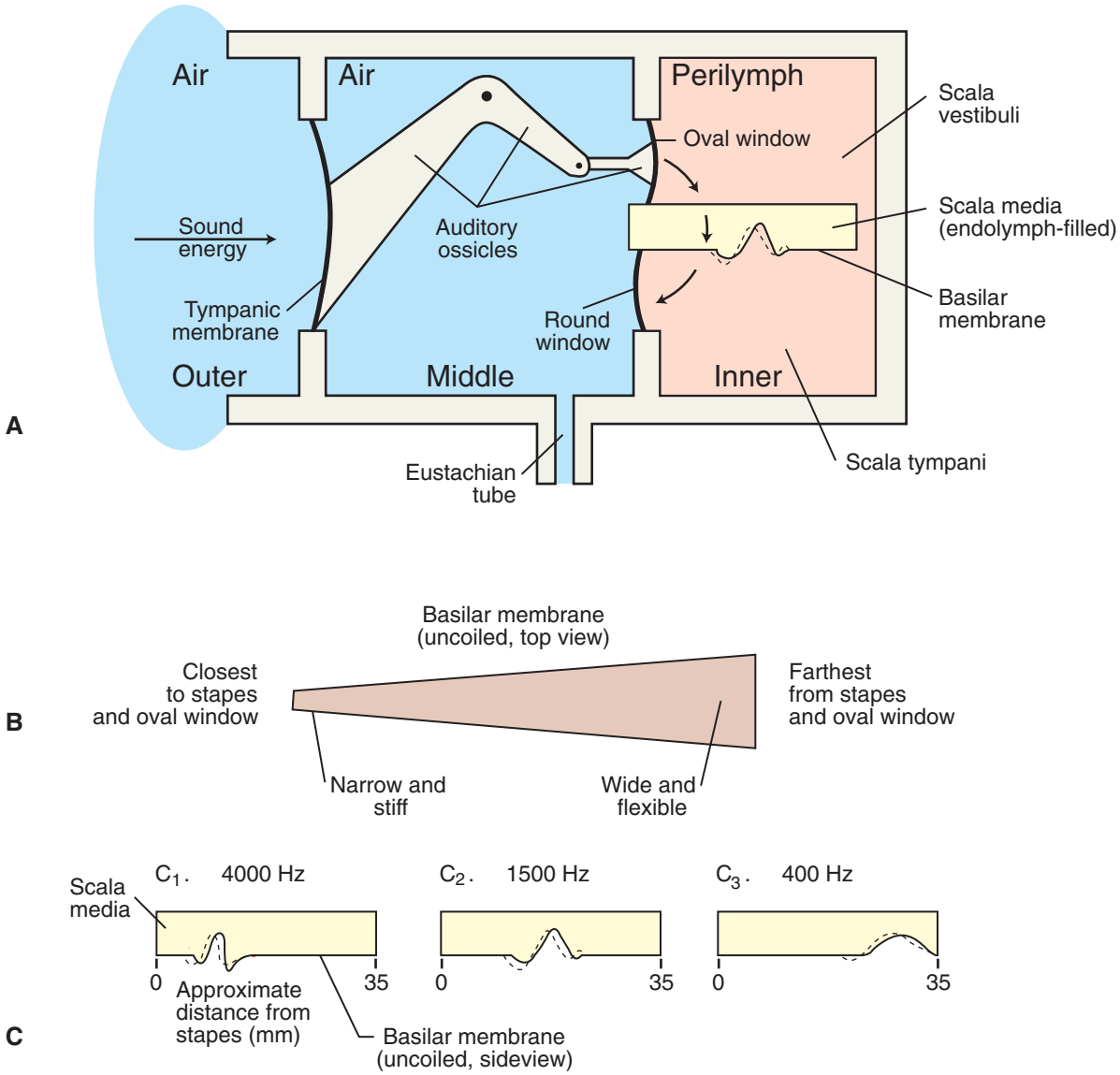
The bony labyrinth

• **Fig. 17.2** Schematic diagrams of a section through the left ear, the auditory ossicles, and the bony labyrinth. *1,* Pinna; *2,* ear canal; *3,* tympanic membrane; *4,* auditory ossicles; *5,* bony labyrinth; *6,* eighth cranial nerve; *7,* tympanic bulla; *8,* eustachian tube; *9,* middle ear; *A,* tympanic membrane; *B,* malleus; *C,* incus; *D,* stapes; *E,* cochlea; *F,* utricle; *G, H,* and *I,* semicircular canals. (From Getty R. *Atlas for Applied Veterinary Anatomy.* 2nd ed. Ames, Iowa: Iowa State University Press; 1964.)

Vet-ebooks.com

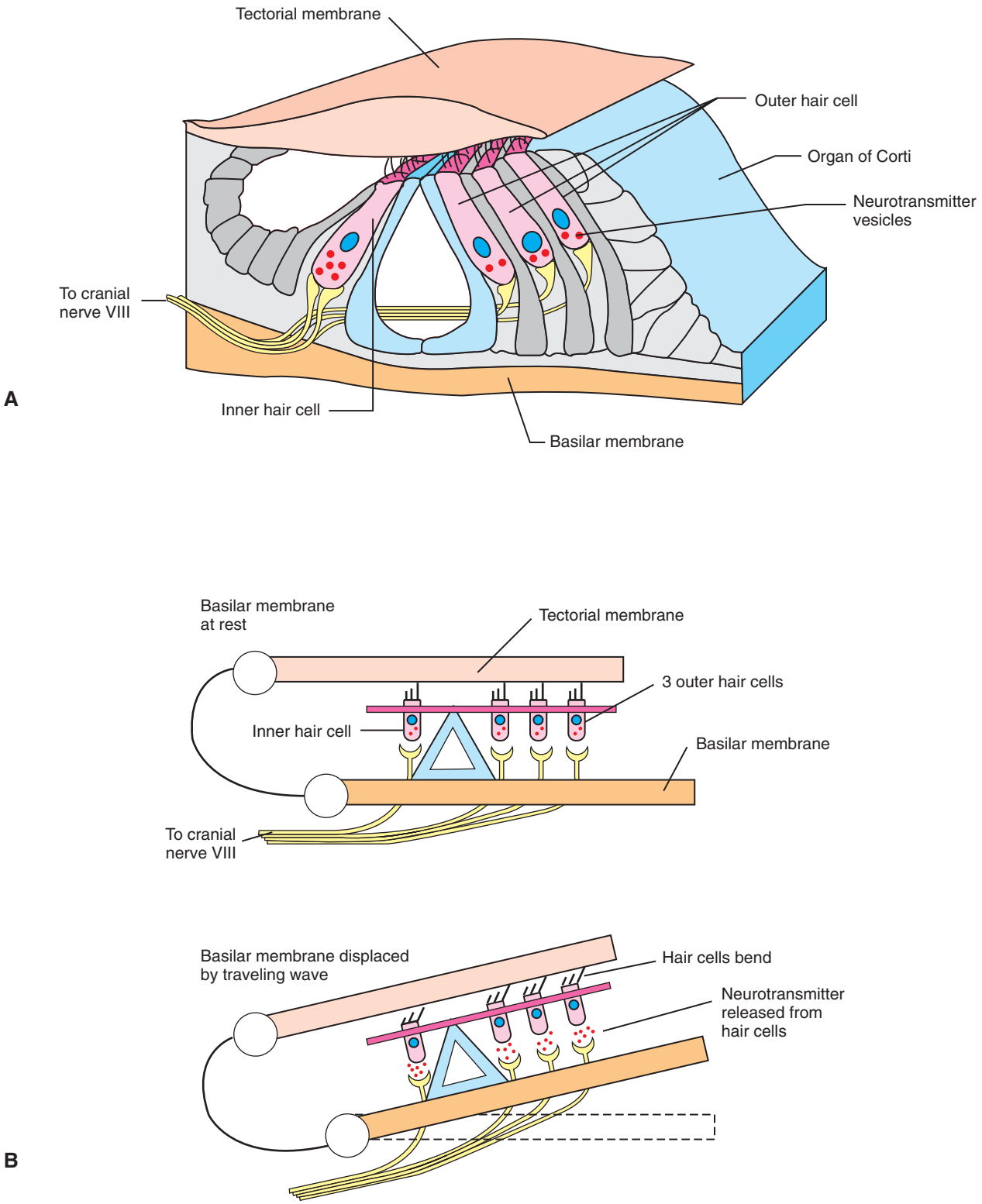


• **Fig. 17.3** Schematic representation of a section through one of the turns of the cochlea. (From Hall JE. *Guyton and Hall Textbook of Medical Physiology.* 12th ed. Philadelphia: Saunders; 2011.)



• **Fig. 17.4** Schematic representation of the transmission of sound energy from the outer ear to the inner ear. (A) Sound energy enters the air-filled outer ear and vibrates the tympanic membrane, which produces movement of the ossicles in the air-filled middle ear. Ossicular movement displaces the oval window of the fluid-filled inner ear, resulting in a traveling wave on the basilar membrane. (B) Basilar membrane is narrow and stiff at the base and wider and more flexible farther from the base. (C) As the frequency of a sound decreases, the region of maximum displacement of the basilar membrane, produced by the traveling wave, is located progressively farther from the base. (Redrawn from Lippold OCJ, Winton FR. *Human Physiology*. 6th ed. New York: Churchill Livingstone; 1972.)

Vet-ebooks.com



• **Fig. 17.5** Transduction of sound into neural activity occurs at the organ of Corti. (A) Structural organization of the organ of Corti and its hair cell receptors, sitting atop the basilar membrane. (B) Traveling waves along the basilar membrane displace the membrane and organ of Corti upward, shearing hair cell cilia against the overlying tectorial membrane. This bending of the cilia induces the release of neurotransmitter from the hair cells onto sensory neurons of the eighth cranial nerve. (Modified from Goldstein EB. *Sensation and Perception*. 6th ed. Pacific Grove, Calif: Wadsworth; 2002.)

Vet-ebooks.com

(normally perceived as a louder sound), a longer area of the basilar membrane is displaced. This results in a larger number of hair cells shearing against the tectorial membrane, which in turn affects the activity of a larger number of eighth nerve neurons. This is one way that sound intensity is coded by the nervous system.

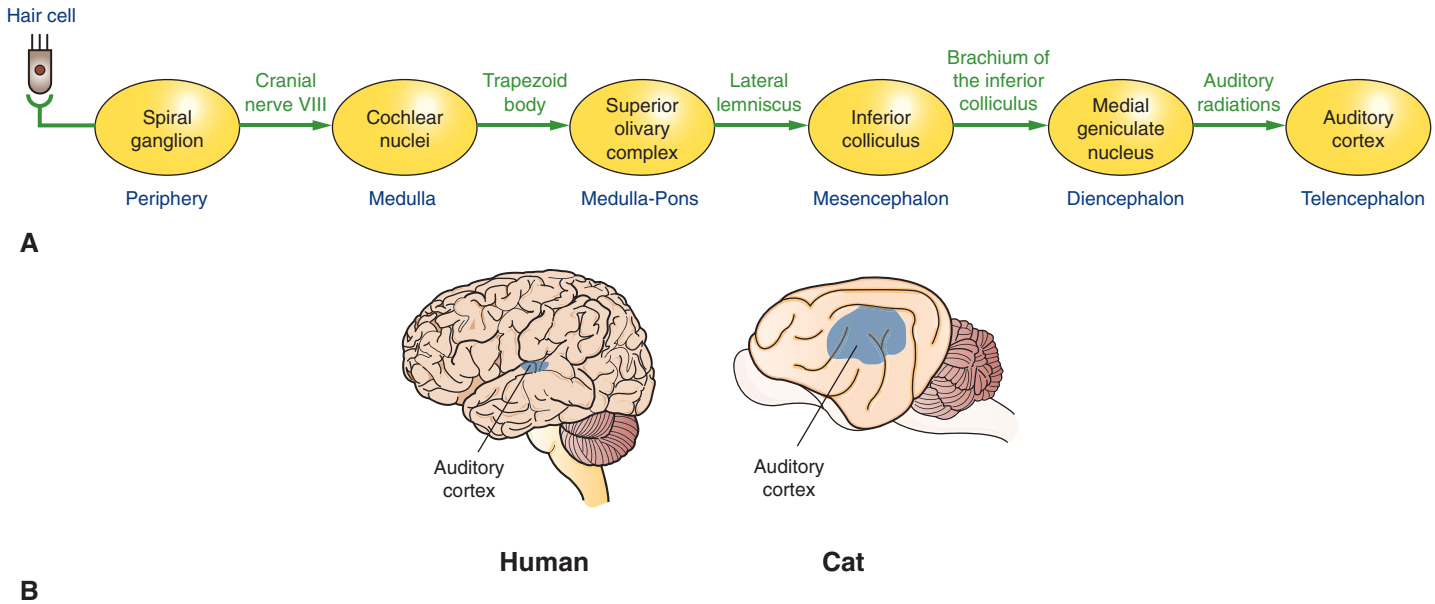
Deciphering of Sound Wave Frequency Begins in the Cochlea

The physical properties of the basilar membrane are not uniform along its length. It is narrower and relatively stiffer at its base (near the oval window), and it becomes progressively wider and more flexible toward its apex (see Fig. 17.4B). Thus the resonant properties of the membrane are not uniform. A region near the base is significantly displaced by a high-frequency sound, and as the sound frequency decreases, the region of significant displacement is located progressively closer to the apex (see Fig. 17.4C). Because the organ of Corti sits atop the basilar membrane, high-frequency sounds are most likely to affect hair cells and their associated eighth nerve neurons near the base of the membrane. As frequency decreases, the activated hair cells and neurons are those located progressively closer to the apex. Given this orderly relationship between the frequency of a sound wave and the region of the cochlea activated by that frequency, the cochlea is said to have a *tonotopic organization.* Therefore a principal means by which the nervous system begins to decipher the frequency of a sound is through the location of the hair cells and neurons most affected by that sound.

Among species, there is a rough positive correlation between the number of coils or turns in the cochlea and the size of the frequency range for hearing, although there are exceptions for some species with specialized cochleas (e.g., horseshoe bat, kangaroo rat). The upper frequency range of hearing appears to be negatively correlated with the distance between the two ears.

Action Potentials From the Cochlea Are Transmitted up Through the Brainstem to the Cerebral Cortex

Action potentials arising in the cochlea travel along the cochlear portion of the eighth cranial nerve to the cochlear nuclei in the medulla oblongata. From there, neural activity is synaptically relayed in a sequential manner to the *superior olivary complex* (a group of nuclei spanning the pontomedullary border region), the *inferior colliculus* of the midbrain, the *medial geniculate nucleus* of the thalamus, and finally to the *auditory cortex* of the temporal lobe (Fig. 17.6). Conscious perception of sound occurs in the cerebral cortex. Because of extensive connections of central auditory neurons across the midline, information originating in the cochlear nuclei on one side can reach other auditory nuclei on both sides of the brain. However, information originating from a given cochlea is predominantly conducted to the contralateral auditory cortex. Each nucleus in the auditory pathway has a tonotopic representation of sound frequency but is specialized to process particular features of sound. For example, the superior olivary complex plays a major role in determining which side of the head an environmental sound source is coming from. Important environmental cues for this directional localization are differences in the intensity and in the time of arrival of a sound at the two different ears. On the other hand, the medial geniculate nucleus is specialized to detect certain combinations of frequencies, as well as timing patterns among sounds.



• **Fig. 17.6** Principal components of the auditory pathway from hair cell to cerebral cortex. (A) *Ovals* represent ganglia (peripheral nervous system) or nuclei (central nervous system) and *arrows* represent named axonal connections between them. Major brain divisions (e.g., medulla, pons, etc.) where the structures reside are also noted. (B) The temporal lobe location of auditory cortex in the human and cat brains. (Modified from Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain.* 3rd ed. Philadelphia: Lippincott, Williams & Wilkins; 2007.)

Vet-ebooks.com

Deafness Results From an Interruption in the Hearing Process

Clinical deafness may result from a loss of sound transmission in the outer or middle ear, called *conduction deafness,* or from malfunction of the cochlear hair cells or eighth nerve fibers, called *nerve deafness* or *sensorineural deafness.* Given that auditory information from one ear is significantly distributed to both sides of the central nervous system, unilateral damage to the auditory system in the brain is often difficult to detect or localize with a traditional hearing test. In veterinary medicine, inflammatory lesions and neoplasms of the outer or middle ears are frequently the cause of conduction deafness. Sometimes the inflammation can secondarily spread to the inner ear to also cause sensorineural deafness. Deafness in young animals is usually caused by a congenital defect in the cochlea (inherited sensorineural deafness), frequently associated with white coat coloration. Certain antibiotics, diuretics, and antineoplastic agents have ototoxic properties capable of damaging cochlear structures (acquired sensorineural deafness). Like humans, dogs and cats are also susceptible to age-related hearing loss (*presbycusis*).

CLINICAL CORRELATIONS

Congenital Deafness

History

An owner brings you an 8-week-old, almost completely white, male Dalmatian and reports that the pup does not appear to hear anything.

Clinical Examination

Your physical examination reveals a seemingly normal, healthy Dalmatian pup except for his apparent deafness. He does not seem to respond to voice commands or loud noises. Vestibular and all other neurological reflexes are within normal limits. A brainstem auditory-evoked response (BAER; see Chapter 16) is flat, which suggests that the brain has not received any signal from the cochlea.

Comment

Congenital deafness is fairly common in dogs and other animals with white coat color. It is usually caused by the partial or complete absence of the cochlea and occasionally by the absence of other neural elements in the auditory pathway. This is known as *nerve deafness* or *sensorineural deafness* and is usually present at birth (congenital). Why it is linked to white coat coloration is under investigation, but the pattern suggests that this is a genetically determined failure, usually bilateral, of the cochlea to develop. This dog can have a relatively normal life as long as the owners are attentive.

Deafness in American Paint Horse Associated With Genetic Defect

History

An owner has recently adopted a 13-year-old paint gelding with an overo spotting pattern from a local rescue organization. She reports that the horse seems less active than normal and frightens easily.

Clinical Examination

Your physical examination is within normal limits except for the horse being easily startled. Routine blood work is normal but the animal fails to show a behavioral response if auditory stimuli are presented outside the visual field. Brainstem auditory-evoked response (BAER) testing reveals that waveforms are virtually absent, and blood submitted for analysis of the endothelin B receptor (EDNBR) demonstrates a mutation.

Comment

In several species, an association has been identified between genetic defects, coat color, and deafness. In some paint horses, there is a defect of the EDNBR gene, which has been associated with overo lethal white (coat) syndrome. The presence and phenotype of the genetic defect has been determined in deaf, suspected deaf, and non-deaf paint and pinto horses. All deaf horses lacked the BAER, had at least one blue eye, and had extensive white markings on the limbs and head. Ninety-one percent of deaf and suspected deaf horses had the EDNBR mutation. Hearing is regulated by the outer ear, which collects sound waves and then causes vibrations of the tympanic membrane. These vibrations are then passed through the middle ear via movement of the ossicles. This results in vibrations at the oval window of the cochlea. The oval window then vibrates, which transfers the sound energy through the endolymph of the scala media, to the base of the basilar membrane. Because the organ of Corti sits on top of the basilar membrane, the sound energy is transmitted to the hair cell cilia, which then bend against the tectorial membrane, and this leads to release of neurotransmitter onto the eighth cranial nerve fibers. Melanocytes are cells that play an important role in determining skin pigmentation. However, a defect in functional melanocytes is thought to contribute to some forms of deafness because these cells normally contribute to the blood vessel-rich stria vascularis of the scala media of the inner ear. The stria vascularis normally regulates the chemical composition of the endolymph of the scala media within which the organ of Corti resides. A lack of normal melanocytes, associated with some pigmentation variations, can alter the development and function of the stria vascularis, as well as the endolymph. This can lead to degeneration of cochlear hair cells and auditory nerve neurons. Therefore the relationship between coat pigmentation pattern, melanocyte function, and EDNBR gene mutation warrants further study in equine congential deafness.

Treatment

Once a horse has been identified as being deaf, the horses can adapt well if accommodations are made (e.g., relying on visual and tactile cues during training). Many horses suspected of being deaf are still used as intended.

Practice Questions

1. Which one of the following is the *first* to be displaced by sound energy entering the ear?
- **a.** Oval window
- **b.** Tectorial membrane
- **c.** Basilar membrane
- **d.** Tympanic membrane
- **e.** Stapes

2. Hair cells similar to those of the organ of Corti are important for the function of which *two* of the following sensory organs?
- **a.** Muscle spindle
- **b.** Retina
- **c.** Crista ampullaris
- **d.** Golgi tendon organ
- **e.** Utricular macula

Vet-ebooks.com

3. Which one of the following statements is *false?*
- **a.** An increase in sound wave frequency is perceived as an increase in pitch.
- **b.** The area of greatest displacement of the basilar membrane moves closer to its base (toward the oval window) as the sound wave frequency decreases.
- **c.** A decrease in sound wave amplitude is perceived as decreased loudness.
- **d.** An increase in sound wave amplitude displaces a longer region of the basilar membrane.
- **e.** Contraction of muscles attached to the ossicles reduces the sound energy reaching the inner ear.

4. Which one of the following cranial nerves transmits sound to the brain?
- **a.** Second
- **b.** Seventh
- **c.** Eighth
- **d.** Tenth

5. Of the following brain nuclei, which one receives auditory information *last?*
- **a.** Medial geniculate nucleus
- **b.** Cochlear nuclei
- **c.** Superior olivary complex
- **d.** Inferior colliculus

Bibliography

Bear MF, Connors BW, Paradiso MA. *Neuroscience: Exploring the Brain*. 4th ed. Philadelphia: Wolters Kluwer; 2016.

Brodal P. *The Central Nervous System: Structure and Function*. 5th ed. New York: Oxford University Press; 2016.

De Lahunta A, Glass E, Kent M. *Veterinary Anatomy and Clinical Neurology*. 4th ed. Philadelphia: Elsevier Saunders; 2015.

Goldstein EB. *Sensation and Perception*. 8th ed. Pacific Grove, Calif: Wadsworth; 2009.

Haines DE, ed. *Fundamental Neuroscience*. 5th ed. New York: Elsevier; 2018.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Magdesian KG, Williams DC, Aleman M, Lecouteur RA, Madigan JE. Evaluation of deafness in American Paint Horses by phenotype, brainstem auditory-evoked responses, and endothelin receptor B genotype. *J Am Vet Med Assoc*. 2009;235(10):1204–1211.

Purves D, Augustine GJ, Fitzpatrick D, et al. *Neuroscience*. 6th ed. New York: Sinauer; 2018.

Strain GM. Deafness prevalence and pigmentation and gender associations in dog breeds at risk. *Vet J*. 2004;167(1):23–32.

Strain GM, Myers LJ. Hearing and equilibrium. In: Reece WO, eds. *Duke's Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Vet-ebooks.com

18 Overview of Cardiovascular Function

ROBERT B. STEPHENSON

KEY POINTS

1. Cardiovascular function is essential for life and health.
2. Impaired cardiovascular function (*cardiovascular pathophysiology*) is encountered frequently in veterinary medicine.
3. The cardiovascular system transports nutrients, waste products, water, electrolytes, hormones, and heat.
4. Two modes of transport are used in the cardiovascular system: bulk flow and diffusion.
5. Because diffusion is very slow, every metabolically active cell in the body must be close to a capillary carrying blood by bulk flow.
6. The pulmonary and systemic circulations are arranged in series, but the various organs within the systemic circulation are arranged in parallel.
7. Cardiac output is the volume of blood pumped each minute by one ventricle.
8. The perfusion pressure for the systemic circulation is much greater than the perfusion pressure for the pulmonary circulation.
9. Each type of blood vessel has physical properties suited to its particular function.
10. Blood is a suspension of cells in extracellular fluid (plasma).
11. The cellular component of blood includes red blood cells, white blood cells, and platelets.
12. Most of the oxygen in blood is carried in chemical combination with the protein hemoglobin within red blood cells.

Cardiovascular Function Is Essential for Life and Health

Cardiovascular physiology is the study of the function of the heart, the blood vessels, and the blood. The primary function of the cardiovascular system can be summarized in one word: *transport.* The bloodstream transports numerous substances essential for life and health, including the oxygen and nutrients required by every cell in the body. Blood also carries carbon dioxide and other metabolic waste products away from metabolically active cells and delivers them to the lungs, kidneys, or liver, where they are excreted. Blood also transports water, electrolytes, hormones, and heat from place to place in the body.

To appreciate the importance of cardiovascular transport, consider what happens if the heart stops contracting and circulation ceases. Unconsciousness results within a minute. Irreversible damage to the brain and other sensitive body tissues occurs within a few minutes, and death follows shortly thereafter. However, circulation does not have to stop completely for significant dysfunction to occur. In each tissue of the body, normal function depends on the delivery of adequate blood flow. The higher the rate of metabolism in a tissue, the greater the requirement for blood flow. The condition of inadequate blood flow to meet a tissue's metabolic needs is called *ischemia.* Even transient ischemia leads to *dysfunction.* For example, the loss of as little as 10% of the normal blood volume can reduce cardiac output and impair exercise performance. Persistent, severe ischemia leads to irreversible tissue damage and eventually to cell death (*necrosis*). The sequence of changes (ischemia → reversible tissue damage → irreversible tissue damage → necrosis) is called *infarction.* An area of ischemic necrosis is called an *infarct.*

Because normal cardiovascular function is essential for life and health, a practical understanding of cardiovascular function and dysfunction is vital to the veterinary clinician. Yet many veterinary students have difficulty understanding cardiovascular physiology. They tend to agree with William Harvey, the father of cardiovascular physiology, whose initial impression was that the motions of the heart and the blood were so complicated that they could be comprehended only by God. Harvey persisted, however, in his careful, deliberate studies of cardiovascular function and, in 1628, he set forth the first proof that the heart propels blood through the blood vessels in a circulatory pattern. Before Harvey's time, it was thought that blood flowed out of the heart into the blood vessels and then returned to the heart by backward flow in the same vessels. In other words, blood was thought to flow in a tidal manner, in much the same way that air flows through a single set of airways: first into the lungs and then back out.

We now take for granted that the cardiovascular system is a *circulatory system,* not a tidal system. However, the circularity of the cardiovascular system is precisely what makes it difficult to understand. It has no clear beginning or ending, and disturbances in one part of the cardiovascular system end up affecting all other parts as well. In recognition of this complexity, Chapters 18 to 26 are written with the goal of identifying the most basic and important concepts of *normal cardiovascular function* and explaining them in a way that best prepares the

Vet-ebooks.com

reader to understand, diagnose, and treat *cardiovascular dysfunction* (cardiovascular disease). The remainder of Chapter 18 reviews the general features of the cardiovascular system. Chapters 19 to 25 discuss the various elements of the cardiovascular system in detail. Chapter 26 summarizes cardiovascular function and dysfunction by describing the overall effects of heart failure, hemorrhage, and exercise.

Impaired Cardiovascular Function (*Cardiovascular Pathophysiology*) is Encountered Frequently in Veterinary Medicine

Some cardiovascular dysfunctions are *primary,* in that the underlying disturbance or disease affects the cardiovascular system directly. One example of primary cardiovascular dysfunction is *hemorrhage* (loss of blood from blood vessels). Primary cardiovascular dysfunctions may be *congenital* (present at birth) or *acquired* after birth. Defective heart valves, which either cannot open fully or cannot close completely, may be present at birth. In contrast, c*ardiomyopathies* (literally, heart muscle diseases) are most often acquired as a result of a toxic chemical, cardiac ischemia (resulting from coronary artery disease), or a viral or bacterial infection that inflames the heart muscle and impairs the ability of the heart to pump blood.

Although a heart that has a congenital defect or an acquired disease may be able to pump an adequate amount of blood when the animal is at rest, it usually cannot deliver the increased blood flow required by the body during exercise. When a dysfunction in the heart impairs its ability to pump blood, the condition is called *heart failure* (or *pump failure*). The patient with heart failure classically exhibits a limited ability or willingness to exercise (*exercise intolerance*).

Parasites are another common cause of acquired cardiovascular dysfunction. In dogs, for example, adult heartworms (*Dirofilaria immitis*) lodge in the right ventricle and pulmonary artery where they impede the flow of blood. These worms also release substances into the circulation that disrupt the body's ability to control blood pressure and blood flow. In horses, bloodworms (*Strongylus vulgaris*) lodge in the mesenteric arteries and decrease the blood flow to the intestine. The resulting intestinal ischemia depresses digestive functions (motility, secretion, and absorption), and the horse exhibits signs of gastrointestinal distress (*colic*).

In many other disease states, cardiovascular complications develop even though the cardiovascular system is not the primary target of the disease. These *secondary cardiovascular dysfunctions* often become the most serious and life-threatening aspects of the disease. For example, severe burns or persistent vomiting or diarrhea lead to substantial losses of water and *electrolytes* (small, soluble ions in the body fluids; e.g., Na^+, Cl^-, K^+, Ca^{2+}) . Even if the blood volume is not depleted to dangerously low levels in these conditions, the alteration in electrolyte concentrations can lead to abnormal heart rhythms (*cardiac arrhythmias*) and therefore ineffective pumping of blood by the heart (heart failure). The electrolyte abnormalities in such a patient can be made even worse if incorrect fluid therapy is given. Incorrect fluid therapy can also lead to an accumulation of excess fluid in the tissues of the body; this "waterlogging" of tissues is called *edema.* If the excess fluid gathers in the lung tissue, the condition is called *pulmonary edema.* Pulmonary edema is life threatening because it slows the flow of oxygen from the pulmonary air sacs (*alveoli*) into the bloodstream.

Pulmonary edema is a secondary complication in many disease states. A further example is *shock-lung syndrome,* which results when toxic substances in the body trigger an increase in the permeability of the lung blood vessels. These "leaky" vessels allow water, electrolytes, plasma proteins, and white blood cells to leave the bloodstream and accumulate in the lung tissue and airways. The resulting pulmonary edema can lead to death.

Shock is a state of profound depression in the vital processes of the body, usually caused by impaired cardiovascular function. Shock-lung syndrome is just one example. *Hemorrhagic shock* is a generalized cardiovascular failure caused by severe blood loss. *Cardiogenic shock* is a cardiovascular collapse caused by heart failure. *Septic shock* is caused by bacterial infections in the bloodstream (*bacteremia*). *Endotoxic shock* occurs when endotoxins (fragments of bacterial cell walls) enter the bloodstream; this often occurs when the epithelial lining of the intestines becomes damaged. Epithelial damage can result from bacterial infections in the intestines or from ischemia in the intestinal walls (as with bloodworms in horses). When the intestinal epithelium breaks down, endotoxins from the intestine can enter the bloodstream. These endotoxins then cause the body to produce substances that depress the pumping ability of the heart. The resulting heart failure leads to low blood flow and ischemia in all the vital body organs. Kidney (*renal*) failure, respiratory failure, central nervous system (*CNS*) depression, and death follow.

Anesthetic overdose is another clinical problem in which secondary cardiovascular complications are the most serious and life-threatening symptoms. Most anesthetics depress the CNS, and the resulting abnormal neural signals to the heart and the blood vessels can depress cardiac output and lower blood pressure. Some anesthetics also depress the pumping ability of the heart directly.

There are many other examples of primary and secondary cardiovascular dysfunction, but those just mentioned illustrate the importance and variety of cardiovascular dysfunctions encountered in veterinary medicine. The distinction between primary and secondary cardiovascular dysfunction is sometimes unclear, but this difficulty simply emphasizes how intimately the cardiovascular system is interconnected with all the other body systems and how dependent all the other systems are on the normal functioning of the cardiovascular system.

The Cardiovascular System Transports Nutrients, Waste Products, Water, Electrolytes, Hormones, and Heat

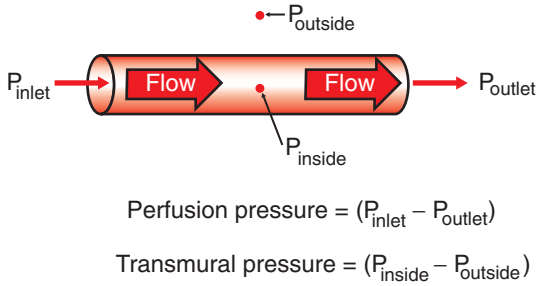
The blood transports the metabolic substrates needed by every cell of the body, including oxygen, glucose, amino acids, fatty acids, and various lipids. The blood also carries away from each cell in the body various metabolic waste products, including carbon dioxide, lactic acid, the nitrogenous wastes of protein metabolism, and heat. Although the heat produced by metabolic processes within cells is not a material waste product, its transport by the cardiovascular system to the body surface is essential, because tissues deep within the body would otherwise become overheated and dysfunctional.

Blood also transports vital chemical messengers: the hormones. Hormones are synthesized and released by cells in one organ and are carried by the bloodstream to cells in other organs, where they alter organ function. For example, insulin, which is produced by cells of the pancreas, is carried by the blood to cells throughout the body, where it promotes the cellular uptake of glucose. Inadequate insulin production (as in type 1 diabetes) results in inadequate entry of glucose into cells, whereas glucose concentrations in the blood rise to very high levels. The low intracellular glucose concentration is particularly disruptive to neural function, and the

Vet-ebooks.com

consequences can be serious (diabetic coma) or lethal. As another example, the hormones epinephrine and norepinephrine are released into the bloodstream by cells in the adrenal medulla during periods of stress. The epinephrine and norepinephrine circulate to various body organs, where they have effects that help prepare a threatened animal for "fight or flight." These effects include an increase in heart rate and cardiac contractility, dilation of skeletal muscle blood vessels, an increase in blood pressure, increased glycogenolysis, dilation of the pupils and airways, and piloerection (hair standing on end).

Finally, blood transports water and essential electrolytes, including Na^+, Cl^-, K^+, Ca^{2+}, H^+, and HCO_3^-. The kidneys are the organs primarily responsible for maintaining normal water and electrolyte composition in the body. The kidneys accomplish this by altering the electrolyte concentrations in blood as it flows through the kidneys. The altered blood then circulates to all other organs in the body, where it normalizes the water and electrolyte content in the extracellular fluids of each tissue.



• **Fig. 18.1** Fluid pressures associated with a blood vessel. P_{inlet}, P_{outlet}, and P_{inside} refer to blood pressures within the vessel. $P_{outside}$ refers to the pressure in the tissue fluid (interstitial fluid) immediately outside the blood vessel. Perfusion pressure is the pressure difference *along the length* of a blood vessel. Transmural pressure (distending pressure) is the pressure difference *across the wall* of the vessel, indicated here at the midpoint of the vessel. Perfusion pressure is the driving force for blood flow through the vessel, whereas transmural pressure is the driving force that would cause blood to flow out of the vessel if there were a hole in it.

Two Modes of Transport Are Used in the Cardiovascular System: Bulk Flow and Diffusion

Blood moves through the heart and blood vessels by bulk flow. The most important feature of bulk flow is that it is rapid over long distances. Blood pumped out of the heart travels quickly through the aorta and its various branches, and within 10 seconds, it reaches distant parts of the body, including the head and limbs. All transport requires energy, and the source of energy for bulk flow is a hydrostatic pressure difference. Unless the pressure at one end of a blood vessel is higher than the pressure at the other end, flow will not occur. The difference in pressure between two points in a blood vessel is called the *perfusion pressure difference* or, more often, simply *perfusion pressure.* Perfusion literally means "through-flow." Perfusion pressure is the pressure difference that causes blood to flow through blood vessels. The muscular pumping action of the heart creates the perfusion pressure that constitutes the driving force for bulk blood flow through the circulation.

It is important to distinguish between perfusion pressure difference and *transmural pressure difference* (usually shortened to *transmural pressure*). Transmural means "across the wall," and transmural pressure is the difference between the blood pressure inside a blood vessel and the fluid pressure in the tissue immediately outside the vessel (transmural pressure equals inside pressure minus outside pressure). Transmural pressure is the pressure difference that would cause blood to flow out of a vessel if a hole were poked in the vessel wall. Transmural pressure is also called *distending pressure,* because it corresponds to the net outward "push" on the wall of a blood vessel. Fig. 18.1 emphasizes the distinction between perfusion pressure and transmural pressure.

Diffusion is the second mode of transport in the cardiovascular system. Diffusion is the primary mechanism by which dissolved substances move across the walls of blood vessels, from the bloodstream into the interstitial fluid, or vice versa. *Interstitial fluid* is the extracellular fluid outside capillaries. It is the fluid that bathes each cell of a tissue. Most of the movement of substances between the blood and the interstitial fluid takes place across the walls of the *capillaries,* the smallest blood vessels. A substance (e.g., oxygen) moves from the bloodstream to a tissue cell by first diffusing across the wall of a capillary and into the tissue interstitial fluid, and then diffusing from the interstitial fluid and into the tissue cell.

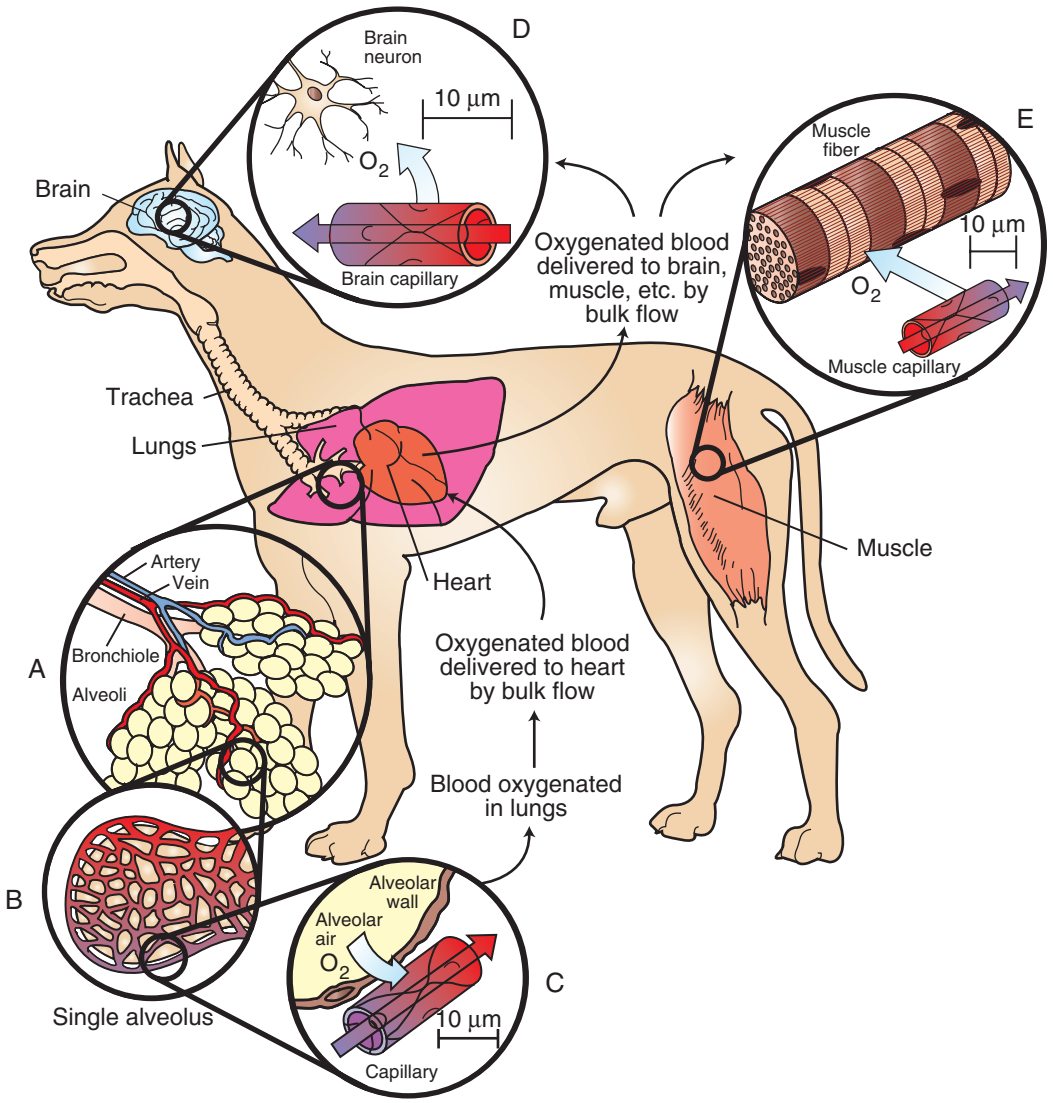
The source of energy for diffusion is a *concentration difference.* A substance diffuses from the bloodstream, across the wall of a capillary, and into the interstitial fluid only if the concentration of the substance is higher in the blood than in the interstitial fluid (and if the capillary wall is permeable to the substance). If the concentration of a substance is higher in the interstitial fluid than in the blood, the substance will diffuse from the interstitial fluid into the capillary blood. It is important to distinguish *diffusion,* in which a substance moves passively from an area of high concentration toward an area of low concentration, from *active transport,* in which substances are forced to move in a direction opposite to their concentration gradient. In general, substances are not transported actively across the walls of capillaries. The movement of substances between the bloodstream and the interstitial fluid occurs by passive diffusion.

Because Diffusion Is Very Slow, Every Metabolically Active Cell in the Body Must Be Close to a Capillary Carrying Blood by Bulk Flow

To understand more fully how the two types of transport (bulk flow and diffusion) are used in the cardiovascular system, consider the transport of oxygen from the outside air to a neuron in the brain. With each inspiration, fresh air containing oxygen (O_2) moves by bulk flow through progressively smaller airways (trachea, bronchi, and bronchioles) and finally enters the alveolar air sacs (Fig. 18.2A). The thin walls separating alveoli contain a meshwork of capillaries (see Fig. 18.2B). Blood flowing through these *alveolar capillaries* passes extremely close (within 1 µm) to the air in the alveoli (see Fig. 18.2C). The blood in an alveolar capillary has just returned from the body tissues, where it gave up some of its oxygen. Therefore the concentration of oxygen in alveolar capillary blood is lower than the concentration of oxygen in alveolar air. This concentration difference causes some oxygen to diffuse from the alveolar air into the capillary blood.

A large dog has about 300 million alveoli, with a total surface area of about 130 m^2 (equal to half the surface area of a tennis court). This huge surface area is laced with pulmonary capillaries. Thus even though only a tiny amount of oxygen diffuses into each pulmonary capillary, the aggregate uptake of oxygen into the pulmonary bloodstream is substantial (typically, 125 mL O_2/min in a large, resting dog, increasing by a factor of 10 or more during strenuous exercise). In summary, both the large alveolar surface area and the proximity of alveolar air to the blood in alveolar capillaries

Vet-ebooks.com



• **Fig. 18.2** Oxygen (O_2) is transported from the atmosphere to cells throughout the body by a combination of bulk flow and diffusion. First, O_2 moves by bulk flow through the airways, from the atmosphere to the alveoli (tiny air sacs) of the lungs (inset A). The wall of each alveolus contains a meshwork of alveolar (pulmonary) capillaries (inset B). O_2 readily diffuses from the alveolar air into the blood that is flowing through the alveolar capillaries (inset C). Bulk flow of blood next carries this O_2 from the lungs to the heart. From there O_2 is delivered by bulk flow into the capillaries of all the body organs (except the lungs). In the brain (inset D), skeletal muscle (inset E), and other tissues, O_2 moves by diffusion from the capillary blood into the interstitial fluid and then into the tissue cells, where it is utilized to support oxidative metabolism. Bulk flow is rapid; it can transport O_2 to all parts of the body within a few seconds. Diffusion is slow; it can transport O_2 efficiently only over distances less than 100 µm (note the distance scales in insets C, D, and E). Oxygenated blood has a bright-red color; deoxygenated blood is darker and bluish red.

promote efficient diffusion of oxygen; it takes less than 1 second for the blood in an alveolar capillary to become oxygenated.

As it leaves the lungs, each 100 mL of oxygenated blood normally carries 20 mL of oxygen. About 1.5% of this oxygen is carried in solution; the other 98.5% is bound to the protein *hemoglobin* within the *erythrocytes* (red blood cells). The oxygenated blood moves by bulk flow from the lungs to the heart. The heart pumps this oxygenated blood out into the aorta, and from there it is distributed via a complex system of branching arteries to all parts of the body (including the brain and skeletal muscles, as illustrated in Fig. 18.2). Capillaries in the brain bring a bulk flow of oxygenated blood very close to each brain neuron (see Fig. 18.2D). Metabolic processes within the neurons consume oxygen, so the oxygen concentration inside neurons is low. The gradient of oxygen concentration (high in the capillary blood and low in the neurons) provides the driving force for oxygen to diffuse first from the blood into the interstitial fluid and then into the neurons.

Each brain neuron must be within about 100 µm (i.e., 0.1 mm) of a capillary carrying blood by bulk flow if diffusion is to deliver oxygen rapidly enough to sustain normal metabolism in the neuron. Diffusional exchange over distances up to 100 µm typically takes only 1 to 5 seconds. If the distance involved were a few millimeters, diffusion would take minutes to occur. Diffusion of oxygen a few centimeters through body fluid would take hours. For this reason,

Vet-ebooks.com

normal life processes require that every metabolically active cell of the body be within about 100 µm of a capillary carrying blood by bulk flow. If this bulk flow is interrupted for any reason, perhaps because of a *thrombus* (blood clot) in the artery that delivers blood to a particular region of a tissue, that region of tissue becomes ischemic. As stated earlier, ischemia leads to cellular dysfunction; persistent, severe ischemia leads to irreversible damage and eventually to necrosis (the sequence called *infarction*). *Cerebral infarction* causes the condition commonly known as *stroke.*

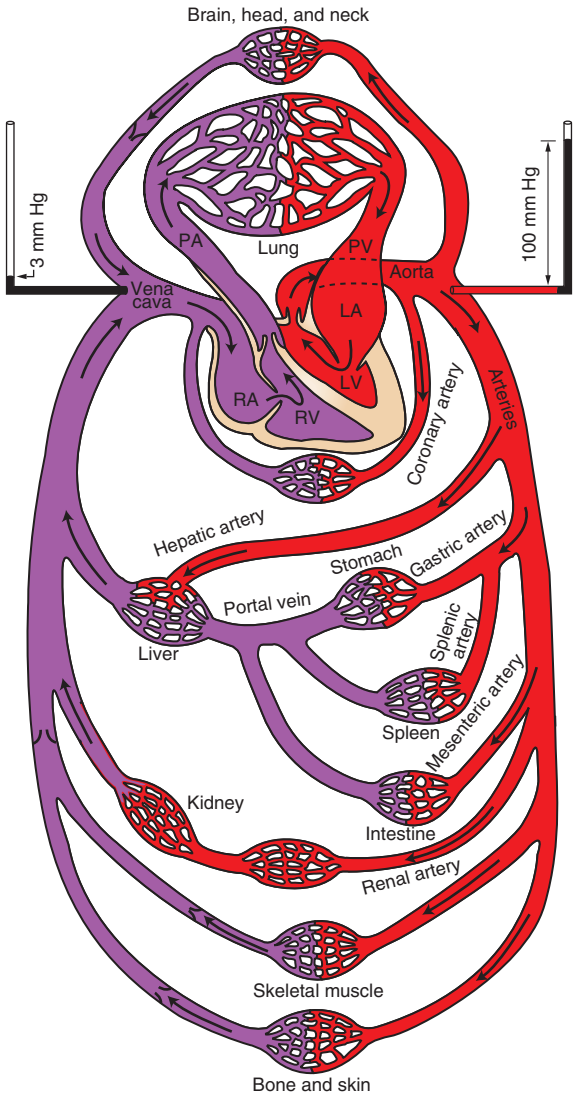
Fig. 18.2E shows a capillary carrying bulk flow of blood past a skeletal muscle cell (muscle fiber). Oxygen moves by diffusion from the capillary blood into the muscle interstitial fluid and then into the muscle cell, where it is consumed in the metabolic reactions that provide energy for muscle contraction. The oxygen consumption of a skeletal muscle depends on the severity of its exercise; at a maximum, oxygen consumption may reach levels 40 times greater than the resting level. Because of its tremendous metabolic capacity, muscle tissue has an especially high density of capillaries. In fact, several capillaries are typically arrayed around each skeletal muscle fiber. This arrangement provides more surface area for diffusional exchange than would be possible with a single capillary and brings the bulk flow of blood extremely close to all parts of each skeletal muscle cell.

Heart muscle, like skeletal muscle, consumes a large amount of oxygen. Oxygenated blood is carried from the aorta to the heart muscle by a network of branching *coronary arteries.* This arterial blood next moves by bulk flow into arterioles and then into *coronary capillaries,* which pass close by each cardiac muscle cell. If a thrombus interrupts the bulk flow of blood in a coronary artery, the heart muscle cells supplied by that artery become ischemic. Ischemia develops even if the cardiac muscle deprived of blood flow lies within a few millimeters of the left ventricular chamber, which is filled with oxygen-rich blood. Oxygen simply cannot diffuse rapidly enough from the ventricular chamber to the ischemic cells to sustain their metabolism. Ischemic cardiac muscle loses its ability to contract forcefully; also, cardiac arrhythmias may develop. Severe myocardial ischemia causes a *myocardial infarction,* or *heart attack.*

Coronary artery disease and *cerebrovascular disease* are encountered more often in human medicine than in veterinary medicine. In contrast, *cardiac disease* (dysfunction of the heart muscle or valves, as distinguished from disease of the coronary arteries) is encountered more often in veterinary medicine than in human medicine. Therefore Chapters 19 to 26 place more emphasis on cardiac physiology than on vascular physiology.

The Pulmonary and Systemic Circulations Are Arranged in Series, but the Various Organs Within the Systemic Circulation Are Arranged in Parallel

As shown in Fig. 18.3, blood is pumped from the left ventricle into the aorta. The aorta divides and subdivides to form many arteries, which deliver fresh, oxygenated blood to each organ of the body, except the lungs. The pattern of arterial branching that delivers blood of the same composition to each organ is called *parallel.* After blood passes through the capillaries within individual organs, it enters veins. Small veins combine to form progressively larger veins, until the entire blood flow is delivered to the right atrium by way of the *venae cavae* (pleural of vena cava that includes both superior vena cava and inferior vena cava). The blood vessels between the aorta and the *venae cavae* (including the blood vessels



• **Fig. 18.3** General layout of the cardiovascular system, showing that the systemic and pulmonary circulations are arranged in series and that the organs within the systemic circulation are arranged in parallel. Oxygenated blood has a bright-red color; deoxygenated blood is darker and bluish red. The drawing also shows that, if an open tube containing mercury (*black*) were stuck into the aorta, the normal blood pressure within the aorta would push mercury nearly 100 mm upward into the tube, at which point the upward force of the blood pressure would be equalized by the downward force of gravity acting on the mercury. In contrast, the blood pressure in the *venae cavae* is much lower (typically about 3 mm Hg), as illustrated on the left side of the drawing. *LA,* Left atrium; *LV,* left ventricle; *PA,* pulmonary artery; *PV,* pulmonary vein; *RA,* right atrium; *RV,* right ventricle. (Modified from Milnor WR. *Cardiovascular Physiology.* New York: Oxford University Press; 1990.)

in all organs of the body except the lungs) are collectively called the *systemic circulation.* From the right atrium, blood passes into the right ventricle, which pumps it into the pulmonary artery. The pulmonary artery branches into progressively smaller arteries and arterioles, which deliver blood to each alveolar (pulmonary) capillary. Blood from pulmonary capillaries is collected in pulmonary veins and brought to the left atrium. Blood then passes into the left ventricle, completing the circuit. The blood vessels of the lungs, including the pulmonary arteries and veins, constitute the *pulmonary circulation.* The pulmonary circulation and the heart are collectively

Vet-ebooks.com

termed the *central circulation.* The pulmonary circulation and the systemic circulation are arranged in *series;* that is, blood must pass through the pulmonary vessels between each passage through the systemic circuit.

In one pass through the systemic circulation, blood generally encounters only one capillary bed before being collected in veins and returned to the heart, although a few exceptions to this rule exist. One exception occurs in the *splanchnic circulation,* which supplies blood to the digestive organs. As shown in Fig. 18.3, blood that leaves the gastric, splenic, or mesenteric capillaries enters the *portal vein.* The portal vein carries splanchnic venous blood to the liver, where the blood passes through another set of capillaries before it returns to the heart. This arrangement of two systemic capillary beds in series is called a *portal system.* The splanchnic portal system allows nutrients that have been absorbed from the gastrointestinal tract to be delivered directly to the liver. There the nutrients are transformed for storage or allowed to pass into the general circulation. The liver also receives some blood directly from the aorta through the hepatic artery.

The kidneys also contain a portal system. As shown in Fig. 18.3, blood enters a kidney by way of a renal artery and passes through two sets of capillaries (called *glomerular* and *tubular*) before returning to the venous side of the systemic circulation. Large amounts of water, electrolytes, and other solutes are filtered out of the blood as it passes through the glomerular capillaries. Most of this filtered material is subsequently reabsorbed into the bloodstream as it flows through the peritubular capillaries. The remainder becomes urine. The kidneys use the *renal portal system* to adjust the amounts of water, electrolytes, and other critical solutes in the blood.

A third portal system is found in the brain and is important in the control of hormone secretion by the pituitary gland. After traversing capillaries in the hypothalamus, blood enters portal vessels that carry it to the anterior pituitary gland (*adenohypophysis*) and to another set of capillaries (see Figs. 33.16 and 33.17). As blood traverses the hypothalamic capillaries, it picks up several signaling chemicals that control the release of pituitary hormones. When this blood reaches capillaries in the anterior pituitary gland, these substances diffuse out of the bloodstream and into the pituitary interstitial fluid, where they act on pituitary cells to increase or decrease their secretion of specific pituitary hormones. This system is called the *hypothalamic-hypophyseal portal system.*

To summarize, except for a few specialized portal systems, blood encounters only one capillary bed in a single pass through the systemic circulation.

Cardiac Output Is the Volume of Blood Pumped Each Minute by One Ventricle

In a resting dog, it takes about 1 minute for blood to traverse the entire circulation (from the left ventricle back to the left ventricle). Because the pulmonary and systemic circulations are in series, equal volumes of blood must be pumped each minute by the right and left sides of the heart. The volume of blood pumped per minute by either the left ventricle or the right ventricle is called *cardiac output.* Among the mammalian species typically encountered in veterinary medicine, cardiac output at rest is approximately 3 liters per minute per square meter (L/min/m^2) of body surface area. A large dog (e.g., German shepherd) typically has a body surface area a little less than 1 m^2 and a cardiac output at rest of about 2.5 L/min.

In an animal at rest, blood entering the aorta is divided so that approximately 20% of it flows through the splanchnic circulation and 20% through the kidneys. Another 20% goes to the skeletal muscles. The brain receives about 15% of the cardiac output, and the coronary arteries carry about 3% of the cardiac output. The remainder goes to skin and bone.

The Perfusion Pressure for the Systemic Circulation Is Much Greater Than the Perfusion Pressure for the Pulmonary Circulation

When the left ventricle contracts and ejects blood into the aorta, the aorta becomes distended with blood, and aortic blood pressure rises to a peak value called *systolic aortic pressure* (typically 120 mm Hg). Between ejections, blood continues to flow out of the aorta into the downstream arteries. This outflow of blood from the aorta causes aortic pressure to decrease. The minimal value of aortic blood pressure, reached just before the next cardiac ejection, is called *diastolic aortic pressure* (typically 80 mm Hg). A typical appearance of the pressure pulsations in the aorta is shown in the middle panel of Fig. 22.7. The *mean aortic pressure* (average value of the pulsatile blood pressure in the aorta) is about 98 mm Hg. This means that, if an open tube containing mercury were stuck into the aorta, the blood pressure within the aorta would push mercury 98 mm upward into the tube; at which point the upward force of the blood pressure would be equalized by the downward force of gravity acting on the mercury.

The mean aortic pressure represents a potential energy for driving blood through the systemic circulation. As blood flows through the systemic blood vessels, this pressure energy is dissipated through friction. The potential energy (blood pressure) remaining by the time the blood reaches the *venae cavae* is only 3 mm Hg. Therefore the perfusion pressure for the systemic circuit is typically 95 mm Hg (i.e., 98 mm Hg minus 3 mm Hg).

Right ventricular contractions cause pulsatile ejections of blood into the pulmonary artery. The resulting pulsations in pulmonary arterial blood pressure typically have a peak (systolic) value of 20 mm Hg and a minimum (diastolic) value of 8 mm Hg. The mean pulmonary artery blood pressure is typically 13 mm Hg. The blood pressure in pulmonary veins (at the point where they enter the left atrium) is typically 5 mm Hg. Therefore under these conditions, the perfusion pressure for blood flow through the lungs is 8 mm Hg (i.e., 13 mm Hg minus 5 mm Hg).

The same volume of blood (the cardiac output) flows each minute through the systemic circulation and through the lungs; however, as is evident from the typical values just given, the perfusion pressure for the systemic circuit is much greater than the perfusion pressure for the lungs. The reason for this difference in perfusion pressure is that the systemic vessels offer much more friction against blood flow (i.e., have a much higher *resistance*) than do the pulmonary vessels. For this reason, the systemic circulation is referred to as the *high-pressure, high-resistance side of the circulation.* The pulmonary circuit is called the *low-pressure, low-resistance side.*

By convention, blood pressures are always measured with reference to atmospheric pressure. Thus an aortic pressure of 98 mm Hg means that the blood pressure in the aorta is 98 mm Hg higher than the atmospheric pressure outside the body. Also, by convention, blood pressure is measured at heart level. This is why, in human medicine, blood pressure cuffs are typically applied over the brachial artery (in the upper arm), the brachial artery being at the same level as the heart. If blood pressure is measured in an artery or vein at a level different from heart level, an arithmetic correction should be made so that the pressure is reported as if it had been measured at heart level. This correction is necessary

Vet-ebooks.com

because gravity pulls downward on blood and therefore affects the actual pressure of blood within vessels. Gravity increases the actual blood pressure in vessels lying below heart level and decreases the actual pressure in vessels above heart level. The gravitational effect is significant in an animal the size of a dog and substantial in an animal the size of a horse. The correction factor for the effect of gravity is 1 mm Hg for each 1.36 cm above or below heart level.

Each Type of Blood Vessel Has Physical Properties Suited to Its Particular Function

In a resting animal, at any one moment, about 25% of the blood volume is in the central circulation and about 75% is in the systemic circulation (Table 18.1). Most of the blood in the systemic circulation is found in the veins. Only 20% of the systemic blood is found in the arteries, arterioles, and capillaries. Therefore systemic veins are known as the *blood reservoirs* of the circulation. Arteries function as *high-pressure conduits* for rapid distribution of blood to the various organs. Arterioles are the "gates" of the systemic circulation; they *constrict* or *dilate* to control the blood flow to each capillary bed. Although only a small fraction of the systemic blood is found in capillaries at any one time, it is within these *exchange vessels* that the important diffusional transport takes place between the bloodstream and the interstitial fluid.

TABLE 18.1 Distribution of Blood Volume in the Cardiovascular System of a Normal Dog

Distribution	Percent
Between Central and Systemic Circulations	
Central circulation	25
Systemic circulation	75
Total blood volume	100
Within the Various Vessels of the Systemic Circulation	
Arteries and arterioles	15
Capillaries	5
Venules and veins	80
Total systemic blood volume	100

Table 18.2 compares the various types of vessels in the systemic circulation of a dog. As the aorta branches into progressively smaller vessels, the diameters of the vessels become smaller, but the number of vessels increases. One aorta supplies blood to 45,000 terminal arteries, each of which gives rise to more than 400 arterioles. Each arteriole typically branches into about 80 capillaries. The capillaries are so small in diameter that red blood cells must pass through in single file. However, because of the sheer number of capillaries, the total cross-sectional area of the capillaries is much greater than the cross-sectional area of the preceding arteries and arterioles. Because capillary blood flow is spread out over such a large cross-sectional area, the flow velocity within capillaries is low. Blood moves rapidly (about 13 cm/s) through the aorta and large arteries. At this speed, blood is delivered from the heart to all parts of the body in less than 10 seconds. The velocity of blood flow decreases as the blood leaves arteries and enters arterioles and capillaries in each tissue. The velocity of blood flow in capillaries is so slow that blood typically takes about 1 second to travel the short (0.5 mm) length of a capillary. During this time, diffusional exchange takes place between the capillary blood and the interstitial fluid. Blood from the capillaries is collected by venules and veins and is carried quite rapidly back to the heart.

An understanding of the normal dynamics of blood flow provides a basis for interpretation of *capillary refill time,* which is measured during a typical clinical physical examination. The examiner locates an area of nonpigmented epithelial membrane (most commonly a nonpigmented area of the gums). Such tissue is normally pink due to an adequate flow of well-oxygenated blood through the small vessels (arterioles, capillaries, and venules). The examiner applies firm finger pressure to the area for 1 or 2 seconds, which compresses all the small blood vessels and squeezes the blood out of them. Immediately upon release of the finger pressure, the tissue appears very pale due to the absence of blood in the small vessels. A normal circulation will then quickly restore blood flow through the small vessels; the pink color will return within 1 to 2 seconds (the normal capillary refill time). A prolonged capillary refill time is indicative of poor perfusion of the tissue and, by inference, a sluggish circulation.

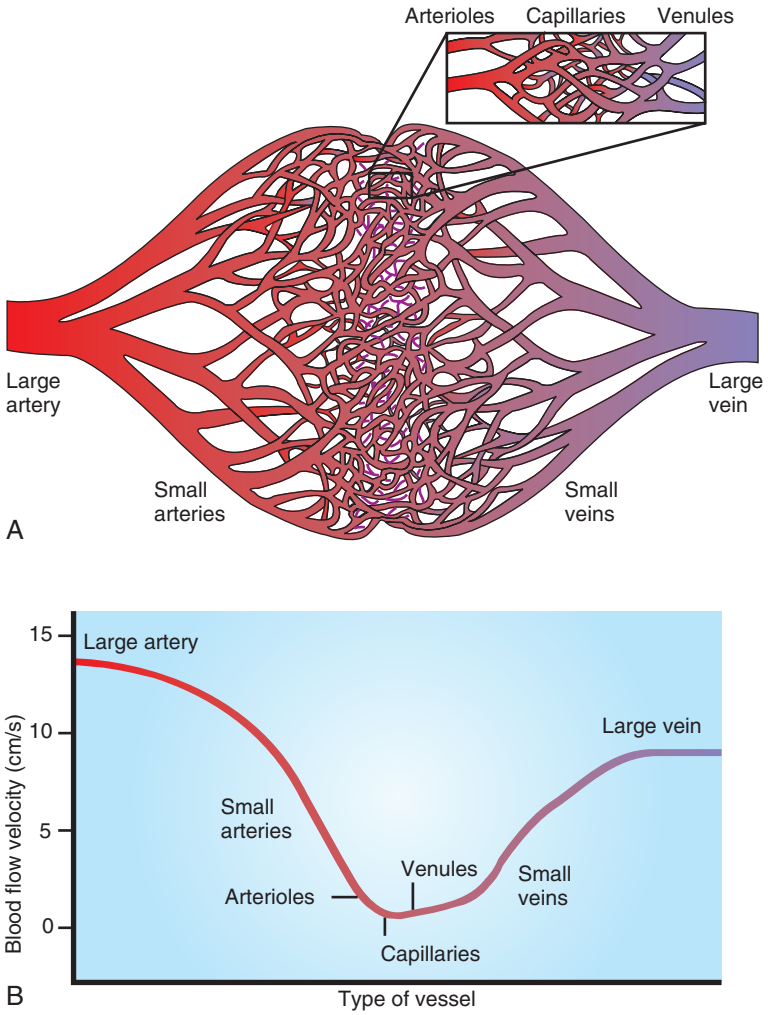
Fig. 18.4 depicts the branching pattern of the systemic vessels and graphs the velocity of blood flow within the different types of vessels. This figure emphasizes the rapidity of bulk flow through large vessels and the relatively slow flow through the capillaries. Note that the *velocity* of blood flow is lowest in the capillaries;

TABLE 18.2 Geometry of Systemic Circulation of a 30-kg Resting Dog

Vessel	Number	Inside Diameter (mm)	Total Cross-Sectional Area (cm^2)	Length (cm)	Velocity of Blood Flow (cm/s)	Mean Blood Pressure (mm Hg)
Aorta	1	20.0	3.1	40.0	13.0	98
Small arteries	45,000	0.14	6.9	1.5	6.0	90
Arterioles	20,000,000	0.030	140.0	0.2	0.3	60
Capillaries	1,700,000,000	0.008	830.0	0.05	0.05	18
Venules	130,000,000	0.020	420.0	0.1	0.1	12
Small veins	73,000	0.27	42.0	1.5	1.0	6
Venae cavae	2	24.0	9.0	34.0	4.5	3

Modified from Minor WR. *Cardiovascular Physiology,* New York: Oxford University Press; 1990.

Vet-ebooks.com



• **Fig. 18.4** As the systemic arteries branch to form small arteries, arterioles, and capillaries (A), the total cross-sectional area of the vessels increases, so the forward velocity of blood flow decreases (B). As blood from the capillaries is collected into venules and veins, the total cross-sectional area is reduced, so the velocity of blood flow increases again. Therefore blood moves quickly from the heart to the microvessels, where it stays for a few seconds before moving rapidly back to the heart.

however, the same *volume* of blood necessarily flows each minute through an artery, the capillaries that it feeds, and the veins draining the capillaries.

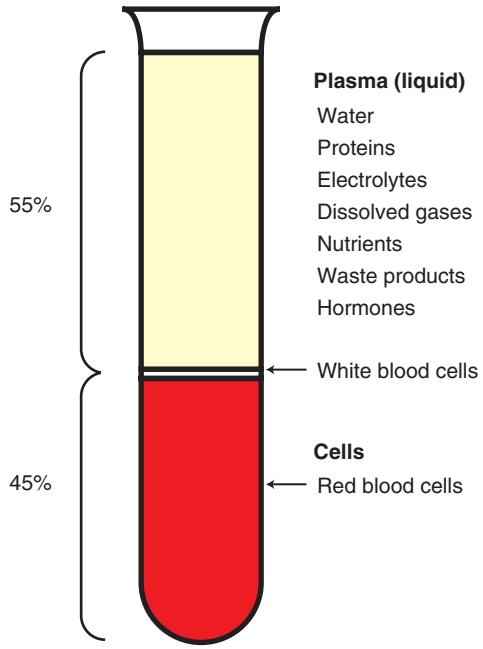
In addition to having a large cross-sectional area (and therefore slow velocity of blood flow), capillaries have a large surface area. The total surface area of the walls of all the capillaries in the systemic circulation of a large dog is about 20 m^2, which is nearly 30 times greater than the dog's body surface area. The large surface area of capillaries helps promote efficient diffusional exchange between the capillary blood and the interstitial fluid.

Blood Is a Suspension of Cells in Extracellular Fluid (Plasma)

As shown in Fig. 18.5, blood can be separated into its cellular and liquid components by centrifugation. The liquid phase of blood is lighter in weight than the cells and therefore ends up on the top of the centrifuge tube. This acellular or extracellular liquid in blood is called *plasma.* Water constitutes 93% of the plasma volume. About 5% to 7% of the plasma volume is made up of protein molecules. The presence of proteins gives plasma its typical pale-yellow color. The *plasma proteins* are synthesized in the liver and are added to the bloodstream as it passes through the liver capillaries. Globulin, albumin, and fibrinogen are the primary plasma proteins. Globulin and albumin are important in the immune responses of the body. Fibrinogen is important in the process of blood clotting. If blood is removed from the body and allowed to stand for a few moments, the soluble fibrinogen molecules polymerize to form an insoluble matrix of fibrin. This causes the blood to congeal, or *coagulate.* Coagulation can be prevented by adding an anticoagulant to the blood; the most common anticoagulants are heparin and citrate. An anticoagulant must be added in preparation for separating blood into its cellular and plasma fractions by centrifugation.

Many important substances, in addition to plasma proteins, are dissolved in plasma. Plasma contains several ions (*electrolytes*) in solution. The dominant cation is sodium (Na^+). The predominant anions are chloride (Cl^-) and bicarbonate (HCO_3^-). Other ions are present in lesser amounts, as indicated in Table 18.3. The concentration of each plasma electrolyte must be kept within narrow limits for body function to be normal, and numerous control systems accomplish this regulation. In general, the plasma electrolytes

Vet-ebooks.com



• **Fig. 18.5** Anticoagulated blood can be separated into an extracellular fluid component (plasma) and a cellular component (cells) by centrifugation. Plasma is a solution of many important substances in water. The presence of proteins gives plasma its typical pale-yellow color. The cells are heavier than the plasma, and they settle to the bottom. Most of the cells are red blood cells. The white blood cells are slightly lighter in weight than the red blood cells, and they form a thin *buffy coat* on the top of the red cell layer. Most of the platelets end up in the buffy coat, although at slow centrifuge speeds ("soft spin"), platelets tend to remain suspended in the plasma. The fraction of cells in blood is called the hematocrit. In this example, the hematocrit is 45%.

can diffuse readily across capillary walls; therefore interstitial fluid and plasma typically have similar electrolyte concentrations.

Plasma contains small amounts of gases (O_2, CO_2, and N_2) in solution. In the lungs, O_2 enters the blood as dissolved O_2, but most of this O_2 quickly combines with hemoglobin (in the red blood cells). As a consequence, about 98.5% of the total O_2 in blood is carried as *oxyhemoglobin* and only about 1.5% as dissolved O_2. Likewise, only a small portion of the carbon dioxide (CO_2) in blood is carried in the dissolved form. Most of the CO_2 becomes hydrated to form HCO_3^- or combines with hemoglobin or plasma proteins to form *carbamino compounds.*

Nutrient substances dissolved in plasma include glucose, amino acids, lipids, and some vitamins. Dissolved metabolic waste products (in addition to CO_2) include urea, creatinine, uric acid, and bilirubin. Plasma also contains many hormones (e.g., insulin, epinephrine, thyroxine), which are present in exceedingly tiny, but critically important amounts. Table 18.3 lists some of the normal constituents of plasma.

TABLE 18.3 Some Constituents of Canine Plasma (in Addition to Water, the Main Constituent)

Component	Normal Range	Units
Plasma Proteins (Carried in Colloidal Suspension)		
Globulin (total)	2.7–4.4	g/dL
Albumin	2.3–3.1	g/dL
Fibrinogen	0.15–0.30	g/dL
Electrolytes (Dissolved)		
Na^+	140–150	mmol/L
K^+	3.9–5.1	mmol/L
Ca^{2+} (ionized)	1.2–1.5	mmol/L
Mg^{2+} (ionized)	0.5–0.9	mmol/L
Cl^-	110–124	mmol/L
HCO_3^-	17–24	mmol/L
HPO_4^{2-} and $H_2PO_4^-$	1–1.4	mmol/L
H^+	38–49	nmol/L[a]
(H^+ expressed as pH)[b]	(7.31–7.42)	
Dissolved Gases (Values for Arterial Plasma)		
O_2	0.26–0.30	mL/dL
CO_2	2–2.5	mL/dL
Examples of Nutrients, Waste Products, Hormones		
Cholesterol	140–280	mg/dL
Glucose	76–120	mg/dL
Triglycerides	40–170	mg/dL
Urea nitrogen	8–28	mg/dL
Creatinine	0.5–1.7	mg/dL
Bile acids (fasting)	0–8	µmol/L
Thyroxine (T_4)	1.5–4	nmol/L[a]

[a]Note that [H^+] and [Thyroxine] are in nanomolar units.
1000 nmol = 1 µmol, and 1000 µmol = 1 mmol.
[b]pH = –log [H^+], where [H^+] is expressed in molar units; pH is dimensionless.
Modified from Latimer KS, Mahaffey EA, Prasse KW. *Duncan & Prasse's Veterinary Laboratory Medicine: Clinical Pathology.* 4th ed. Ames, Iowa: Wiley-Blackwell; 2003.

The Cellular Component of Blood Includes Red Blood Cells, White Blood Cells, and Platelets

Cells normally constitute 30% to 60% of the blood volume (depending on the species). The fraction of cells in blood is called the *hematocrit* (see Fig. 18.5). The hematocrit is determined by adding an anticoagulant to some blood and then centrifuging it in a tube. The cells are somewhat heavier than plasma and settle to the bottom of the tube during centrifugation. Because centrifugation results in a packing of the blood cells in the bottom of the tube, the hematocrit is sometimes called the *packed cell volume.* Most of the cell component looks red because most of the blood cells are *erythrocytes* (red blood cells, or RBCs). Erythrocytes acquire their red color from hemoglobin.

The *leukocytes* (white blood cells, or WBCs) are slightly lighter in weight than the RBCs; in a centrifuge tube, the WBCs gather in a white *buffy coat* on top of the RBCs. The buffy coat is normally very thin because there are about 1000 times more RBCs than WBCs. Leukocytes are critical in immune and allergic responses of the body. The subtypes of leukocytes include neutrophils, lymphocytes, monocytes, eosinophils, and basophils. A laboratory analysis of the total number and relative distribution of the various WBC subtypes (*differential WBC count*) provides important clues

Vet-ebooks.com

in the diagnosis of disease. Both erythrocytes and leukocytes are made in the bone marrow. They develop, by mitosis and differentiation, from a common line of progenitor cells, the *pluripotent* (uncommitted) *stem cells*.

The *platelets,* or *thrombocytes,* in blood are cellular fragments from their precursor cells, the *megakaryocytes*. The megakaryocytes reside in the bone marrow, and they shed bits of their cytoplasm, bounded by cell membrane, into the bloodstream. Platelets participate in *hemostasis* (the control of blood loss from injured or severed blood vessels). In this process, a clumping together of platelets (*platelet aggregation*) begins to create a physical barrier across openings in blood vessels. The platelets also release the substance *serotonin,* which causes the blood vessels to constrict, thereby reducing blood pressure and blood flow at the site of injury. Additional substances released from the platelets interact with fibrinogen and several clotting factors in the plasma, which leads to the coagulation of blood and the formation of a stable, fibrin-based blood clot.

Coagulation and clotting involve complex, interconnected sequences of chemical reactions (*the coagulation cascade*). A key step in the coagulation cascade is the formation in the plasma of *thrombin,* an enzyme that catalyzes the transformation of fibrinogen to fibrin. Several laboratory tests are used to assess the status of an animal's coagulation system. Two common tests involve determination of the *prothrombin time* (PT) and the *partial thromboplastin time* (PTT).

If blood is allowed to coagulate and then is centrifuged, the fibrin and other plasma-clotting factors settle to the bottom along with the RBCs, WBCs, and platelets. The liquid portion remaining above (essentially plasma without fibrinogen and other clotting factors) is called *serum.* Most of the common clinical blood chemistry analyses are performed on serum. Examples include the determination of concentrations of electrolytes and cholesterol.

If blood is treated with an anticoagulant and then allowed simply to sit in a tube (without centrifugation), the erythrocytes slowly begin to settle. For reasons that are not completely understood, the rate of their settling tends to be increased to above normal in certain disease states and decreased to below normal in others. Therefore the *erythrocyte sedimentation rate* (ESR) is a clinically useful diagnostic measurement. An important caveat is that the normal ESR varies substantially between species; for example, it is much more rapid in equine blood than in canine blood.

Blood cell counts are performed by manual or automated scanning of a very small volume (e.g., 1 μL) of anticoagulated whole blood. Table 18.4 presents a summary of normal hematologic values for the dog.

TABLE 18.4 Canine Hematology

Test	Normal Range	Units
Hematocrit	35–57	%
Blood Cell Counts		
Red blood cells	5000–7900	$\times 10^3/\mu L$
White blood cells	5–14	$\times 10^3/\mu L$
Platelets	210–620	$\times 10^3/\mu L$
Hemoglobin Measures		
Blood hemoglobin	12–19	g/dL
MCH (mean corpuscular hemoglobin)	21–26	pg[a]
MCHC (mean corpuscular hemoglobin concentration)	32–36	g/dL

[a]Note that MCH is expressed in pg (picogram) units. 1 pg = 10^{-12} g.
Modified from Latimer KS, Mahaffey EA, Prasse KW. *Duncan & Prasse's Veterinary Laboratory Medicine: Clinical Pathology.* 4th ed. Ames: Wiley-Blackwell; 2003.

Most of the Oxygen in Blood Is Carried in Chemical Combination With the Protein Hemoglobin Within Red Blood Cells

Of the 20 mL of O_2 normally carried in each 100 mL of oxygenated blood, only 1.5% (0.3 mL) is carried in dissolved form. The remaining 98.5% is carried in chemical combination with hemoglobin (in RBCs). *Oxygenated hemoglobin* (*oxyhemoglobin,* HbO_2) is bright red. When O_2 is released, HbO_2 becomes *reduced hemoglobin* (Hb), which is dark bluish red. The adequacy of oxygenation of an animal's blood can be judged somewhat by looking at the color of its nonpigmented epithelial membranes (e.g., gums, nostrils, or inside surfaces of eyelids). Well-oxygenated tissues appear pink. Poorly oxygenated tissues appear bluish (*cyanotic*) because of the prevalence of reduced Hb.

The ability of blood to carry oxygen is determined by the amount of hemoglobin in the blood and by the chemical characteristics of that Hb. For example, each deciliter (dL) of normal dog blood contains about 15 g of Hb. Each gram of Hb can combine with 1.34 mL of O_2 when fully saturated. Thus each deciliter of fully oxygenated, normal blood carries 20 mL of O_2. Several disease states (*hemoglobinopathies*) result in the synthesis of chemically abnormal Hb, with a diminished capacity to bind O_2. Also, several common toxins, including carbon monoxide (CO) and nitrates, cause life-threatening alterations in the ability of Hb to bind O_2.

Because hemoglobin is localized inside RBCs, it is possible to derive several clinically useful relationships among the blood Hb content, RBC count, Hb content of each RBC, and hematocrit. For example, if a normal dog has 15 g of Hb in each deciliter of blood and an RBC count of 6 million cells per microliter (μL) blood, it follows that each RBC (on average) contains 25 picograms (pg) of Hb:

$$\frac{15 \text{ g of hemoglobin/dL of blood}}{6\times10^6 \text{ red blood cells/}\mu\text{L of blood}} = 25\times10^{-12} \text{ g of Hb/RBC}$$

The value calculated in this way is called the *mean corpuscular hemoglobin* (MCH).

An easier calculation, which serves the same purpose, is to determine how much hemoglobin is contained in each deciliter of packed RBCs. For example, if a dog has 15 g Hb/dL of blood and has a hematocrit of 50%, the Hb concentration in the RBC portion of the blood must be 30 g of Hb/dL of packed RBCs:

$$\frac{15 \text{ g of hemoglobin/dL of blood}}{0.5 \text{ dL of red blood cells/dL of blood}} = 30 \text{ g of Hb/dL of RBCs}$$

The value calculated in this way is called the *mean corpuscular hemoglobin concentration* (MCHC). For simplicity, the calculation is often summarized as follows:

$$\text{MCHC} = [\text{hemoglobin}]/\text{hematocrit}$$

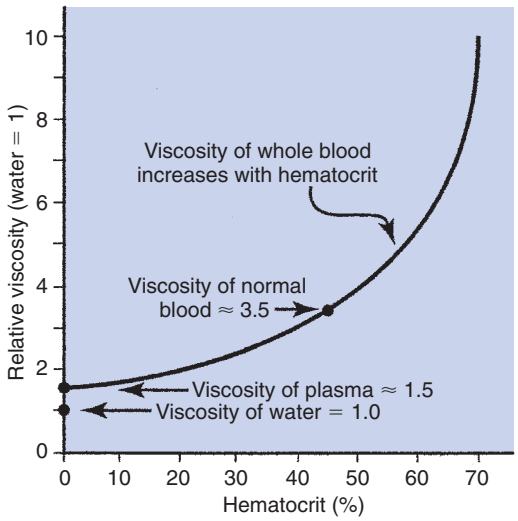
The brackets around "hemoglobin" denote concentration.

An abnormally low value of MCH or MCHC is clinically important because it points to a deficit in hemoglobin synthesis (i.e., not enough Hb being made to load up each RBC). In contrast,

Vet-ebooks.com

an abnormally low value for hemoglobin concentration by itself is less informative; [Hb] in the blood could fall below normal for several reasons, including a deficit in Hb synthesis, a deficit in RBC synthesis, or a "watering down" of the blood either by addition of excess plasma fluid or by loss of RBCs.

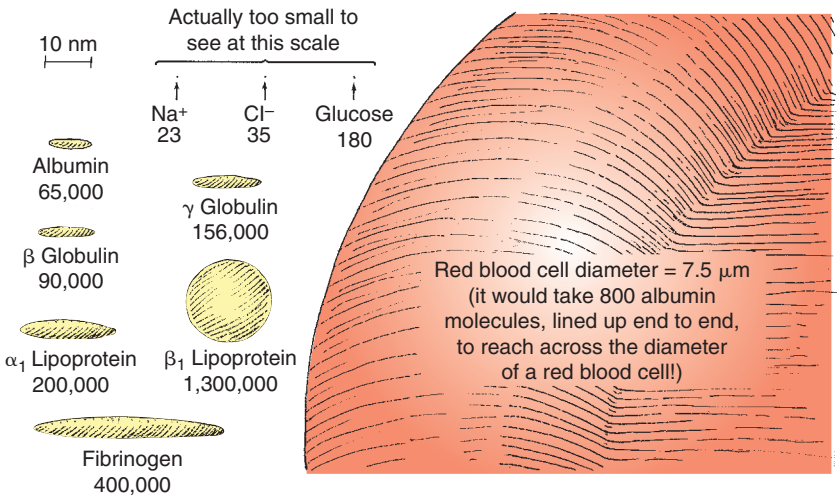
Deviations from a normal hematocrit (Hct) have important consequences in terms of the ability of blood to carry oxygen. But hematocrit also affects the viscosity of blood, as shown in Fig. 18.6. *Viscosity* is a measure of resistance to flow. For example, honey is more viscous (more resistant to flow) than water. Plasma, by itself, is about 1.5 times more viscous than water because of the presence of plasma protein molecules (albumin, globulin, fibrinogen). The presence of cells in blood has an even greater effect on viscosity. Blood with an Hct of 40% has twice the viscosity of plasma. For Hct exceeding 50%, viscosity increases rapidly. An abnormally high hematocrit is called *polycythemia,* which literally means "many cells in the blood." The blood of a patient with polycythemia can carry more than the normal 20 mL of O_2/dL of blood (provided that the MCHC is normal), and this may be viewed as beneficial. However, the increased viscosity makes it difficult for the heart to pump polycythemic blood. Therefore polycythemia creates a heavy workload for the heart and can lead to heart failure, particularly if the cardiac muscle is not healthy.



• **Fig. 18.6** Plasma is more viscous than water because of the presence of plasma proteins. Blood is more viscous than plasma because of the presence of blood cells. Blood viscosity increases sharply when the fraction of cells (hematocrit) increases above 50%.

The opposite problem, in which the hematocrit is too low, is called *anemia.* Anemia literally means "no blood," but the word is used to refer to any condition in which there are abnormally few RBCs in each deciliter or a condition in which there is an abnormally low hemoglobin concentration in each RBC (i.e., MCH and/or MCHC is low). Each deciliter of blood of an anemic patient carries less than the normal 20 mL of O_2. Therefore cardiac output must be increased above normal to deliver the normal amount of O_2 to the tissues each minute. The necessity to increase cardiac output also imposes an increased workload on the heart and can lead to the failure of a diseased heart. Thus Hct within the normal range provides the blood with enough Hb to carry an adequate amount of O_2 without putting an undue workload on the heart. For additional information about the transport of O_2 and CO_2 in blood, see Chapter 48.

Fig. 18.7 provides an idea of the relative sizes and shapes of the major constituents of blood. The plasma proteins are much, much larger than the ions and nutrient molecules dissolved in plasma. RBCs and WBCs are many, many times larger than the plasma proteins. In fact, as mentioned earlier, blood cells are so large that they can barely squeeze through a typical capillary.



• **Fig. 18.7** Relative size and shape of the major constituents of blood. The figure emphasizes two points: first, that the plasma protein molecules are huge compared with the other plasma solutes, such as glucose, Na^+, and Cl^-; and second, that the blood cells (red and white) are huge compared with plasma protein molecules. Numbers under constituents are their molecular weights (in daltons). The scale (*upper left*) indicates a length of 10 nm. In comparison, the diameter of the red blood cell is 7.5 μm, which is 750 times larger than the scale marker.

Vet-ebooks.com

CLINICAL CORRELATIONS

Lethargic Kid Goat

History

A 6-month-old female kid goat is presented for lethargy and difficulty breathing. Two months ago, in April, the owners bought this goat and another at a sale as pets for their children. The goats have been provided with a small amount of goat feed daily, along with access to a pasture. The owners noticed that both goats were initially very playful, but both have seemed progressively lethargic during the last month. Also, they seem to have more difficulty breathing, even at rest. No vaccinations, deworming, or other treatments have been given.

Clinical Examination

The goat is somewhat thin and is reluctant to stand. There is a swelling (likely edema fluid) under the jaw. The goat's temperature is slightly elevated. The pulse and respiratory rates are moderately increased. The mucous membranes are very pale, which makes the capillary refill time difficult to assess. Respiratory sounds are increased (suggesting possible pulmonary edema). There are no other abnormal findings on physical examination.

Comment

The very pale mucous membranes suggest marked anemia. Indeed, centrifugation of a blood sample reveals that the goat's packed cell volume (Hct) is only 12%. Plasma protein concentration is also below normal at 4.5 g/dL. Given the lack of deworming, you suspect parasitical infection associated with *Haemonchus contortus, Ostertagia,* or *Trichostrongylus.* A fecal analysis is positive for *Haemonchus* and *Ostertagia.*

Parasitism is a common problem in sheep and goats. The parasites mentioned damage the abomasum, which results in blood loss. The consequent anemia would explain the goat's lethargy, because anemia limits O_2 delivery to the organs, especially during exercise. The elevated respiratory rate and heart rate reflect the animal's attempts to compensate for low O_2 delivery to the tissues by increasing air flow into the lungs and blood flow through the circulation. Plasma protein is lost along with red blood cells (RBCs). This *hypoproteinemia* could account for the edema, because the proteins in plasma exert an important osmotic effect to oppose the tendency for plasma water to leak out of capillaries and into the tissue (interstitial) fluid (see Chapter 23 for details).

Treatment

Ideally, a transfusion of whole blood would be given to help restore both RBCs and plasma proteins; the kid would then be dewormed. However, even if appropriate whole blood were available, transfusion in such an animal is risky. This goat's ability to deal with stress has been severely compromised, and even the physical restraint needed to administer a transfusion might trigger physical collapse or even death. On the other hand, without the transfusion, the animal has little chance of recovery if only treated for the parasites.

Colic and Endotoxic Shock in Horse Secondary to Strongylus Parasitism

History

A 1-year-old Standardbred filly is brought to your clinic by its new owner because the horse has been restless, rolling, kicking at its belly, and pawing the ground. The owner reports that the horse has had a poor appetite for several days and now refuses both hay and grain. The owner says he has dewormed the filly recently, but her previous deworming history is unknown.

Clinical Examination

The horse is underweight and has a dull hair coat. It is obvious that she is in pain. Physical examination reveals an abnormally high temperature (103.5° F), rapid labored breathing (40 breaths/min), and an elevated heart rate (80 beats/min). All limbs feel cool to the touch. The mucous membranes are abnormally dark, and the capillary refill time is prolonged (both these observations indicate sluggish circulation). Gastrointestinal auscultation of all four quadrants yields abnormal findings; no gastrointestinal *borborygmus* is heard on either the left or the right side, dorsally or ventrally. A rectal examination reveals several distended loops of bowel.

You perform *abdominocentesis* and withdraw some peritoneal fluid. Normally, peritoneal fluid is clear and straw-colored; the fluid from this horse is darker yellow than normal and has a turbid appearance. Measurements with a refractometer indicate that the peritoneal fluid contains five times more protein than normal. Microscopical examination of the fluid reveals the presence of four times the normal number of white blood cells (WBCs), specifically neutrophils, and the cells contain bacteria.

Outcome

You tell the owner that the filly appears to have a badly damaged bowel and that the prognosis is poor. You inform him that surgical treatment is possible, but expensive postoperative complications are likely because infection appears to have spread into the peritoneum. After considering the options, the owner decides against surgery. You institute supportive therapy with intravenous (IV) fluids, analgesics, and antibiotics. Depending on the extent of compromise to the bowel, horses can respond to medical management. However, based on the signs that this filly is showing, including that the filly already has signs of peritonitis, the prognosis is grave.

The horse's condition deteriorates over the next 12 hours. The heart rate increases progressively to 100 beats/min. The mucous membranes show evidence of declining blood flow (darker color and longer capillary refill time). The horse begins to wheeze and becomes lethargic. Bowel sounds continue to be absent. Despite the delivery of IV fluids, there is no output of urine. With the owner's consent, you euthanize the horse.

Necropsy examination indicates that this horse had thrombi (vascular obstructions) in several major branches of her mesenteric arteries, probably secondary to a severe infestation of bloodworms (*Strongylus vulgaris*). Several areas of the intestine were necrotic. Gram-negative bacteria were cultured from both the peritoneal fluid and the blood. The lungs were edematous, and excessive fluid was found in the airways and intrapleural space.

Comment

In horses, *S. vulgaris* lodges in mesenteric arteries and decreases the blood flow to the intestine. Deworming a severely infested horse can precipitate acute intestinal ischemia, because the dead/dying worms break away from the walls of major mesenteric arteries and drift into smaller arteries, which they occlude. Also, the dying worms release substances that trigger the formation of blood clots in the arteries. Digestive processes become disrupted and may cease entirely. Intestinal ischemia and gaseous distention of the bowel cause severe pain. With persistent ischemia, segments of the bowel suffer irreversible damage. Ischemic damage to the intestinal epithelium allows intestinal bacteria and bacterial products (endotoxins) to enter the peritoneum and the blood. WBCs move from the bloodstream into the peritoneal fluid, where they combat the bacteria by engulfing them (phagocytosis). However, the infection overwhelms the immune system. Bacteria and endotoxins (from gram-negative bacteria) cause the body to produce substances that depress the heart and disrupt the capillary endothelium, especially in the lungs. The resultant combination of heart failure and pulmonary edema leads to respiratory failure and subsequent renal failure. The progression of dysfunction becomes irreversible.

Vet-ebooks.com

Practice Questions

1. What is the name of the condition when blood flow to a tissue is inadequate to meet the metabolic needs of the tissue?
- **a.** Anemia
- **b.** Edema
- **c.** Necrosis
- **d.** Ischemia
- **e.** Infarction

2. The volume of blood pumped by the left ventricle in 1 minute would equal:
- **a.** The volume of blood that flowed through the coronary circulation (in the same minute).
- **b.** The volume of blood that flowed through all organs of the systemic circulation, except for coronary blood flow.
- **c.** The volume of blood that flowed through the lungs.
- **d.** One-half of the cardiac output.
- **e.** Two times the cardiac output.

3. A transfusion of normal plasma into a normal dog would:
- **a.** Decrease the hematocrit of the recipient's blood.
- **b.** Increase the viscosity of the recipient's blood.
- **c.** Decrease the mean corpuscular hemoglobin concentration (MCHC) in the recipient's plasma.
- **d.** Increase the number of cells in the recipient's blood.
- **e.** Decrease the concentration of proteins in the recipient's plasma.

4. Which of the following sequences of capillary beds might a red blood cell encounter in a normal circulation?
- **a.** Lungs, skin, lungs, brain
- **b.** Spleen, liver, mesentery, lungs
- **c.** Coronary, kidney (glomerular), kidney (tubular), lungs
- **d.** Lungs, coronary, stomach, liver
- **e.** Brain, lungs, liver, coronary

5. The walls of most capillaries have pores or clefts in them, which are approximately 4 nm (4×10^{-9} m) in diameter. According to Fig. 18.7:
- **a.** A capillary pore is many times larger in diameter than a sodium ion.
- **b.** A capillary pore is many times larger in diameter than a glucose molecule.
- **c.** The diameter of a red blood cell is many times greater than the diameter of a capillary pore.
- **d.** A molecule of β globulin or γ globulin could just about squeeze through a capillary pore if it were lined up exactly right.
- **e.** All of the above are correct.

6. Suppose that the following conditions exist in a particular blood vessel: blood pressure (BP) inside vessel at inlet = 60 mm Hg, BP inside vessel at midpoint = 50 mm Hg, and BP inside vessel at outlet = 40 mm Hg. Pressure in the tissue fluid (interstitial fluid) immediately outside vessel = 5 mm Hg. Under these conditions:
- **a.** Perfusion pressure for blood flow through this vessel = 20 mm Hg.
- **b.** Perfusion pressure for blood flow through this vessel = 15 mm Hg.
- **c.** Distending pressure at the vessel midpoint = 55 mm Hg.
- **d.** Distending pressure at the vessel midpoint = 45 mm Hg.
- **e.** Both a and d are correct.

7. Compared with the systemic circulation, the pulmonary circulation:
- **a.** Carries more blood flow per minute.
- **b.** Has a lower perfusion pressure.
- **c.** Has a higher resistance to blood flow.
- **d.** Carries blood that has a lower hematocrit.
- **e.** Contains a higher blood volume.

Bibliography

Bowman DD. *Georgis' Parasitology for Veterinarians*. 10th ed. Elsevier; 2013.

Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. Elsevier; 2017.

Hill RW, Wyse GA, Anderson M. *Animal Physiology*. 4th ed. Sinauer; 2016.

Kaneko JJ, Harvey JW, Bruss ML. *Clinical Biochemistry of Domestic Animals*. 6th ed. Academic Press/Elsevier; 2008.

Klabunde RE. *Cardiovascular Physiology Concepts*. 2nd ed. Lippincott Williams & Wilkins; 2011.

Kumar V, Abbas AK, Aster J. *Robbins Basic Pathology*. 10th ed. Elsevier; 2017.

Latimer KS, ed. *Duncan & Prasse's Veterinary Laboratory Medicine: Clinical Pathology*. 5th ed. Wiley-Blackwell; 2011.

Milnor WR. *Cardiovascular Physiology*. New York: Oxford University Press; 1990.

Mohrman DE, Heller LJ. *Cardiovascular Physiology*. 8th ed. McGraw-Hill; 2013.

Patteson MW. *Equine Cardiology*. Oxford, UK: Blackwell Science; 1996.

Physick-Sheard PW. Parasitic arteritis. In: Colahan PT, Merritt AM, Moore JN, et al, eds. *Equine Medicine and Surgery*. Vol. 1. 5th ed. St Louis: Mosby–Year Book; 1999.

Pugh DG, Baird N. *Sheep and Goat Medicine*. 2nd ed. Elsevier; 2011.

Reece WO, Erickson HH, Goff JP, eds. *Dukes' Physiology of Domestic Animals*. 13th ed. Wiley-Blackwell; 2015.

Reed SM, Bayly WM, Sellon DC. *Equine Internal Medicine*. 4th ed. Elsevier; 2017.

Smith BP, ed. *Large Animal Internal Medicine*. 5th ed. Elsevier; 2014.

Thrall MA, Weiser G, Allison R, Campbell T, eds. *Veterinary Hematology and Clinical Chemistry*. 2nd ed. Wiley-Blackwell; 2012.

Weiss DJ, Wardrop KJ. *Schalm's Veterinary Hematology*. 6th ed. Ames, Iowa: Wiley-Blackwell; 2010.

Zachary JF, ed. *Pathologic Basis of Veterinary Disease, Expert Consult*. 6th ed. Elsevier; 2016.

Vet-ebooks.com

19

Electrical Activity of the Heart

ROBERT B. STEPHENSON

KEY POINTS

1. In heart muscle, an action potential causes a contraction (no action potential, no heartbeat).
2. The contractile machinery in cardiac muscle is similar to that in skeletal muscle.
3. Cardiac muscle forms a functional syncytium.
4. Cardiac contractions are initiated by action potentials that arise spontaneously in specialized pacemaker cells.
5. A system of specialized cardiac muscle cells initiates and organizes each heartbeat.
6. Cardiac action potentials are extremely long.
7. Membrane calcium channels play a special role in cardiac muscle.
8. The long duration of the cardiac action potential guarantees a period of relaxation (and refilling) between heartbeats.
9. Atrial cells have shorter action potentials than ventricular cells.
10. Specialized ion channels cause cardiac pacemaker cells to depolarize to threshold and form action potentials.
11. Sympathetic and parasympathetic nerves act on cardiac pacemaker cells to increase or decrease the heart rate.
12. Cells of the atrioventricular node act as auxiliary pacemakers and stabilize ventricular rate.
13. Sympathetic nerves act on all cardiac muscle cells to cause quicker, more forceful contractions.
14. Parasympathetic effects are opposite to those of sympathetic activation but are most evident in the sinoatrial node, atrioventricular node, and atria.
15. Dysfunction in the specialized conducting system leads to abnormalities in cardiac rhythm (arrhythmias).
16. Atrioventricular node block is a common cause of cardiac arrhythmias.
17. Cardiac tachyarrhythmias result from abnormal pacemaker activity.
18. Antiarrhythmic drugs affect the ion channels responsible for cardiac action potential formation and propagation.

In Heart Muscle, an Action Potential Causes a Contraction (No Action Potential, No Heartbeat)

The heart is a muscular pump that propels blood through the blood vessels by alternately relaxing and contracting. As the heart muscle relaxes, the atria and ventricles fill with venous blood. During cardiac contraction, some of this blood is ejected into the arteries. Cardiac contraction takes place in two stages: (1) the right and left atria begin to contract, and (2) after a delay of about 0.1 seconds, the right and left ventricles begin to contract. Atrial contraction helps to finish filling the ventricles with blood. The delay allows time for this "topping up" of ventricular volume. Ventricular contraction ejects blood out of the left ventricle into the aorta and out of the right ventricle into the pulmonary artery. After the atria and ventricles contract, they relax and begin to refill. The entire contractile sequence is initiated and organized by an electrical signal, an *action potential,* which propagates from muscle cell to muscle cell, through the heart.

This chapter begins with a brief description of how cardiac muscle contracts followed by a detailed description of the action potentials that initiate and organize the heart's contractions. Several common electrical dysfunctions of the heart are then discussed.

Throughout this chapter, comparisons are made between cardiac and skeletal muscle (Table 19.1). In both cardiac and skeletal muscle, an electrical action potential in each muscle cell is necessary to trigger a contraction. The molecular mechanisms that carry out the contraction are also similar in both types of muscle. However, important differences exist between cardiac and skeletal muscle in the characteristics of the action potentials that initiate contractions.

The Contractile Machinery in Cardiac Muscle Is Similar to That in Skeletal Muscle

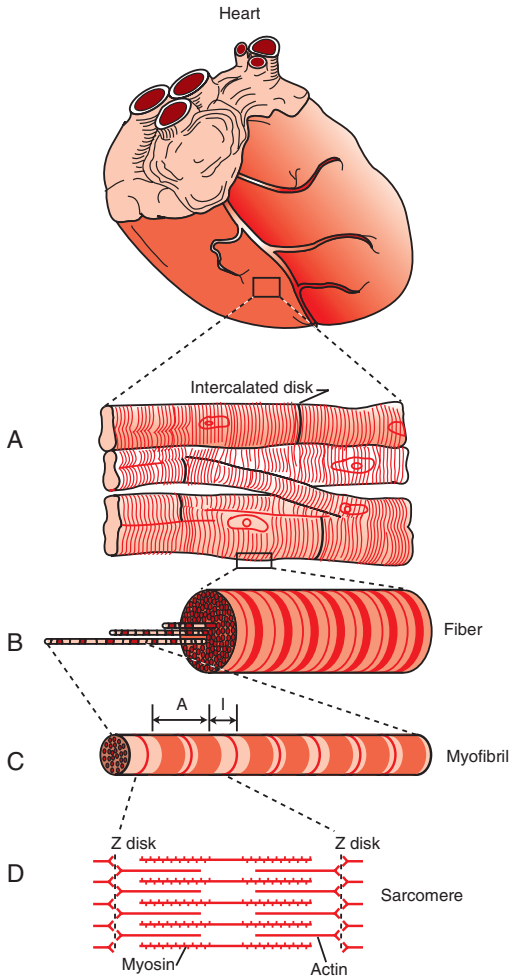
Cardiac muscle, like skeletal muscle, has a *striated* appearance under the light microscope (Fig. 19.1). These cross-striations have the same structural basis in cardiac muscle as in skeletal muscle (see Fig. 6.2). Each striated cardiac muscle cell (*muscle fiber*) is made up of a few hundred myofibrils. Each *myofibril* has a repetitive pattern of light and dark bands. The various bands within a myofibril are given letter designations (A band, I band, Z disk). The alignment of these bands in adjacent myofibrils accounts for the striated appearance of the whole muscle fiber. Each repeating unit of myofibrillar bands is called a *sarcomere.* This name, which means "little muscle," is apt because a single sarcomere constitutes the contractile subunit of the cardiac muscle. By definition, a sarcomere extends from one Z disk to the next, a distance of approximately 2 µm (0.002 mm). About 50 sarcomeres comprise a myofibril, and each myofibril extends for the length of the cardiac muscle cell.

As in skeletal muscle, each cardiac muscle sarcomere is composed of an array of thick and thin filaments. The *thin filaments* are attached to the Z disks; they interdigitate with the thick filaments. The thin filaments are composed of *actin* molecules. The *thick filaments* are composed of *myosin* molecules. In the presence of adenosine triphosphate (ATP) and calcium ions (Ca^{2+}), myosin and actin interact in a sequence of steps called the *cross-bridge*

Vet-ebooks.com

TABLE 19.1 Sequence of Events in Contraction of Skeletal Muscle and Cardiac Muscle

Skeletal Muscle	Cardiac Muscle
Action potential is generated in somatic motor neuron	*Note:* Action potentials in autonomic motor neurons are not needed to initiate heartbeats
Acetylcholine is released	*Note:* Neurotransmitters are *not* needed to make the heart beat
Nicotinic cholinergic receptors on muscle cell membrane are activated	*Note:* Activation of receptors is *not* needed—a completely isolated or denervated heart still beats
Ligand-gated Na^+ channels in muscle membrane open	Pacemaker Na^+ channels spontaneously open (and K^+ channels close) in membranes of pacemaker cells
Muscle membrane depolarizes to threshold level for formation of action potential	Pacemaker cell membranes depolarize to threshold for formation of action potential
Action potential forms in muscle cell but does not enter other cells	Action potential forms in a pacemaker cell and then propagates from cell to cell throughout the whole heart
Note: Skeletal muscle cells *do not* have slow Ca^{2+} channels	During action potential, extracellular Ca^{2+} ("trigger" Ca^{2+}) enters cell through "slow" Ca^{2+} channels
Action potential causes Ca^{2+} release from sarcoplasmic reticulum; Ca^{2+} binds to troponin	Entry of extracellular trigger Ca^{2+} causes release of more Ca^{2+} from sarcoplasmic reticulum; Ca^{2+} binds to troponin
Actin's binding sites are made available for actin-myosin cross-bridge formation	Actin's binding sites are made available for actin-myosin cross-bridge formation
Cross-bridge cycling generates contractile force between actin and myosin filaments	Cross-bridge cycling generates contractile force between actin and myosin filaments
Muscle contracts (brief "twitch"); Ca^{2+} is taken up by sarcoplasmic reticulum	Heart contracts (complete "beat" or "systole"); Ca^{2+} is taken up by sarcoplasmic reticulum or pumped back out of cell into extracellular fluid
Muscle relaxes	Heart relaxes



• **Fig. 19.1** **A**, If a section of heart muscle is viewed under light microscopy, cardiac muscle fibers (muscle cells) are seen to be striated, similar to skeletal muscle. Unlike skeletal muscle fibers, however, cardiac muscle fibers often branch, and they are attached end to end with neighboring fibers at structures called intercalated disks. Unseen within the intercalated disks are gap junctions (minute channels, filled with intracellular fluid), which allow action potentials to propagate from cell to cell. **B–D**, Electron microscopy reveals that the striations in cardiac muscle (like striations in skeletal muscle) result from an orderly arrangement of actin (thin) filaments and myosin (thick) filaments into muscular subunits called sarcomeres. Sarcomeres are the structural and functional subunit of cardiac muscle (just as for skeletal muscle).

cycle, which results in contraction and force generation in each sarcomere and therefore in the whole muscle cell, just as in skeletal muscle (for details, see Figs. 1.3–1.5, and 6.6).

Cardiac Muscle Forms a Functional Syncytium

Although the molecular basis of contraction is the same for cardiac and skeletal muscle, the two muscle types differ in regard to electrical linkages between neighboring cells, and this difference has important consequences. Individual skeletal muscle cells are electrically isolated (insulated) from one another, so action potentials cannot "jump" from one skeletal muscle cell to another. As described in Chapter 5, an action potential in a skeletal muscle cell is initiated only in response to an action potential in the somatic motor neuron that innervates the skeletal muscle cell. Each neural action potential causes release of the neurotransmitter acetylcholine, which activates nicotinic cholinergic receptors on the skeletal muscle cell, which in turn depolarizes the muscle cell to threshold for the formation of an action potential. When formed, the action potential propagates along the length of that particular muscle cell and then stops. The muscle action potential causes the cell to contract. Neighboring cells may or may not contract at the same time, depending on whether action potentials are initiated in the neighboring cells by their motor neurons.

In contrast, cardiac muscle cells are electrically linked to one another. Specialized junctions between cardiac muscle cells allow ionic currents created by an action potential in one cell to flow into the neighboring cells and initiate action potentials in those

Vet-ebooks.com

cells too. Because a cardiac action potential propagates from cell to cell through cardiac tissue, neighboring cardiac muscle cells all contract in synchrony, as a unit, and then they all relax. In this regard, cardiac muscle tissue behaves as if it were a single cell. Cardiac muscle is therefore said to form a *functional syncytium* (literally, "acts like same cell").

When viewed with light microscopy, cardiac muscle appears as an array of striated fibers (individual cardiac muscle cells) arranged approximately in parallel but with some branching (see Fig. 19.1). Dark bands (called *intercalated disks*) are evident at sites where neighboring cells attach end-to-end. These disks contain not only sites of strong physical attachment between adjacent cells but also *gap junctions*, which are tiny open channels, through which the intracellular fluid of adjacent cells comes into contact. When an action potential depolarizes the cell on one side of an intercalated disk, positive ions flow through the gap junctions and into the neighboring cell. This local, ionic current depolarizes the neighboring cell to threshold for the formation of an action potential. In effect, an action potential propagates from cell to cell through the gap junctions located within the intercalated disks. Skeletal muscle does not have intercalated disks or gap junctions between cells.

Cardiac Contractions Are Initiated by Action Potentials That Arise Spontaneously in Specialized Pacemaker Cells

Because cardiac muscle tissue forms a functional syncytium, and because a cardiac action potential leads to contraction, any one cardiac muscle cell can initiate a heartbeat. In other words, if a single cardiac muscle cell depolarizes to threshold and forms an action potential, that action potential will propagate from cell to cell, throughout the heart, and cause the whole heart to contract. Most cardiac muscle cells have the property of remaining stable at a resting membrane potential, so they never form action potentials by themselves. However, a few specialized cardiac muscle cells have the property of depolarizing spontaneously toward the threshold for the formation of action potentials. When any one of these specialized cells reaches threshold and forms an action potential, a heartbeat results. Cardiac cells that depolarize spontaneously toward threshold are called *pacemaker cells* because they initiate heartbeats and therefore determine the rate, or pace, of the heart.

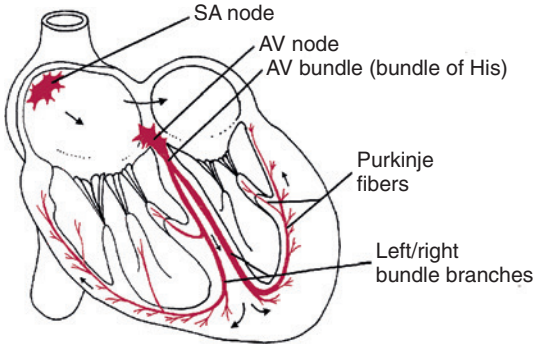
Although all spontaneously depolarizing cells in the heart are called pacemaker cells, only one pacemaker cell, the one that reaches threshold first, actually triggers a particular heartbeat. In the normal heart, the pacemaker cells that depolarize most quickly to threshold are located in the *sinoatrial* (SA) *node.* The SA node is in the right atrial wall, near where the *venae cavae* enter the right atrium.

Because it has spontaneously depolarizing pacemaker cells, the heart initiates its own muscle action potentials and contractions. Motor neurons are not necessary for initiating cardiac contractions, whereas they are needed for initiating skeletal muscle contractions. Motor neurons (sympathetic and parasympathetic) do affect the heart rate by influencing the rapidity with which the pacemaker cells depolarize to threshold, but the pacemaker cells initiate action potentials, and therefore heartbeats, even without any sympathetic or parasympathetic influences. Thus a denervated heart still beats, whereas a denervated skeletal muscle remains relaxed (in fact, paralyzed). The ability of the heart to beat without neural input enables surgically transplanted hearts to function. When a donor's heart is connected to a recipient's circulation during cardiac transplantation, no nerves are attached to the transplanted heart. The pacemaker cells in the transplanted heart initiate its action potentials and contractions. The only factor missing is control of the heart rate through cardiac sympathetic and parasympathetic nerves.

A System of Specialized Cardiac Muscle Cells Initiates and Organizes Each Heartbeat

Each normal heartbeat is initiated by an action potential that arises spontaneously in one of the pacemaker cells in the SA node (Fig. 19.2). When formed, the action potential propagates rapidly, from cell to cell, across the right and left atria, causing both atria to contract. Next, the action potential propagates *slowly,* from cell to cell, through a special pathway of cardiac muscle cells that lies between the atria and the ventricles. This pathway consists of the *atrioventricular* (AV) *node* and the first part of the *AV bundle,* also called the *bundle of His.* The AV node and AV bundle provide the only route for the propagation of action potentials from the atria to the ventricles. Elsewhere, a layer of connective tissue, which can neither form nor propagate action potentials, separates the atria and ventricles. In addition to providing the only conductive pathway between the atria and the ventricles, the AV node and the first part of the AV bundle have the special property of very slow conduction of action potentials. It takes 50 to 150 ms for an atrial action potential to travel through the AV node and the first part of the AV bundle; that is, it takes 50 to 150 ms for an atrial action potential to propagate into the ventricles. Slow conduction through the *AV junction* creates a delay between atrial and ventricular contractions.

After passing through the slowly conducting cells of the AV junction, the cardiac action potential enters a branching network of specialized cardiac cells that have the property of extremely rapid propagation of action potentials. The transition zone from slowly conducting to rapidly conducting cells is located within the AV bundle, which has slowly conducting cells in its first portion (adjacent to the AV node) and rapidly conducting cells beyond that. The rapidly conducting portion of the AV bundle splits to form the *left* and *right bundle branches.* At the ventricular apex, the bundle branches break up into a dispersed network of *Purkinje fibers,* which carry the action potential rapidly along the inner walls (subendocardial layers) of both ventricles. From the Purkinje fibers, action potentials then propagate outward, from cell to cell, through the ventricular walls. As the action potential reaches each ventricular muscle fiber, that fiber contracts. The very rapid



• **Fig. 19.2** Specialized conduction system of the heart is responsible for the initiation and organization of cardiac contractions. The system is composed of specialized cardiac muscle fibers, not nerves. *AV,* Atrioventricular; *SA,* sinoatrial.

Vet-ebooks.com

conduction of the cardiac action potential, from cell to cell, through the latter portion of the AV bundle, the bundle branches, and the Purkinje system results in a nearly synchronous contraction of all the fibers in both ventricles.

The SA and AV nodes, AV bundle, bundle branches, and Purkinje fibers are collectively called the *specialized conduction system of the heart.* This system is composed of specialized cardiac muscle cells, not nerves. The particular characteristics of the components in the specialized conduction system cause each heartbeat to follow a specific, patterned sequence. In a normal beat, both atria contract, almost simultaneously. Next, there is a brief pause (caused by slow propagation of the action potential through the AV node). The two ventricles then contract, almost simultaneously. Finally, the entire heart relaxes and refills.

Fig. 19.3 reemphasizes the role of the specialized conduction system in initiating and organizing a normal cardiac contraction. In this "time lapse" illustration, atrial excitation begins at time $t = 0$, when SA node cells have reached threshold and an action potential is just beginning to propagate out of the SA node and into regular atrial tissue. Within 0.1 second, the action potential has propagated completely across the right and left atria, and a coordinated contraction of both atria is just beginning. As the action potential propagates across the atria, it also depolarizes the first cells in the AV node, beginning at time $t = 0.04$ second. While the atria are in a depolarized (excited) state, the action potential is propagating slowly from cell to cell through the AV node and first part of the AV bundle. After traversing this slowly conducting region, the action potential propagates rapidly through the remainder of the bundle of His and its branches. As illustrated in Fig. 19.3, the action potential arrives at the ventricular apex at time $t = 0.17$ second. Note that it takes about 0.13 second [(0.17 − 0.04) second] for the action potential to travel through the AV node and bundles; that is, 0.13 second represents a typical delay between atrial depolarization and ventricular depolarization (*A-V delay*). From the ventricular apex, the Purkinje fibers propagate the action potential rapidly throughout both ventricles. Ventricular excitation (depolarization) is complete by time $t = 0.22$ second, and both ventricles contract. By this time the atrial cells have ended their action potential; they have repolarized to a resting state and are relaxing. After ventricular excitation and contraction, the ventricular cells also repolarize to their resting state and the ventricles relax. The whole heart remains in a resting, relaxed state until the spontaneously depolarizing SA node pacemaker cells originate the next beat.

Cardiac Action Potentials Are Extremely Long

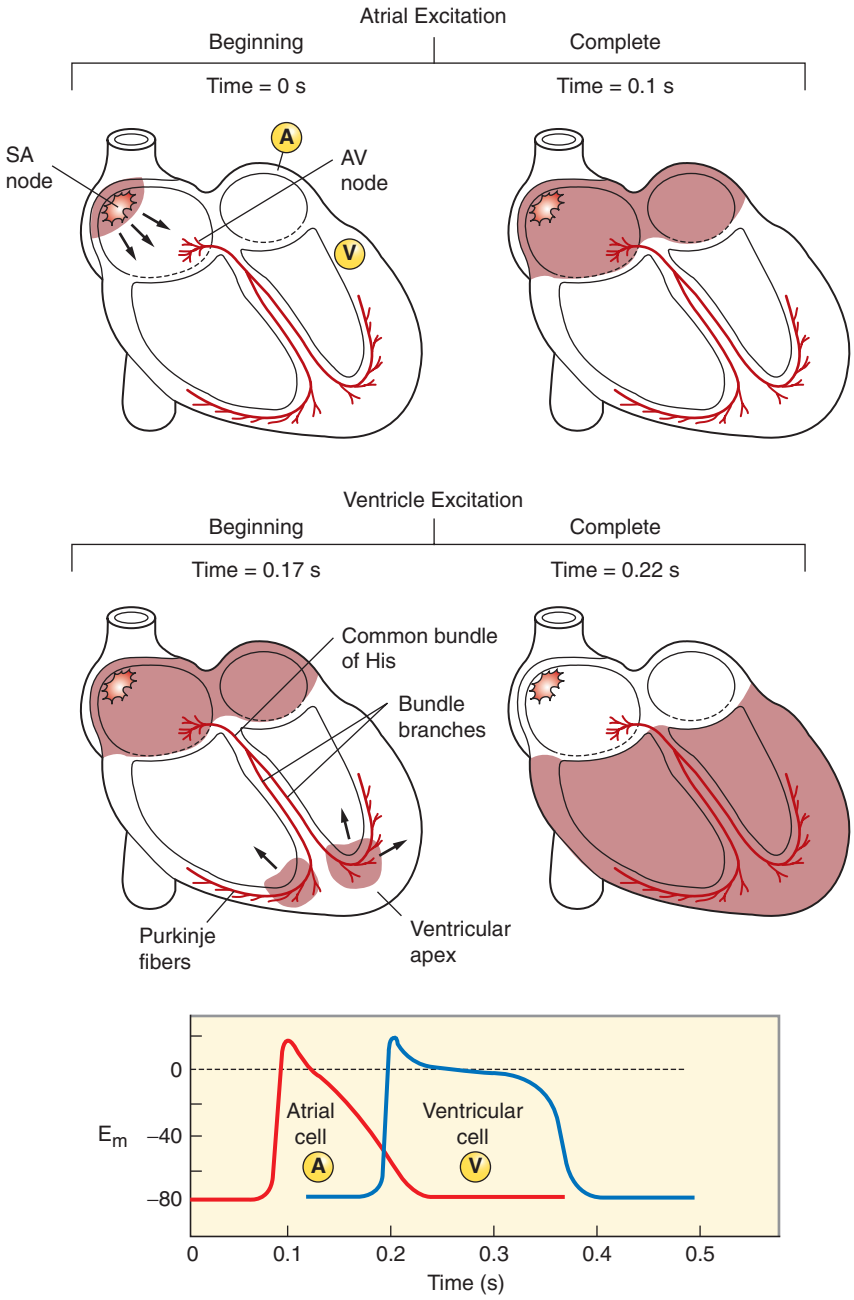
Two major differences between action potentials in skeletal muscle and cardiac muscle have already been mentioned: first, action potentials propagate from cell to cell in cardiac muscle, whereas skeletal muscle cells are electrically isolated from one another. Second, the heart has pacemaker cells that spontaneously form action potentials, whereas skeletal muscle does not. A skeletal muscle cell only depolarizes and forms an action potential when "commanded" to do so by its motor neuron.

A third important difference between skeletal and cardiac action potentials is their duration (Fig. 19.4). The entire action potential in a skeletal muscle lasts only 1 to 2 ms. A cardiac action potential lasts about 100 times longer (100–250 ms). Prolongation of the cardiac action potential is brought about by prolonged changes in the permeability of the cardiac muscle membrane to potassium and calcium ions (K^+ and Ca^{2+}). Cardiac muscle cell membranes have fast, voltage-gated Na^+ channels, K^+ "leak" channels, and voltage-gated K^+ channels similar to those found in skeletal muscle, but in addition, cardiac cell membranes also have other types of K^+ channels, as well as special Ca^{2+} channels that are not present in skeletal muscle. These Ca^{2+} channels allow extracellular Ca^{2+} to enter the cytoplasm of cardiac muscle cells, and this Ca^{2+} entry prolongs the cardiac action potential and controls the strength of the resulting cardiac contraction. The presence of Ca^{2+} channels (and their influence on action potential duration and contractile force) is the fourth major difference between cardiac and skeletal muscle.

Before delving more deeply into the special significance of the membrane Ca^{2+} channels in cardiac muscle, it is useful to review the roles of the Na^+, K^+ exchange pumps and of the Na^+ and K^+ channels in skeletal muscle, because these pumps and channels function similarly in cardiac and skeletal muscle. As explained in Chapter 4, the Na^+, K^+ exchange pumps keep K^+ concentration higher inside cells (than in extracellular fluid) and keep Na^+ concentration lower inside cells (than in extracellular fluid). With each pump cycle, a Na^+, K^+ exchanger transfers three Na^+ out of a cell and two K^+ into a cell. This imbalance creates a small, negative resting membrane potential (inside of the cell membrane is negative in comparison with the outside). In resting skeletal and cardiac cells, there are many open K^+ channels (leakage channels) but very few open Na^+ channels. As a result, the resting cell is much more permeable to K^+ than to Na^+, so there is a greater tendency for positive K^+ to exit from the cell than for positive Na^+ to enter. This charge imbalance is the main contributor to the normal, negative resting membrane potential. The resting membrane potential in skeletal muscle cells is typically between –70 and –80 mV (see Fig. 19.4). An action potential is created when something *depolarizes* the cell (makes it less negative inside). Specifically, depolarization to the *threshold voltage* for opening the voltage-gated Na^+ channels allows an influx of extracellular Na^+ into the cell. This rapid entry of positive ions causes the cell membrane to become positively charged on its inside surface. This positive membrane potential persists for only a moment, however, because the Na^+ voltage-gated channels become *inactivated* very quickly. Na^+ entry ceases, and the cell rapidly repolarizes toward its resting membrane potential. Repolarization is also promoted by the opening of longer latency, voltage-gated K^+ channels. In fact, this opening of extra K^+ channels may cause neurons and skeletal muscle cells to become *hyperpolarized* (even more negative than normal resting membrane potential) for a few milliseconds at the end of each action potential (see Fig. 19.4, panel C).

In a resting skeletal muscle cell, calcium ions are sequestered within the *sarcoplasmic reticulum.* The occurrence of an action potential in the skeletal muscle cell causes Ca^{2+} to be released from the sarcoplasmic reticulum into the free intracellular fluid, which is called the *cytosol.* The increase in cytosolic Ca^{2+} concentration initiates muscle contraction (see Fig. 1.5). The skeletal muscle contraction initiated by a single action potential is very brief, because the cytosolic Ca^{2+} is rapidly pumped back into the sarcoplasmic reticulum by active transport, which causes the muscle cell to relax. Note that the Ca^{2+} responsible for initiating skeletal muscle contraction comes entirely from an intracellular storage site, the sarcoplasmic reticulum. No extracellular Ca^{2+} enters the cell during the action potential, because skeletal muscle cells do not have membrane Ca^{2+} channels. By contrast, in cardiac muscle, membrane Ca^{2+} channels and the entry of extracellular Ca^{2+} into the cells play key roles in both action potentials and contractions.

Vet-ebooks.com

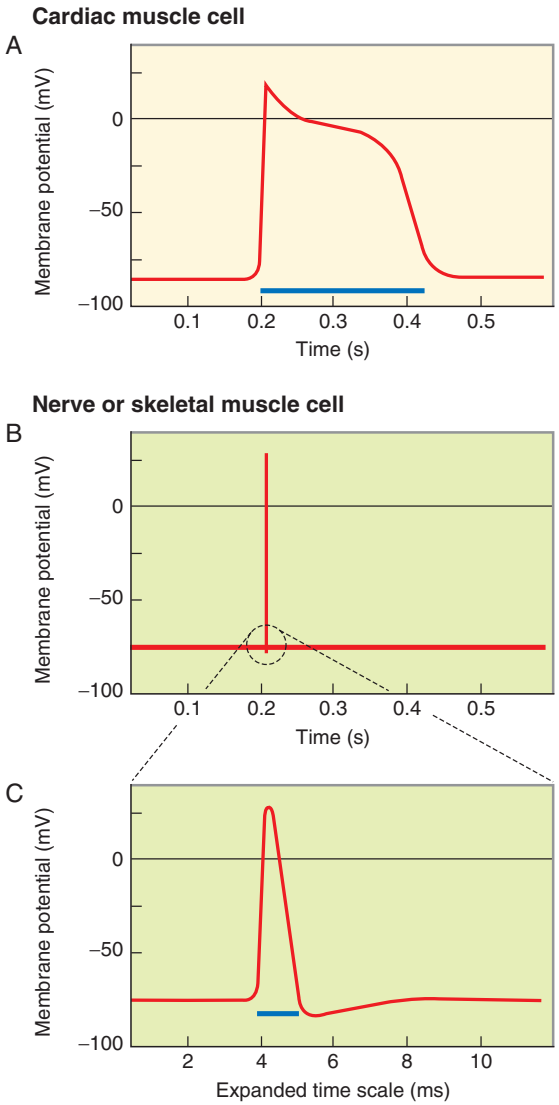


• **Fig. 19.3 Top.** Heart is pictured at four instants during initiation of a normal contraction. *Shading* indicates areas of heart where an action potential is underway. *Time = 0 sec,* Pacemaker cells in the sinoatrial (*SA*) node have just reached threshold, and an action potential has begun to propagate outward across the atria. *Time = 0.1 sec,* Action potential has reached all parts of both atria (action potential underway in all atrial cells). *Time = 0.17 sec,* Action potential has passed through the atrioventricular (*AV*) node and down the bundle branches and has just reached the ventricular apex. *Time = 0.22 sec,* Action potential has just finished propagating outward through the walls of both ventricles (action potential is underway in all ventricular cells). All atrial cells have finished their action potential. **Bottom.** Graph shows the timing of action potentials in a left atrial cell and a left ventricular cell (at locations labeled *A* and *V* in the top drawing). The locations of cells A and V make these among the last atrial and ventricular cells to be depolarized as an action potential propagates across the atria and ventricles, respectively. E_m, Membrane potential in millivolts.

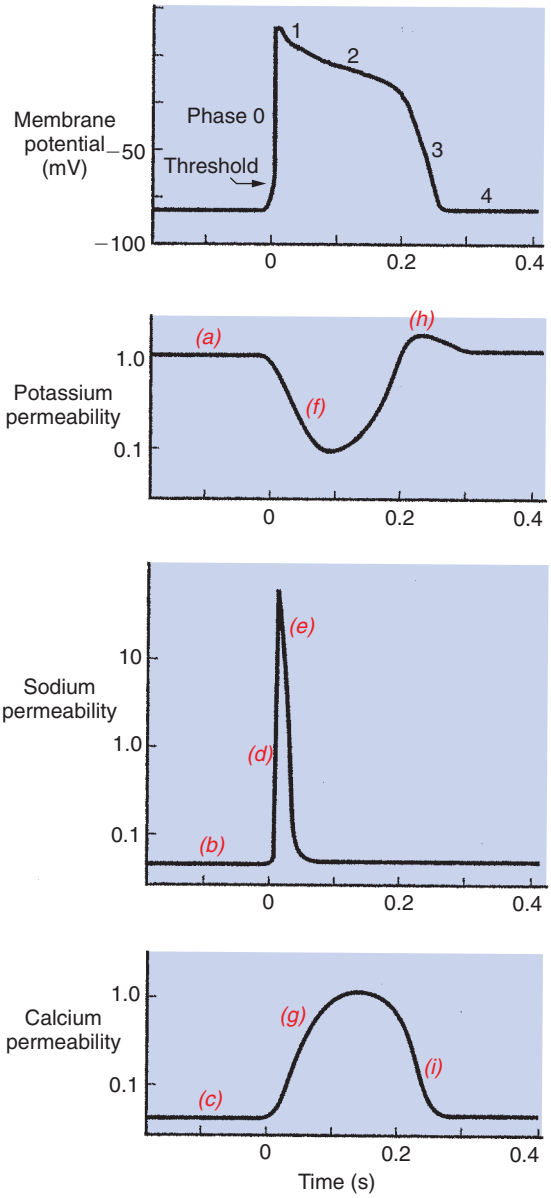
Membrane Calcium Channels Play a Special Role in Cardiac Muscle

Fig. 19.5 depicts a cardiac muscle cell action potential, aligned in time with graphs of the sequence of changes in K^+, Na^+, and Ca^{2+} permeability that are responsible for the action potential. As the time line begins (on the left side of each graph), the cardiac cell is at a normal, negative resting membrane potential of about –80 mV. The cardiac membrane potential is negative at rest for the same reason that skeletal muscle cells have negative resting membrane

Vet-ebooks.com



• **Fig. 19.4** Action potential in a cardiac muscle cell (A) lasts 100 times longer than an action potential in a nerve or skeletal muscle cell (B). The prolonged phase of depolarization in cardiac muscle cells is called the plateau of the action potential. **C**, The nerve or skeletal muscle action potential is shown on a greatly expanded time scale to illustrate that an action potential in a nerve or skeletal muscle cell has a different shape than a cardiac action potential, as well as a much shorter duration. The *dark bars* under each action potential indicate the length of the absolute refractory period.



• **Fig. 19.5** Top graph: Changes in membrane potential in a cardiac muscle cell during an action potential. The cell is initially at resting membrane potential. A depolarization to *Threshold* (at time = 0) triggers the action potential. *Phase 0* of the action potential is called rapid depolarization. *Phase 1* is partial repolarization. *Phase 2* is the plateau. *Phase 3* is repolarization. *Phase 4* is return to resting membrane potential. This sequence of changes in the cell's membrane potential is caused by a patterned sequence of changes in the relative permeabilities of the cell membrane to K^+, Na^+, and Ca^{2+}, as shown in the three lower graphs. Key features of these permeability changes (labeled *a* to *i*) are discussed in the text.

potentials: many K^+ channels are open at rest, and most of the Na^+ channels are closed. As a result, membrane permeability to K^+ is much higher than membrane permeability to Na^+ (see Fig. 19.5, *labels a and b*). In resting cardiac cells, the membrane Ca^{2+} channels are closed, so Ca^{2+} permeability is very low (*label c*); extracellular Ca^{2+} is prevented from entering the cardiac cells.

As in skeletal muscle, a cardiac action potential is created when the cell is depolarized to the threshold voltage for opening the voltage-gated Na^+ channels. The resulting sudden increase in Na^+ permeability (*label d*) allows a rapid influx of extracellular Na^+ into the cell, which causes the rapid depolarization phase (*Phase 0*) of the cardiac action potential. Within a few milliseconds, the Na^+ channels inactivate, which causes the Na^+ permeability to decrease quickly (*label e*). The resulting decrease in Na^+ influx into the cell causes the membrane to begin to repolarize (*Phase 1*). However, this repolarization is interrupted, and is followed by a prolonged plateau of depolarization, which lasts about 200 ms (*Phase 2*). The plateau of the cardiac action potential is brought about by two conditions that do not occur in nerves or skeletal muscle fibers: (1) some

Vet-ebooks.com

types of gated K^+ channels close, so K^+ permeability decreases (*label f*); and (2) many gated Ca^{2+} channels open, so Ca^{2+} permeability increases (*label g*). Because the Ca^{2+} concentration is higher in the extracellular fluid than in the intracellular fluid, Ca^{2+} flows through the open Ca^{2+} channels and into the cytosol. The combination of reducing the exit of K^+ from the cell and allowing the entrance of Ca^{2+} into the cell keeps the cell membrane in a depolarized state. After about 200 ms, the K^+ channels reopen, and the Ca^{2+} channels close; K^+ permeability increases, and Ca^{2+} permeability decreases (*labels h and i*). The combination of increasing the exit of K^+ from the cell and shutting off the entrance of Ca^{2+} into the cell causes the cell to repolarize (*Phase 3*) and eventually to return to its stable, negative resting membrane potential (*Phase 4*).

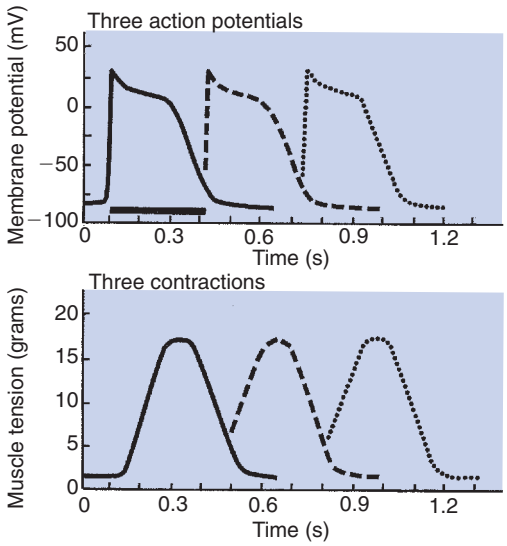
The specialized Ca^{2+} channels in cardiac muscle cell membranes are called *slow Ca^{2+} channels* (or *L-type Ca^{2+} channels*) because they take much longer to open than the Na^+ channels, and they stay open much longer. As shown in Fig. 19.5, Na^+ permeability increases and then decreases (Na^+ channels open and then inactivate) within a few milliseconds. Ca^{2+} permeability, in comparison, is slow to increase (Ca^{2+} channels are slow to open) and Ca^{2+} permeability remains elevated for about 200 ms (the time Ca^{2+} channels stay open). In recognition of their much quicker responses, the Na^+ channels in cardiac muscle are sometimes called *fast Na^+ channels.*

The Ca^{2+} that enters a cardiac cell during an action potential triggers the release of additional Ca^{2+} from the sarcoplasmic reticulum. This process is called *calcium-triggered calcium release* (or *calcium-induced calcium release*). In less than 0.1 second, the concentration of free Ca^{2+} in the cytosol increases about 1000-fold. As in skeletal muscle, this increase in cytosolic Ca^{2+} initiates contraction. When the Ca^{2+} channels close, at the conclusion of the action potential, most of the cytosolic Ca^{2+} is pumped back into the sarcoplasmic reticulum or pumped back across the cell membrane into the extracellular fluid. Both these processes involve active transport because the Ca^{2+} is being pumped against its electrochemical gradient. Once the cytosolic Ca^{2+} concentration is returned to its low, resting level, the cardiac muscle relaxes.

The Long Duration of the Cardiac Action Potential Guarantees a Period of Relaxation (and Refilling) Between Heartbeats

Na^+ channels become inactivated at the peak of the cardiac action potential. Na^+ cannot pass through an inactivated channel; therefore, as long as the Na^+ channels remain inactivated, another action potential cannot occur. The inactivated state ends, and Na^+ channels become susceptible to reopening, only when the cell membrane potential returns to (or nearly to) its resting level. Thus Na^+ inactivation guarantees that the upstroke of a second action potential cannot occur until the first action potential is completed (or very nearly completed).

While the Na^+ channels are inactivated, the cell is said to be *refractory* (resistant) with regard to the formation of an action potential. The time after the beginning of one action potential during which another action potential cannot be initiated is called the *refractory period* (sometimes designated as *absolute refractory period*). Because Na^+ inactivation lasts until the membrane potential returns to (or nearly to) its resting level, the refractory period lasts about as long as an action potential. Thus the refractory period in a cardiac muscle cell lasts about 200 ms, whereas the refractory period in a skeletal muscle cell lasts only about 2 ms (see Fig. 19.4).



• **Fig. 19.6** Relationship between action potentials (top graph) and contractions (bottom graph) in a cardiac muscle cell. The first of three cardiac action potentials (*solid line, top*) causes a cardiac contraction (*solid line, bottom*). Note that the action potential and contraction have similar durations. The *heavy horizontal bar* under the first action potential shows the duration of the absolute refractory period. The *dashed line* and *dotted line* in the top graph show the earliest possible occurrence of a second and a third action potential, each occurring right after the absolute refractory period for the preceding action potential. The *dashed line* and *dotted line* in the bottom graph depict the corresponding cardiac contractions. Because of the long refractory period, each contraction is almost over before the earliest possible next contraction can begin. This guarantees a period of cardiac relaxation between contractions.

The long refractory period in cardiac muscle guarantees a period of relaxation (and cardiac refilling) between cardiac contractions. Fig. 19.6 (top) depicts the quickest possible succession of three action potentials in a cardiac muscle cell; the second action potential begins immediately after the conclusion of the refractory period for the first action potential. Likewise, the third action potential begins immediately after the conclusion of the refractory period for the second. The bottom graph in Fig. 19.6 shows the pattern of muscle contraction that results from these three action potentials. Note that contractile strength reaches a peak late in the plateau phase of each action potential, and that the contractile strength decreases (the muscle begins to relax) during the repolarization phase of each action potential. As a result, the cardiac muscle cell becomes partially relaxed before the earliest possible subsequent contraction can begin; that is, each cardiac action potential produces a contraction that is distinctly separated from the preceding contraction. Because of its long refractory period, cardiac muscle cannot sustain a continuous (*tetanic*) contraction. Thus the heart has a guaranteed period of relaxation (and refilling) between heartbeats.

The pattern of changes in muscle tension depicted in the bottom graph of Fig. 19.6 corresponds closely to the changes in the cytosolic Ca^{2+} concentration. This makes sense, considering that the increase in cytosolic Ca^{2+} concentration initiates muscle contraction, and the subsequent removal of Ca^{2+} from the cytosol permits the muscle to relax. Cytosolic Ca^{2+} concentration increases during the plateau of the action potential (because of Ca^{2+}-triggered Ca^{2+} release) and decreases back to its resting level during the repolarization phase

Vet-ebooks.com

of the action potential (as active transport pumps move Ca^{2+} back into the sarcoplasmic reticulum or out into the extracellular fluid).

In skeletal muscle cells, an action potential lasts only 1 to 2 ms. The membrane is repolarized (and the refractory period is over) even before the release of Ca^{2+} from the sarcoplasmic reticulum is finished, and many milliseconds before the released Ca^{2+} is pumped back into the sarcoplasmic reticulum. As a result, the cytosolic Ca^{2+} concentration in skeletal muscle reaches its peak level after the action potential is over, and the contractile tension resulting from the action potential also reaches its peak after the action potential is over. Because a contractile twitch lasts much longer than the refractory period in skeletal muscle, several action potentials can occur during the time of a single contractile twitch. Multiple action potentials in quick succession cause cytosolic Ca^{2+} concentration to build to a high level and stay there. The resulting contractile tension is stronger than the tension that results from a single action potential, and it is sustained for a longer time. In effect, the muscle twitches caused by successive action potentials "fuse" together. This phenomenon is called *temporal summation.* Fusion and temporal summation are the mechanisms that permit graded and prolonged tension development in skeletal muscle. In contrast, the long refractory period in cardiac muscle cells prevents the fusion and summation of cardiac contractions. Each contraction of the heart (each heartbeat) is followed immediately by a relaxation. Temporal summation of muscle contraction does not occur in the heart.

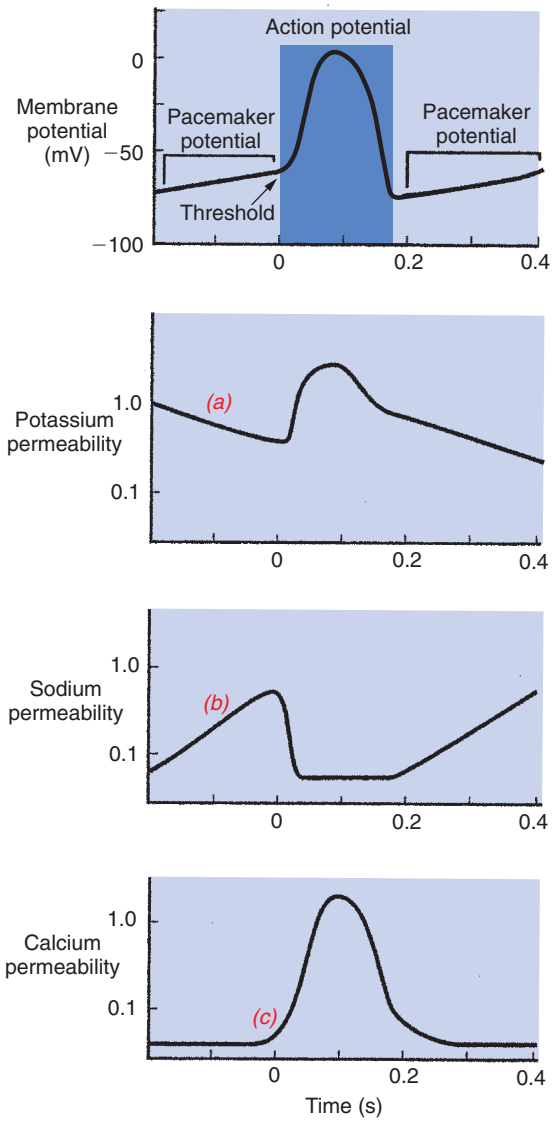
Atrial Cells Have Shorter Action Potentials Than Ventricular Cells

The previous description of cardiac ion channels, action potentials, and contractions is based on properties of normal ventricular cells. Atrial cells are basically similar, except that their action potentials are shorter than action potentials in ventricular cells. Like ventricular cells, atrial cells have fast Na^+ channels that open briefly at the beginning of an action potential and then become inactivated. Likewise, atrial slow Ca^{2+} channels open during the action potential, and K^+ channels close. The differences between atrial and ventricular cells are that atrial slow Ca^{2+} channels stay open for a shorter time than those in ventricular cells, and atrial K^+ channels stay closed for a shorter time. As a result, the plateau of an atrial cell's action potential is shorter in duration and not as "flat" as the plateau of a ventricular cell's action potential (see Fig. 19.3, bottom). Atrial cells also have shorter refractory periods than ventricular cells. Therefore the atrial cells are capable of forming more action potentials per minute than ventricular cells; that is, the atria can "beat" faster than the ventricles. The implications of this difference are discussed later in this chapter.

Specialized Ion Channels Cause Cardiac Pacemaker Cells to Depolarize to Threshold and Form Action Potentials

As mentioned earlier, the cardiac pacemaker cells of the SA node spontaneously depolarize to threshold and form action potentials. The spontaneous depolarization is called a *pacemaker potential,* and it is the key distinguishing feature of a pacemaker cell (Fig. 19.7, top). Also, the action potentials of cardiac pacemaker cells typically have a rounded appearance; they lack the very rapid (phase 0) depolarization seen in ventricular and atrial cells.

The spontaneous depolarizations and rounded action potentials of pacemaker cells are consequences of the particular ion channels found in these cells. Pacemaker cells have an unusual set of K^+



• **Fig. 19.7** Top graph: Typical pattern of changes in membrane potential in cardiac pacemaker cell. A pacemaker cell does not have a stable resting membrane potential. Instead, it spontaneously depolarizes to *Threshold*, and this initiates an *Action Potential*. The spontaneous depolarization (called *Pacemaker Potential*) and the action potential are caused by a patterned sequence of changes in the permeabilities of the cell membrane to K^+, Na^+, and Ca^{2+}, as shown in the three lower graphs. Key features of the permeability changes that cause the pacemaker potential are labeled *a, b,* and *c,* and are discussed in the text. The action potential is driven primarily by a large, prolonged increase in Ca^{2+} permeability. The absence of fast Na^+ channels in pacemaker cells accounts for the much slower upstroke of the action potential in pacemaker cells, as compared with nonpacemaker cells. (Compare with Fig. 19.5.)

channels. At the conclusion of an action potential in a pacemaker cell, K^+ permeability is quite high, because many K^+ channels are open. Then some K^+ channels begin to close. As K^+ permeability decreases (see Fig. 19.7, *label a*), less and less K^+ leaves the cells, which makes the cells progressively less and less negatively charged inside. Pacemaker cells lack the usual voltage-gated fast Na^+ channels. Instead, these cells have *pacemaker Na^+ channels* (also called *funny Na^+ channels*), which close during an action potential and then begin to open again, spontaneously, once an action potential has finished. The spontaneous opening of the pacemaker Na^+ channels causes a progressive increase in the cell's Na^+ permeability (*label*

Vet-ebooks.com

b in Fig. 19.7). The increase in Na^+ permeability allows Na^+ to enter the cell from the extracellular fluid, which contributes to the depolarization of the cell toward threshold. Ca^{2+} channels also make a small contribution to the pacemaker potential. Late in the pacemaker potential, just before a pacemaker cell reaches threshold, slow Ca^{2+} channels begin to open, and Ca^{2+} permeability begins to increase (*label c* in Fig. 19.7). The resulting entry of Ca^{2+} into the cell speeds its final approach to threshold. Thus the pacemaker potential is caused by the closing of K^+ channels, the opening of pacemaker Na^+ channels, and (late in the process of reaching threshold) the opening of Ca^{2+} channels. These spontaneous changes in Na^+, K^+, and Ca^{2+} channels in pacemaker cells are in contrast to the stable status of the ion channels in normal, resting atrial or ventricular cells.

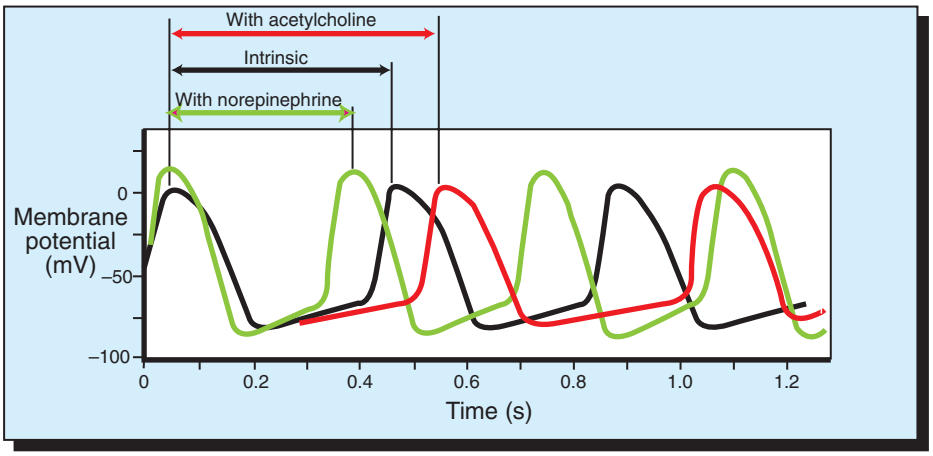
When threshold is reached in a pacemaker cell, an action potential occurs. The upstroke of the action potential is quite slow compared with the rapid (phase 0) depolarization in a normal atrial or ventricular cell, because there are no fast, voltage-gated Na^+ channels in pacemaker cells and therefore no sudden influx of Na^+ upon reaching threshold. The ion primarily responsible for the action potential in a pacemaker cell is Ca^{2+}. When threshold is reached, many of the cell's slow Ca^{2+} channels open. The permeability to Ca^{2+} increases, and extracellular Ca^{2+} flows into the cell. The action potentials in pacemaker cells are often called slow action potentials, because they lack a rapid, phase 0 depolarization and because they are caused primarily by the opening of slow Ca^{2+} channels. In contrast, normal atrial or ventricular action potentials are called fast action potentials. Note, however, that all cardiac action potentials (whether "slow" or "fast") have a very long duration compared with action potentials in nerve or skeletal muscle cells.

Sympathetic and Parasympathetic Nerves Act on Cardiac Pacemaker Cells to Increase or Decrease the Heart Rate

Fig. 19.8 shows how the neurotransmitters *norepinephrine* and *acetylcholine* affect the pacemaker cells of the heart. Norepinephrine exerts its effect by activating *β-adrenergic receptors* on the cell membranes of pacemaker cells. Activation of such receptors speeds up the ion channel changes that are responsible for the spontaneous depolarization of pacemaker cells. Consequently, the pacemaker cells reach threshold more quickly in the presence of norepinephrine; there is a shorter interval between heartbeats. Therefore heart rate is elevated above its intrinsic or spontaneous level.

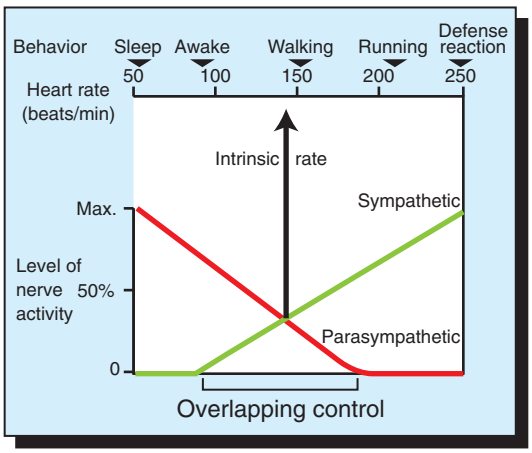
Acetylcholine has the opposite effect. Acetylcholine activates *muscarinic cholinergic receptors* on the cell membranes of pacemaker cells, which slows the ion channel changes that are responsible for the pacemaker cell's spontaneous depolarization. Because the pacemaker cells take longer to reach threshold in the presence of acetylcholine, there is a longer interval between heartbeats. Therefore heart rate is decreased below its intrinsic or spontaneous level.

Sympathetic neurons release norepinephrine at the SA node cells, and thus sympathetic nerve activity increases the heart rate. The same effect is brought about by the hormone mixture (epinephrine plus norepinephrine), which is secreted by the adrenal medulla glands in response to physical or emotional stress. Parasympathetic neurons release acetylcholine at the SA node cells, and thus parasympathetic activity decreases the heart rate. Fig. 19.9 illustrates how sympathetic and parasympathetic neurons interact in the control of the heart rate. In the absence of norepinephrine and acetylcholine, the heart beats at its intrinsic rate. For a large dog, this rate is typically about 140 beats per minute (beats/min). Heart rates below the intrinsic rate are achieved by activation of parasympathetic neurons and release of acetylcholine. Accordingly, the graph in Fig. 19.9 indicates that parasympathetic activity is high during awake rest (heart rate of 90 beats/min) and very high during sleep (heart rate of 55 beats/min). Heart rates above the intrinsic rate (e.g., during exercise or emotional stress) are achieved by activation of the sympathetic nerves to the heart and release of norepinephrine (or by epinephrine and norepinephrine released into the circulation from the adrenal glands). The highest possible level of sympathetic activity, and therefore the highest possible heart rate, occurs during maximal exercise or a *defense alarm reaction* (*"fear, fight, or flight" response*).



• **Fig. 19.8** In the absence of neural or hormonal influences, a pacemaker cell of the sinoatrial (SA) node spontaneously depolarizes to threshold and initiates a series of action potentials, three of which are shown by the *black line.* The interval between action potentials under these conditions determines the intrinsic, or spontaneous, heart rate (in this case, 0.43 s between action potentials, which corresponds to a heart rate of 140 beats/min). Norepinephrine increases the rate of depolarization (*green line*) and therefore shortens the interval between action potentials (i.e., increases heart rate). Acetylcholine decreases the rate of depolarization (*red line*) and therefore lengthens the interval between action potentials (i.e., decreases heart rate). *SA,* Sinoatrial.

Vet-ebooks.com

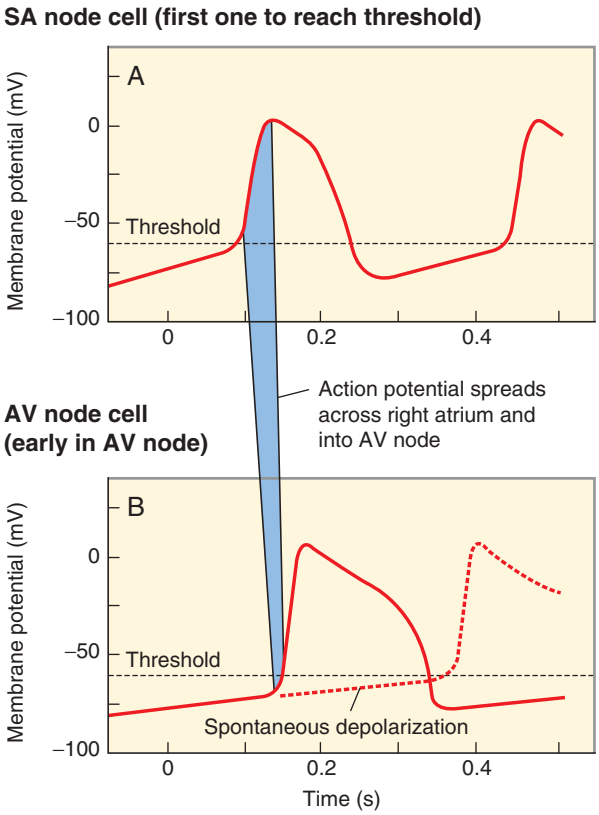


• **Fig. 19.9** The upper scale shows that the heart rate of a normal, large dog ranges from 50 to 250 beats/min, depending on behavioral state. The graph illustrates that this wide range of heart rates is brought about by the interactions between sympathetic nerve activity, which speeds the heart above its intrinsic rate, and parasympathetic nerve activity, which slows the heart below its intrinsic rate. Sympathetic and parasympathetic nerves are simultaneously active over a considerable portion of the heart rate range (*overlapping control*). Note that the heart beats at its intrinsic rate (about 140 beats/min) either in the absence of any neural influence or when sympathetic and parasympathetic effects are equal and opposite.

By changing the levels of sympathetic and parasympathetic activity (*tone),* the dog adjusts its heart rate over a wide range, as appropriate for each behavioral situation. When both sympathetic and parasympathetic nerves are active, the resulting heart rate represents the outcome of a "tug-of-war" between sympathetic action to increase the heart rate and parasympathetic action to decrease the heart rate. Typically, the sympathetic and parasympathetic systems are both partially active during awake states, ranging from quiet rest (heart rate about 90 beats/min) to moderate exercise (heart rate about 175 beats/min). Parasympathetic activity predominates in the lower part of this range, and sympathetic activity predominates in the higher part. When sympathetic and parasympathetic activity are equal, their effects cancel, and the heart will beat at its intrinsic (spontaneous) rate. Simultaneous, partial activation of sympathetic and parasympathetic neurons appears to give the nervous system tight control over the heart rate under a wide variety of behavioral conditions.

Cells of the Atrioventricular Node Act as Auxiliary Pacemakers and Stabilize Ventricular Rate

As with SA node cells, the cells of the AV node normally exhibit pacemaker activity and slow action potentials. However, as shown in Fig. 19.10, the AV node cells spontaneously depolarize toward threshold much more slowly than SA node cells. Therefore under normal circumstances, the SA node cells reach threshold first and initiate an action potential, which then propagates from cell to cell across the atria and into the AV node. Within the AV node, this action potential encounters cells that are already slowly, spontaneously depolarizing toward threshold. The arriving action potential quickly depolarizes these AV node pacemaker cells to threshold, and they form an action potential, which then propagates into the AV bundle, bundle branches, and Purkinje system. Thus under normal conditions, the pacemaker activity of the AV node



• **Fig. 19.10** Both *SA*, sinoatrial, and *AV*, atrioventricular, node cells exhibit pacemaker activity (spontaneous depolarization toward threshold). Normally the SA node cells (panel A) depolarize more quickly and reach the threshold first. The resulting atrial action potential propagates into the AV node (as represented by *blue band*) and depolarizes the AV node cells quickly to their threshold, causing them to form an action potential (as shown by solid line in panel *B*). However, if the SA node pacemaker cells become nonfunctional or if atrial action potentials are not conducted into the AV node, the AV node cells eventually depolarize to threshold and initiate action potentials on their own *(dashed line* in panel B). In this way the AV node cells serve as an auxiliary (i.e., emergency) ventricular pacemaker.

cells is preempted, beat after beat, by the more frequent action potentials initiated by the SA node pacemaker cells.

Under certain abnormal conditions, AV node pacemaker function becomes critical for survival. For example, if the SA node pacemaker cells fail to depolarize to threshold (due to injury or disease), or if atrial action potentials do not successfully propagate into the AV node (complete AV node block), AV node pacemaker cells will spontaneously initiate action potentials that propagate into the ventricles, causing them to contract. This *auxiliary pacemaker function of AV node* cells is the only thing that keeps the ventricles beating in a circumstance where atrial action potentials fail to form or fail to enter the AV node.

Because the AV node pacemaker cells depolarize more slowly than do normal SA node cells, the ventricular rate resulting from AV node pacemaker activity is very low, about 30 to 40 beats/min in a resting dog, compared with 80 to 90 beats/min when the SA node cells are the pacemakers. Also, action potentials initiated by the AV node pacemakers usually do not propagate "backward" into the atria; therefore atrial contractions may be absent. Nevertheless, a ventricular contraction rate of 30 to 40 beats/min is usually enough to sustain life temporarily, until the cardiac electrical

Vet-ebooks.com

problem can be addressed medically. Thus AV node cells are sometimes called the heart's *emergency pacemakers.*

Another important feature of the AV node cells is that they have longer refractory periods than normal atrial cells. The longer refractory period of AV node cells helps protect the ventricles from being stimulated to contract at rates that are too rapid to allow time for adequate ventricular refilling between beats. This *protective function of the AV node* is critical to an animal's survival when atrial dysfunctions cause atrial action potentials to arrive at the AV node very frequently and/or irregularly (see later discussion of atrial flutter/fibrillation).

The long refractory period of the AV node cells plays an important role, even in a normal heart. When a normal action potential reaches the ventricles, it is prevented from "circling back" (and reactivating the atria) by the prolonged refractory state of the AV node cells.

Table 19.2 summarizes the four important electrical characteristics of the AV node previously discussed. Note that three of these characteristics are influenced by the nervous system. As indicated in the table, sympathetic activity increases the conduction velocity of the AV node cells, shortens their refractory period, and speeds their auxiliary pacemaker activity. Parasympathetic activation has the opposite effects. These sympathetic and parasympathetic effects are appropriate for different heart rates. For example, during exercise, when sympathetic activity is high and the SA node pacemakers are initiating heartbeats frequently, the whole process of cardiac contraction and relaxation must occur more quickly. Thus it is appropriate that sympathetic action not only increases heart rate, but also increases the velocity of action potential conduction through the AV node, which shortens AV delay. In addition, sympathetic activation shortens the AV node refractory period, which allows each of the frequent atrial action potentials to be conducted to the ventricles. Finally, sympathetic activation enhances AV node auxiliary pacemaker activity, which provides the animal with a high enough ventricular rate to cope with some stress, even if the SA node pacemaker has failed or if the AV node is blocked. Conversely, when parasympathetic activation causes the SA node pacemakers to decrease the heart rate, all aspects of cardiac contraction and relaxation can proceed at a more leisurely pace. Under these conditions, it is appropriate for AV node conduction velocity to be slowed and the AV node refractory period to be lengthened.

Sympathetic Nerves Act on All Cardiac Muscle Cells to Cause Quicker, More Forceful Contractions

Sympathetic neurons release norepinephrine in all regions of the heart, not only at the SA and AV nodes, and all cardiac muscle cells have β-adrenergic receptors that are activated by norepinephrine. Circulating epinephrine or norepinephrine (whether released from the adrenal medulla or administered as a drug) can also activate these same receptors. The effects of β-receptor activation on the SA and AV node cells have already been described (see Fig. 19.8 and Table 19.2). In all other atrial and ventricular cells, β-receptor activation leads to taller, shorter duration action potentials and to stronger, quicker contractions. One reason for these effects is that activation of β receptors increases the number of L-type Ca^{2+} channels that open during the plateau (phase 2) of an action potential, which increases the amount of extracellular Ca^{2+} that enters the cell. Because Ca^{2+} entry is the primary depolarizing influence during the plateau, increased Ca^{2+} entry raises the plateau (makes the membrane potential more positive). The action potential also becomes shorter because of a complicated effect of the elevated plateau on K^+ channels. Recall that some K^+ channels close at the beginning of a cardiac action potential and then, after a time, reopen (see Fig. 19.5). Reopening of the K^+ channels helps repolarize the cell to a resting state at the end of the action potential. When the membrane potential is more positive than normal during the plateau, the K^+ channels reopen sooner. This shortens the action potential and speeds repolarization. Overall, β-receptor activation makes each action potential taller (more positive plateau) and shorter in duration. An action potential of higher amplitude propagates more quickly along each cell and from cell to cell, leading to faster conduction velocity. The shorter duration action potential means a shorter refractory period, which permits more heartbeats per minute.

β-receptor activation opens more Ca^{2+} channels and increases the entry of extracellular Ca^{2+} into cardiac muscle cells during an action potential. The entry of more extracellular "trigger" Ca^{2+} creates a greater stimulus for the release of Ca^{2+} stores from the sarcoplasmic reticulum. Therefore the cytosolic Ca^{2+} concentration increases more rapidly and reaches an exceptionally high level during the action potential. This leads to a quicker, stronger contraction. In addition, the duration of the contraction is shortened,

TABLE 19.2 Electrical Characteristics of the Atrioventricular (AV) Node

Characteristic (Significance)	Sympathetic Effect[a]	Parasympathetic Effect[b]
Is the only conducting pathway between atria and ventricles (directs atrial action potentials into the rapidly conducting AV bundle and bundle branches)	—	—
Has a slow conduction velocity (creates AV delay)	Increases velocity (shortens AV delay)	Decreases velocity (lengthens AV delay)
Spontaneously depolarizes to threshold (acts as auxiliary pacemaker for the ventricles)	Faster depolarization (speeds auxiliary pacemaker)	Slower depolarization (slows auxiliary pacemaker)
Has a very long refractory period (protective effects: limits maximal rate to which atria can drive ventricles and prevents ventricular action potentials from reexciting atria)	Shortens refractory period (appropriate for high heart rates)	Lengthens refractory period (appropriate for low heart rates)

[a]Through activation of β-adrenergic receptors on AV node cells.
[b]Through activation of muscarinic cholinergic receptors on AV node cells.

Vet-ebooks.com

because β-receptor activation speeds up the pumps that move cytosolic Ca^{2+} back into the sarcoplasmic reticulum and back out of the cell into the extracellular fluid. Thus even though more Ca^{2+} than normal enters the cytosol during an action potential, its removal at the end of the action potential is faster than normal. Overall, β-receptor activation makes each cardiac contraction quicker to develop, stronger, and shorter.

In summary, sympathetic nerves act (1) on the SA node pacemaker cells to increase the heart rate, (2) on the AV node cells to increase the conduction velocity and shorten the AV delay, and (3) on all cardiac cells to shorten the refractory period and to make each cardiac contraction quicker, stronger, and shorter. All these changes cause the heart to pump more blood at a higher pressure, which is an animal's normal response during exercise or emotional stress.

Because sympathetic effects on the heart are all brought about through activation of the β-adrenergic receptors on the cardiac muscle cells, the administration of any drug that activates β receptors (*β-adrenergic agonist*) has the same effects as sympathetic activation. Epinephrine, norepinephrine, and isoproterenol are common β-adrenergic agonists. Conversely, the administration of a drug that binds to and blocks β receptors reduces all the effects of sympathetic activation. Propranolol and atenolol are common *β-adrenergic antagonists.* Examples of their use are provided later.

Parasympathetic Effects Are Opposite to Those of Sympathetic Activation But Are Most Evident in the Sinoatrial Node, Atrioventricular Node, and Atria

Parasympathetic nerves affect the heart by the release of acetylcholine, which activates muscarinic cholinergic receptors on cardiac muscle cells. Qualitatively, all the effects of parasympathetic activation are opposite to those of sympathetic activation, because the effects of activating muscarinic cholinergic receptors are opposite to the effects of activating β-adrenergic receptors. Parasympathetic nerves have very powerful effects on the SA node pacemaker cells (see Fig. 19.8) and on the AV node cells (see Table 19.2). In addition, parasympathetic nerves exert strong, antisympathetic influences on all the atrial cells. However, parasympathetic nerves have relatively weak effects on the ventricular muscle cells, because very few ventricular cells receive direct parasympathetic innervation. By contrast, all ventricular muscle cells receive direct sympathetic innervation. In summary, the predominant parasympathetic influences on the heart are exerted at the SA node (to decrease the rate), at the AV node (to slow conduction and lengthen the refractory period), and on all supraventricular cells (to lengthen the refractory period and make their contractions weaker and slower).

Parasympathetic neurons do exert a curious, indirect effect on ventricular muscle cells. In the ventricles, parasympathetic neurons release their acetylcholine onto sympathetic neuron terminals. This acetylcholine activates muscarinic cholinergic receptors that are located on the sympathetic neuron terminals. The effect of this activation is to inhibit the release of norepinephrine from the terminals, which weakens the effects of sympathetic activation on ventricular cells.

Parasympathetic effects on the heart can be mimicked by the administration of a *muscarinic cholinergic agonist* (e.g., acetylcholine or muscarine) and blocked by the administration of a *muscarinic cholinergic antagonist* (e.g., atropine). Some therapeutic applications are mentioned later.

Dysfunction in the Specialized Conducting System Leads to Abnormalities in Cardiac Rhythm (Arrhythmias)

Cardiac arrhythmias result either from problems with the formation of action potentials or from problems with the propagation (conduction) of action potentials. One example of a problem with action potential formation is *sinus arrest,* in which the SA node completely fails to form action potentials. In a patient with sinus arrest, the auxiliary (emergency) pacemaker function of the AV node keeps the ventricles beating, although at an abnormally low rate. Sinus arrest is the extreme case of the condition called *sick sinus syndrome.* In its less extreme form, sick sinus syndrome is characterized by sluggish depolarization of the SA node pacemaker cells, which leads to an abnormally low intrinsic heart rate. Patients typically exhibit an abnormally low heart rate at rest (*bradycardia*) and an insufficient increase in heart rate during exercise.

Even though the problem in sick sinus syndrome is intrinsic to the sinus itself, one treatment strategy is to administer a cholinergic muscarinic antagonist drug (such as atropine) in order to block parasympathetic action on the heart. Table 19.3 illustrates the logic behind this treatment. In a normal, healthy large dog, the intrinsic rate of the heart is about 140 beats/min. However, resting heart rate is lower (about 90 beats/min) because parasympathetic tone slows the SA node pacemaker to a rate below its intrinsic rate. A drug that blocks parasympathetic effects on the heart would return the heart rate of a resting dog to 140 beats/min. A dog with a sick sinus has a low intrinsic heart rate, perhaps 90 beats/min. Parasympathetic tone makes the resting heart rate even lower, perhaps 40 beats/min. A drug that blocks parasympathetic effects restores the heart rate to its intrinsic level, 90 beats/min. This example illustrates that a dog with sick sinus syndrome, treated with atropine, can have a heart rate that closely matches the rate of a normal resting dog.

Another possible therapeutic approach is to increase the heart rate by administering a β-adrenergic agonist drug (e.g., isoproterenol). Enough isoproterenol would be given to increase the resting rate from 40 to 90 beats/min.

If drug treatment of sick sinus syndrome is ineffective, an alternative way to increase the heart rate is through the use of an *artificial cardiac pacemaker.* Such a device periodically applies an electric shock to the heart, which depolarizes cardiac muscle to threshold. Shocks applied to the atria initiate atrial action potentials. If the AV node is functioning normally, these atrial action potentials are conducted to the ventricles, and the ventricles contract. For temporary or emergency treatment, the pacemaker electrodes can

TABLE 19.3 Treatment of Sick Sinus Syndrome by Blocking Parasympathetic Effects on Heart Rate With a Cholinergic Muscarinic Antagonist

Approximate Heart Rate (beats/min)	Normal Dog	Dog With Sick Sinus Syndrome
Intrinsic rate	140	90
Resting rate (with parasympathetic tone)	90	40
Resting rate after atropine	140	90

Vet-ebooks.com

be inserted intravenously (e.g., via the jugular vein) and advanced into the right atrial chamber. For long-term treatment, a battery-powered electrical stimulator can be surgically implanted under the patient's skin and attached to electrodes that are either inserted into one of the heart's chambers or attached to the outside surface of the heart.

Atrioventricular Node Block Is a Common Cause of Cardiac Arrhythmias

Whereas sick sinus syndrome exemplifies a dysfunction of action potential formation, *AV node block* is a common dysfunction of action potential propagation (conduction). If damage to the AV node prevents (*blocks*) conduction of atrial action potentials into the ventricles, the atria continue to beat at a rate determined by the SA node pacemaker cells. The ventricles also continue to beat, but at a much lower rate. In such a case the ventricular action potentials (and contractions) are being initiated by auxiliary pacemaker cells low in the AV node (i.e., below the level of the block). Because the AV node pacemaker cells depolarize more slowly than the SA node pacemaker cells, the ventricles in a resting dog with AV node block typically beat at only 30 to 40 beats/min. Furthermore, these ventricular beats are not synchronized with the atrial contractions.

Three degrees of severity of AV node block are recognized. Complete block of the AV node, in which no atrial action potentials are conducted to the ventricles, is called *third-degree AV node block.* If action potentials are conducted sporadically from the atria to the ventricles, so that the AV node transmits some atrial action potentials but not all of them, the condition is called *second-degree AV node block.* In a patient with second-degree block, some atrial contractions are followed by ventricular contractions, and others are not. Strong parasympathetic activity can create or exaggerate second-degree AV node block because parasympathetic activity increases the refractory period of the AV node cells. For example, in quietly resting horses, parasympathetic tone is often so strong, and the AV node refractory period so long, that some atrial beats are not conducted to the ventricles. Therefore if the pulse of a relaxed, resting horse is palpated, some "missing" ventricular contractions are likely to be noticed. When this horse begins to exercise, the AV node block will likely disappear because parasympathetic activity has been decreased and sympathetic activity increased. Both these changes shorten the refractory period of the AV node and make it much more certain that every atrial action potential will be conducted to the ventricles.

Second-degree or third-degree AV node block often involves the electrical phenomenon known as *decremental conduction.* As mentioned earlier, AV node cells have "slow" action potentials, characterized by a less rapid upstroke, a lower voltage amplitude, and a slower velocity of conduction than the action potentials in regular atrial or ventricular cells. All these differences make propagation of the action potential from cell to cell less reliable in the AV node than in regular atrial or ventricular tissue. When the AV node cells are in an electrically depressed state, an atrial action potential may simply die out within the AV node and not be conducted to the ventricles. This fading and eventual stoppage of a cardiac action potential in a slowly conducting region is called decremental conduction.

The mildest degree of AV node block is *first-degree block,* in which every atrial action potential is transmitted to the ventricles, but the action potential propagates even more slowly than normal through the AV node. Therefore in first-degree block, the delay between atrial contraction and ventricular contraction is abnormally long. Behavioral state influences the amount by which AV delay is prolonged because parasympathetic activity decreases (and sympathetic activity increases) the velocity of AV node conduction.

AV node block can be caused by cardiac trauma, toxins, viral or bacterial infections, ischemia, congenital heart defects, or cardiac fibrosis. AV node block is sometimes caused by inadvertent damage of AV node tissue during a surgical repair of a ventricular septal defect.

AV node block must be treated if the resulting ventricular rate is too low to maintain adequate blood flow to the body. In such a patient, administration of a muscarinic cholinergic antagonist (e.g., atropine) may reduce the AV node refractory period and decremental conduction sufficiently to overcome the blocked state. The same effect might be achieved with a drug that activates β-adrenergic receptors (e.g., isoproterenol; see Table 19.2). If drug treatment fails to correct AV node block, an artificial pacemaker is needed. In the case of AV node block, the pacemaker needs to be applied to the ventricles; pacing the atria would not be beneficial because atrial action potentials are not being reliably conducted to the ventricles.

Tachyarrhythmias Result From Abnormal Pacemaker Activity

Tachyarrhythmias are arrhythmias in which the atrial rate or the ventricular rate (or both) is abnormally high. An occasional extra atrial or ventricular beat is called a *premature contraction* (or *precontraction*). Occasional precontractions are common both in animals and in humans and usually have no clinical significance. If the precontractions become frequent or continuous, the condition is called *tachycardia,* which means "rapid heart." Tachycardia is a heart rate that is more rapid than is appropriate for the behavioral circumstances (e.g., 160 beats/min in a resting dog). Tachycardia is a clinically significant sign.

The pacemaker initiating the rapid or "extra" beats can be the SA node itself. Alternatively, a region of dysfunctional cardiac muscle outside the SA node can act as an *ectopic pacemaker* by spontaneously depolarizing to threshold before the regular SA node pacemaker does. Common causes of ectopic pacemaker activity include cardiac infection or trauma, reaction to a drug or toxin, electrolyte imbalances, myocardial ischemia, and myocardial infarction.

The tachyarrhythmias are named for the site of the pacemaker at which they originate. Hence, if tachycardia appears to be caused by abnormally rapid depolarizations of SA node pacemaker cells, the condition is called *sinus tachycardia.* If tachycardia originates from an ectopic pacemaker within the atria, it is called *atrial tachycardia.* Atrial tachycardia is common in some canine breeds, including boxers and wolfhounds. *Junctional tachycardia* originates from ectopic pacemakers within the AV node or first part of the AV bundle. *Supraventricular tachycardia* is a collective term that encompasses sinus tachycardia, atrial tachycardia, and junctional tachycardia. If the ectopic pacemaker causing tachycardia is within the ventricles, the condition is called *ventricular tachycardia.* In this situation, the ventricles beat at a rapid rate, as dictated by the ectopic ventricular pacemaker. In occasional patients, some of the action potentials initiated by an ectopic ventricular pacemaker may be conducted backward through the AV node and may cause atrial precontractions. Usually, however, the AV node does not conduct action potentials backward; the atria continue to beat at

Vet-ebooks.com

the rate dictated by the normal SA node pacemaker. In either case, ventricular contractions are not preceded in the normal way by atrial contractions. The major dysfunction associated with ventricular tachycardia is that the ventricles do not relax long enough between contractions for adequate filling, and this problem is exacerbated by the absence of appropriately timed atrial contractions.

An extremely rapid atrial tachycardia is called *atrial flutter.* Atrial flutter does not lead to ventricular flutter because of the long refractory period of the AV node cells; the AV node conducts some, but not all, of the frequent atrial depolarizations to the ventricles. This is an example of the AV node protecting the ventricles from beating at too rapid a rate. If atrial contractions become so rapid that they lose synchrony, the condition is called *atrial fibrillation.* Atrial fibrillation is characterized by the continuous, random passage of action potentials through the atria. Fibrillating atria appear to quiver; there is no effective, coordinated contraction, and no blood is pumped. Atrial fibrillation is common in horses and in certain breeds of dogs, including Doberman pinschers. Atrial fibrillation usually does not lead to ventricular fibrillation. Because of the long refractory period of the AV node, some, but not all, of the randomly occurring atrial action potentials are propagated through the AV node. In consequence, the ventricles continue to contract with a synchronized pumping stroke despite the electrical and contractile chaos occurring in the atria. Typically, in a patient with atrial fibrillation, ventricular contractions are frequent, but irregularly spaced.

Synchronous ventricular contractions are essential for life. If the synchrony of ventricular contractions is disrupted and the ventricles begin to fibrillate, ventricular pumping stops. In *ventricular fibrillation* (*"V-fib"*), each tiny region of the ventricular wall contracts and relaxes at random in response to action potentials that propagate randomly and continuously throughout the ventricles. The condition of ventricular fibrillation is synonymous with *sudden cardiac death.*

In most cases, ventricular fibrillation can be reversed only by electrical *defibrillation.* In this process, a strong electrical current is passed briefly through the heart muscle. This current depolarizes all the cardiac cells simultaneously and holds them in a depolarized state for several milliseconds. It is hoped that when the current is turned off, all the cardiac muscle cells will simultaneously repolarize to a resting membrane potential, so that the normal pacemaker of the heart will again have a chance to initiate beats in an organized and synchronized manner. Sometimes it works; however, if the cardiac problems that caused ventricular fibrillation to develop in the first place are still present, fibrillation is likely to recur. Usually, defibrillation is performed by placing stimulating electrodes *(paddles)* on either side of the thorax. Therefore the stimulating current passes through, and depolarizes, the skeletal muscles of the thorax as well as the cardiac muscle of the heart. The resulting, involuntary contraction of the skeletal muscles causes the patient to "jump" at the moment of defibrillation.

Ectopic pacemaker activity is likely to arise when a region of ischemic or damaged cardiac muscle develops the abnormal, twin properties of slow conduction of action potentials and an ability to conduct action potentials in only one direction. Fig. 19.11 illustrates how a region of slow, one-way conduction in the wall of one cardiac chamber can function as an ectopic pacemaker. The process begins with a normally originating action potential arriving at the damaged or ischemic region. Because of the property of one-way conduction, the action potential can only enter the abnormal region from one side. Because of the property of slow conduction through the abnormal region, all the normal, surrounding muscle may be past its refractory period by the time the action potential emerges from the abnormal region. The emerging action potential can then trigger another action potential in the normal muscle. If this second action potential then propagates around the cardiac chamber and back into the abnormal region, a self-perpetuating cycle can develop. The action potential again propagates slowly through the abnormal region, and again it emerges from the abnormal region after the normal muscle is past its refractory period. The result is a sequence of *reentrant* action potentials, each one initiating a contraction (an "extra" beat) as it propagates through the normal cardiac muscle. The reentrant pathway does not necessarily have to be all the way around the circumference of a cardiac chamber. A damaged, ischemic, or infarcted area of cardiac muscle can form the nonconducting center around which reentrant action potentials can travel. The passage of an action potential around and around a nonconducting center is called a *circus movement.* For the circus movement of the action potential to be self-regenerating, however, a portion of the circular, conducting pathway must have the twin properties of slow and one-way conduction. In effect, an area of slow, one-way conduction within a circular conducting pathway functions as an ectopic pacemaker. Reentry of cardiac action potentials can lead to occasional precontractions, continuous tachycardia, or even fibrillation. In any of these cases, the resulting tachyarrhythmia is called a *reentrant arrhythmia.*

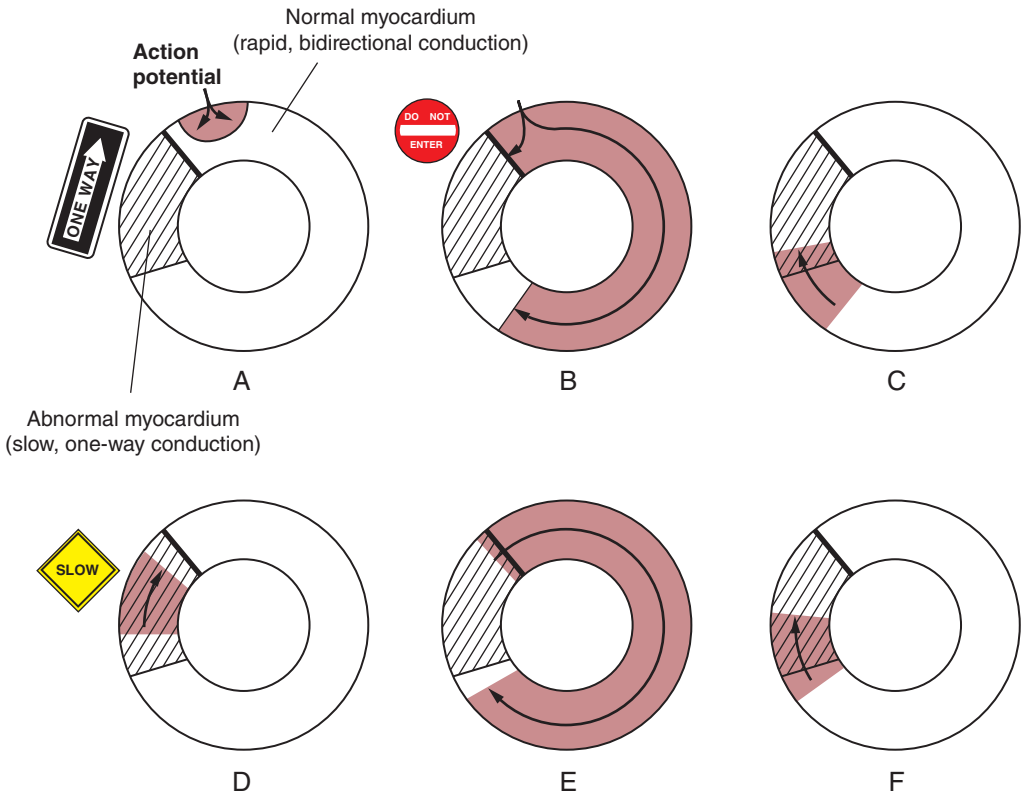
Antiarrhythmic Drugs Affect the Ion Channels Responsible for Cardiac Action Potential Formation and Propagation

Whereas ventricular fibrillation is generally lethal without electrical defibrillation, other tachycardias can often be treated successfully with *antiarrhythmic drugs.* Because tachyarrhythmias result from extra cardiac action potentials, effective antiarrhythmic drugs must work by counteracting either the formation or the propagation of the extra action potentials.

Local anesthetics (e.g., quinidine, lidocaine, procaine) constitute one category of antiarrhythmic drugs. They act by binding to some of the voltage-gated Na^+ channels (fast Na^+ channels) in cardiac muscle cells and preventing them from opening, which counteracts membrane depolarization and action potential formation. In effect, blocking some of the Na^+ channels raises the threshold for action potential formation. This tends to "quiet" ectopic pacemakers and to stifle reentrant arrhythmias. Na^+ channel blockers such as lidocaine or procaine (Novocain) are called *local anesthetics* because, when applied to sensory neurons, they prevent the propagation of neural action potentials that would signal pain to the brain. The cardiac, antiarrhythmic effect of local anesthetics is not the result of their blockage of pain pathways.

Calcium channel blockers constitute a second category of antiarrhythmic drugs. Examples include verapamil, diltiazem, and nifedipine. These drugs bind to some of the L-type (slow) Ca^{2+} channels and prevent them from opening, which decreases the entry of Ca^{2+} into cardiac muscle cells during an action potential. Because Ca^{2+} entry is the primary depolarizing influence during the plateau (phase 2) of the cardiac action potential, one major effect of a Ca^{2+} channel blocker is to lower the plateau (make the membrane potential less positive). A secondary consequence is to lengthen the action potential. The action potential is longer because of a complicated effect of the height of the plateau on K^+ channels, as discussed earlier in connection with sympathetic effects on cardiac action potentials. Lengthening the cardiac action potential also lengthens the refractory period, which makes it less likely that

Vet-ebooks.com



• **Fig. 19.11** Cross-section of a cardiac chamber (atrium or ventricle) is pictured at six successive instants to illustrate how reentrant arrhythmias occur. A region of abnormal myocardium (*cross-hatched* area) conducts action potentials slowly and only in one direction (*clockwise* in this example). *Colored shading* indicates areas of heart where an action potential is underway. A, Normal action potential has just entered this ring of tissue, and only the *shaded area* is depolarized. B, Action potential propagates rapidly in both directions through the normal cardiac tissue but cannot propagate into the abnormal myocardium in a counterclockwise direction. C, The clockwise-going action potential can enter the abnormal region. D, While the action potential propagates slowly, in a clockwise direction through the abnormal region, the normal cardiac tissue repolarizes to a resting state (indicated by *lack of shading*). E, Action potential emerges from the abnormal region into normal cardiac tissue and propagates through the normal tissue for a second time. Meanwhile, the abnormal tissue repolarizes to a resting state. F, Action potential begins to move slowly through the abnormal region for a second time. States D, E, and F repeat themselves. Thus the abnormal region functions as an ectopic pacemaker.

early extra action potentials will be formed in ectopic pacemakers, or that they will propagate even if they are formed.

Particular calcium channel blockers have especially strong effects on the cells of the SA and AV nodes. The consequent increase in refractory period decreases the likelihood that pacemaker cells will form "extra" action potentials. An increased refractory period in the AV node is especially effective in limiting the ventricular rate in cases of atrial flutter or fibrillation. Many of the extra atrial action potentials simply *die out* (through decremental conduction) in the AV node.

By reducing the entry of extracellular Ca^{2+} into cardiac muscle cells during an action potential, Ca^{2+} channel blockers not only suppress tachyarrhythmias, but also decrease the strength of cardiac contractions. Less entry of extracellular "trigger" Ca^{2+} means a less powerful stimulus for the release of stored Ca^{2+} from the sarcoplasmic reticulum. Therefore the cytosolic Ca^{2+} concentration does not increase as much as normal during the action potential, so there is a less forceful contraction. Some clinical situations in which it is desirable to decrease cardiac contractility are discussed in Chapter 21.

The *cardiac glycosides* (e.g., digitalis) constitute a third category of antiarrhythmic drugs. Cardiac glycosides act by inhibiting the Na^+, K^+ pump in cell membranes. As mentioned in Chapters 1 and 4, the Na^+, K^+ pump functions to keep intracellular Na^+ concentration low and intracellular K^+ concentration high. The pump also indirectly supplies energy to transport Ca^{2+} back out of cardiac cells after it enters during an action potential. Inhibition of the Na^+, K^+ pump with a cardiac glycoside has several important effects on cardiac function. The effects are listed here without much explanation because the mechanisms are quite complex. First, cardiac muscle cells do not repolarize fully at the end of an action potential; the resting membrane potential is not as negative as normal. As a consequence, some Na^+ channels remain inactivated, which makes the cells somewhat refractory with regard to the formation of subsequent action potentials. This tends to quiet ectopic pacemakers. Second, effects on the central nervous system lead to an increase in parasympathetic tone. This slows the heart rate, quiets atrial ectopic pacemakers, slows conduction through the AV node, and increases the refractory period of AV node cells. The overall effect is to suppress ectopic atrial action potentials or

Vet-ebooks.com

cause extra atrial action potentials to die out in the AV node and not to be conducted to the ventricles. A third effect of cardiac glycosides is to allow more Ca^{2+} than normal to accumulate inside cardiac cells, resulting in stronger cardiac contractions. In summary, the cardiac glycosides are antiarrhythmic and increase cardiac contractility.

Beta-adrenergic antagonists (e.g., propranolol) constitute a fourth class of antiarrhythmic drug. Beta (β) blockers, as they are called, bind to some of the β-adrenergic receptors on cardiac cells and prevent their activation by norepinephrine from sympathetic nerves or by circulating epinephrine and norepinephrine from the adrenal medulla. Sympathetic activation tends to promote tachyarrhythmias by increasing heart rate, shortening refractory period, and speeding conduction of action potentials, especially through the AV node. Beta blockers reduce these effects and therefore reduce the likelihood that extra action potentials will form or propagate. An additional effect of β blockers is to reverse sympathetic-induced increases in cardiac contractility.

In summary, of the four categories of drugs used to treat tachyarrhythmias, three also have pronounced effects on cardiac contractility. Calcium channel blockers and β blockers decrease cardiac contractility, whereas cardiac glycosides increase contractility. Local anesthetics have little effect on contractility. This variety of effects allows a clinician to select the type of antiarrhythmic drug that is best matched to each patient's cardiac contractile state.

Electrical dysfunction of the heart has been discussed in considerable detail to illustrate how specific abnormalities in the specialized cardiac conduction system can result in specific and serious arrhythmias. Electrical dysfunction of the heart is encountered often in clinical practice, and its consequences are often serious or even lethal. Chapter 20 is devoted to an explanation of the electrocardiogram, which is the most commonly used tool for evaluating electrical dysfunction of the heart.

CLINICAL CORRELATIONS

Third-Degree Atrioventricular Block

History

A 5-year-old male English bulldog has fainted several times during the past 3 weeks. On each occasion, he collapsed, was apparently unconscious for a few seconds, and then slowly recovered. These episodes occur most often during exertion. In general, he tends to be less active than normal, but he has no other obvious signs of illness.

Clinical Examination

The dog is moderately obese. There are no obvious neurological deficits. His mucous membranes are pink, and the capillary refill time is normal (1.5 seconds). Auscultation of the chest reveals a slow, regular heart rate of 45 beats/min. The femoral pulse rate is also 45 beats/min and strong. Thoracic radiography reveals a mildly enlarged heart, but the radiographs are otherwise within normal limits.

The electrocardiogram (ECG) reveals a disparity between the atrial rate (atrial depolarizations occurring regularly, 140 per minute) and the ventricular rate (ventricular depolarizations occurring regularly, 45 per minute). There is no consistent time interval between the atrial and ventricular depolarizations.

Comment

As will be discussed in Chapter 20, atrial and ventricular depolarizations produce characteristic voltage fluctuations at the body surface, which are detected by the ECG. The ECG of this dog shows a complete dissociation between atrial and ventricular depolarizations, which provides definitive diagnostic evidence of complete (third-degree) AV node block. The dog's atria are depolarizing 140 times per minute in response to action potentials being initiated in the normal manner by pacemaker cells of the SA node. However, the atrial action potentials are not being conducted through the AV node. Ventricular action potentials are being initiated, at the slow rate of 45 per minute, by auxiliary pacemaker cells located below the blocked region of the AV node.

The low ventricular rate in this dog allows a longer-than-normal time for ventricular filling between beats. Therefore the volume of blood ejected by the ventricles with each beat (the stroke volume) is greater than normal. This increased stroke volume causes the femoral pulse to be very strong.

In a normal dog, sympathetic and parasympathetic nerves acting on the SA node pacemaker cells adjust the heart rate so that cardiac output is matched to the metabolic requirements of the body. In a dog with complete AV block, the ventricles do not respond to these autonomically mediated changes in SA node pacemaker rate. In such a case, the rate of ventricular contractions is characteristically low at rest and does not increase much during exercise. Therefore cardiac output does not increase enough during exertion to meet the increased metabolic needs of exercising skeletal muscle. As a consequence, arterial blood pressure decreases during attempted exercise and brain blood flow may fall below the level needed to sustain consciousness. The dog faints.

Treatment

Drug therapy for AV node block involves either blocking the effects of parasympathetic nerves on the AV node (with a muscarinic cholinergic antagonist drug such as atropine) or mimicking the effects of sympathetic activation (with cautious use of a β-adrenergic agonist such as isoproterenol or dopamine). The rationale for these treatments is based on the following physiology: AV node block occurs because atrial action potentials die out in the AV node (*decremental conduction*). Parasympathetic activation increases the tendency for decremental conduction because parasympathetic nerves act on AV node cells to increase their refractory period and to decrease the velocity with which action potentials spread from cell to cell. Therefore blocking parasympathetic effects is occasionally effective in reversing AV node block. In contrast, sympathetic activation decreases the tendency for decremental conduction by decreasing the refractory period of AV node cells and increasing their conduction velocity. A sympathomimetic drug (one that mimics sympathetic effects by activating β-adrenergic receptors) may therefore unblock the AV node. Even if administration of a sympathomimetic drug does not reverse the AV node block, it usually increases the rate of the auxiliary (emergency) pacemaker cells in the AV node or bundle, which are initiating the ventricular contractions. The increased ventricular rate improves cardiac output.

Many cases of third-degree AV block cannot be managed effectively with drugs, so an artificial ventricular pacemaker must be installed. The procedure is straightforward; pacemaker electrodes can be placed into the right ventricle via a systemic vein (e.g., external jugular) with only sedation and local anesthesia. The electrode wires are attached to a battery-powered pacemaker unit that is then implanted under the skin. The pacemaker delivers pulses of electric current via the electrodes at a rate set by the clinician. Each current pulse depolarizes right ventricular muscle. The resulting ventricular action potential is then propagated from cell to cell throughout both the right and left ventricles, and both ventricles contract.

Vet-ebooks.com

Practice Questions

1. Which of the following is *true* for *both* cardiac muscle and skeletal muscle?
- **a.** The muscle forms a functional syncytium.
- **b.** An action potential in the muscle cell membrane is required to initiate contraction.
- **c.** Pacemaker cells spontaneously depolarize to threshold and initiate action potentials.
- **d.** Frequent action potentials in motor neurons can cause a sustained (tetanic) muscle contraction.
- **e.** Extracellular Ca^{2+} enters the muscle cell during an action potential and triggers the release of additional Ca^{2+} from the sarcoplasmic reticulum.

2. The normal pathway followed by a cardiac action potential is to begin in the SA node and then propagate:
- **a.** Across the atria in the bundle of His.
- **b.** Through the connective tissue layers that separate the atria and ventricles.
- **c.** Across the atria and into the AV node.
- **d.** From the left atrium to the right atrium.
- **e.** From the left atrium to the left ventricle and from the right atrium to the right ventricle.

3. At the moment when an action potential begins propagating slowly through the AV node in a normal resting dog, the ventricular muscle cells are:
- **a.** At their resting membrane potential.
- **b.** Depolarizing slowly toward threshold for formation of an action potential.
- **c.** Undergoing rapid depolarization at the beginning of an action potential.
- **d.** At the plateau of an action potential.
- **e.** Just ending their action potential (i.e., repolarizing back toward resting membrane potential).

4. During which phase of a normal ventricular action potential is it most likely that fast Na^+ channels are in an inactivated state, slow Ca^{2+} channels are open, and most K^+ channels are closed?
- **a.** Phase 0 (rapid depolarization)
- **b.** Phase 1 (partial repolarization)
- **c.** Phase 2 (plateau)
- **d.** Phase 3 (repolarization)
- **e.** Phase 4 (rest)

5. An increase in heart rate could result from:
- **a.** An increase in sympathetic nerve activity to the heart.
- **b.** A decrease in parasympathetic nerve activity to the heart.
- **c.** An abnormally rapid decrease in permeability of SA node cells to K^+ during diastole.
- **d.** An abnormally rapid increase in permeability of SA node cells to Na^+ during diastole.
- **e.** All the above are correct.

6. In which of the following arrhythmias will there be more atrial beats per minute than ventricular beats?
- **a.** Complete (third-degree) AV block
- **b.** Frequent premature ventricular contractions
- **c.** Sick sinus syndrome (sinus bradycardia)
- **d.** First-degree AV block
- **e.** Ventricular tachycardia

7. Which of the following types of drugs would be the best choice to treat a patient with both supraventricular tachycardia and inadequate cardiac contractility?
- **a.** Local anesthetic (fast Na^+ channel blocker)
- **b.** Muscarinic cholinergic antagonist
- **c.** Beta-adrenergic agonist
- **d.** Cardiac glycoside (inhibits Na^+, K^+ pump)
- **e.** Calcium channel blocker

Bibliography

Baca WJ Jr, Bacha LM. *Color Atlas of Veterinary Histology*. 3rd ed. Wiley-Blackwell; 2012.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.

Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. Elsevier; 2017.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Elsevier; 2015.

Katz AM. *Physiology of the Heart*. 5th ed. Baltimore: Lippincott, Williams & Wilkins; 2010.

Koeppen BM, Stanton BA. *Berne & Levy Physiology*. 7th ed. Elsevier; 2017.

Lilly LS, ed. *Pathophysiology of Heart Disease: A Collaborative Project of Medical Students and Faculty*. 6th ed. Lippincott, Williams & Wilkins; 2015.

Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.

Reece WO, Erickson HH, Goff JP, eds. *Dukes' Physiology of Domestic Animals*. 13th ed. Wiley-Blackwell; 2015.

Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2015.

Tilley LP, Smith FWK, eds. *Blackwell's Five-Minute Veterinary Consult: Canine and Feline*. 6th ed. Wiley-Blackwell; 2015.

Vet-ebooks.com

20 The Electrocardiogram

ROBERT B. STEPHENSON

KEY POINTS

1. The electrocardiogram (ECG) is the most frequently used clinical tool for diagnosing electrical dysfunctions of the heart.
2. Atrial depolarization, ventricular depolarization, and ventricular repolarization cause characteristic voltage fluctuations in the electrocardiogram.
3. The electrocardiogram reveals the timing of electrical events in the heart.
4. Six standardized electrocardiographic leads are used in veterinary medicine.
5. Abnormal voltages in the electrocardiogram are indicative of cardiac structural or electrical abnormalities.
6. Electrical dysfunctions in the heart cause abnormal patterns of electrocardiogram waves.
7. In large animals, there is considerable variability in the polarity and amplitude of the electrocardiogram waves.

The Electrocardiogram Is the Most Frequently Used Clinical Tool for Diagnosing Electrical Dysfunctions of the Heart

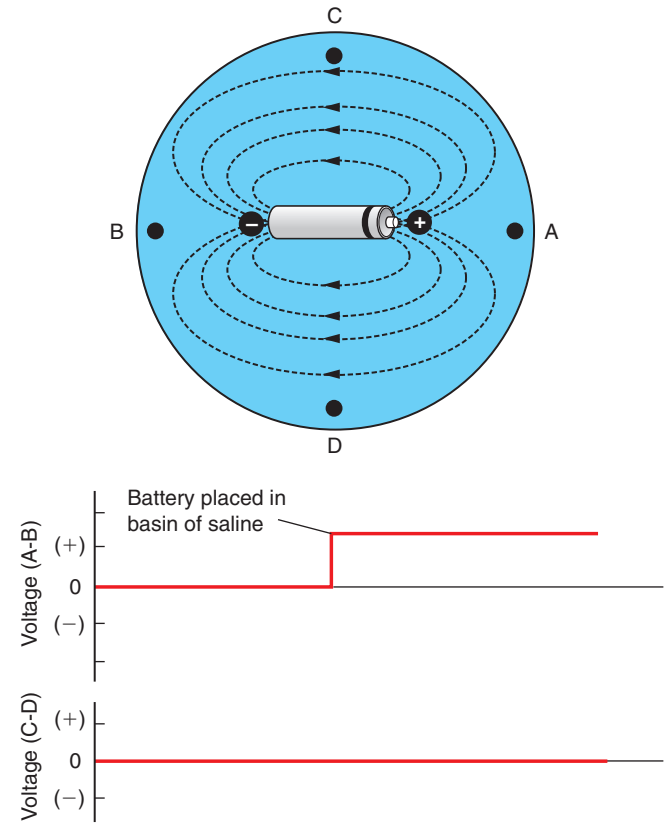
An *electrocardiogram* is simply a graph of tiny voltage fluctuations at the surface of the body that are caused by the passage of action potentials from cell to cell through the heart. The specialized voltmeter that detects and graphs these voltage fluctuations is called an *electrocardiograph.* The procedure for making and interpreting an electrocardiogram is called electrocardiography. The same abbreviation "*ECG*" is used for all three: electrocardiogram, electrocardiograph, and electrocardiography. To make matters more confusing, many people use the abbreviation "*EKG*" instead of ECG, because the original name, *Elektrokardiogramm* (German), was given by its inventor, Willem Einthoven, a Dutch physiologist and Nobel Laureate. In its most common application, two or more metal electrodes are applied to the skin surface, and the voltages recorded by the electrodes are displayed on a video screen or drawn on a paper strip. The physics of how the heart produces voltages detectable at the body surface is extraordinarily complex. However, it is not difficult to develop an intuitive understanding of how electrocardiography works; this intuitive approach is adequate for most clinical applications.

An understanding of the ECG begins with the concept of an *electrical dipole* in a *conductive medium* (Fig. 20.1). A dipole is a pair of electrical charges (a positive charge and a negative charge) separated by a distance. A common flashlight battery is a good example of a dipole. A battery has an excess of positive charges at one end (the "+" end) and an excess of negative charges at its other end (the "–" end), and the two ends are separated by a distance. If this dipole (battery) is placed in a conductive medium (e.g., a bowl containing a solution of sodium chloride in water), ionic currents will flow through the solution. Positive ions (Na^+) in the solution flow toward the negative end of the dipole, and negative ions (Cl^-) flow toward the positive end. The flow of ions creates voltage differences within the salt solution. These voltage differences can be detected by placing the electrodes of a simple voltmeter at the perimeter of the salt solution. As depicted in Fig. 20.1, an electrode placed at point A is closer (more exposed) to the positive end of the dipole, and an electrode at point B is closer (more exposed) to the negative end of the dipole. Therefore the voltage at point A will be positive in comparison with the voltage at point B. That is, the voltmeter would detect a positive voltage difference between point A and point B. Using *V* as an abbreviation for *voltage,* we would summarize this condition by saying, "$V_{A–B}$ is positive." Points C and D are equally near (equally exposed to) the positive and negative ends of the dipole, so no voltage difference would exist between electrodes placed at points C and D. We would say, "$V_{C–D}$ is zero."

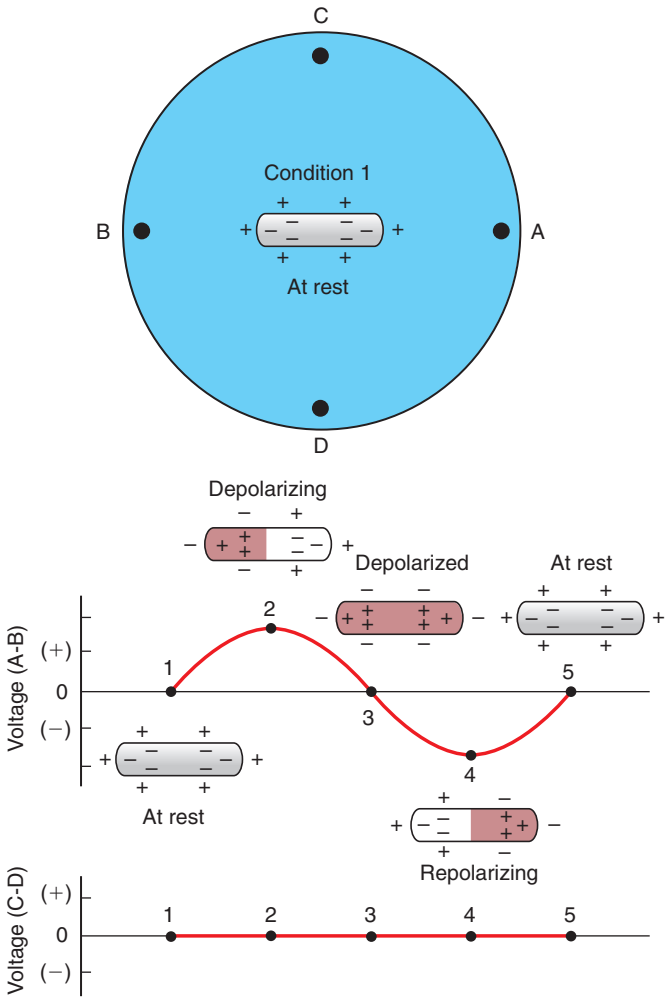
In Fig. 20.2, the battery in the NaCl solution has been replaced with an elongated strip of cardiac muscle. Again, a voltmeter is set up to detect any voltage differences that are created at point A compared with point B, and at point C compared with point D. The voltage differences (A–B and C–D) are plotted for five different conditions. In *condition 1,* all the cells in the strip of cardiac muscle are at a resting membrane potential; each cell is charged negatively on its inside and positively on its outside. Because cardiac cells are electrically interconnected by gap junctions, the strip of cardiac muscle behaves electrically as if it were one large cell (a functional syncytium). From the outside, the strip of cells looks like one large cell that is symmetrically charged positively around its perimeter. Therefore no dipole exists. There would be no voltage difference between point A and point B (i.e., $V_{A–B}$ would be zero). There would also be no voltage difference between point C and point D (i.e., $V_{C–D}$ would also be zero).

In *condition 2,* a pacemaker cell at the left end of the muscle strip has depolarized to threshold level and formed an action potential, and the action potential is propagating from cell to cell, through the muscle strip, from left to right. In other words, the cells at the left end of the strip are depolarized and are at the plateau of their action potential, whereas the cells at the right end of the strip are still at their resting membrane potential. The outside of each depolarized cell is charged negatively, whereas the outside of each resting cell is still charged positively. Under this condition, the strip of muscle has created an electrical dipole, positive at its

Vet-ebooks.com



• **Fig. 20.1** If an electric dipole (e.g., battery) is placed into a conductive medium (e.g., solution of NaCl in water), the charge difference between the two ends of the dipole (battery) will cause positive ions (Na^+) to flow within the solution, as indicated by the *dashed lines* and *arrows.* Negative ions (Cl^-) will flow in the opposite direction. These ionic currents will create voltage differences within the solution. A simple voltmeter can be used to detect these voltage differences, as shown in the lower graphs. In this example, the ionic currents would create a positive voltage at point A compared with point B, because point A is "exposed to more positive" than is point B (i.e., voltage A–B is positive). No voltage difference would exist between point C and point D, because these two points are "equally exposed to positive" (i.e., voltage C–D is zero).



• **Fig. 20.2** Strip of cardiac muscle cells in sodium chloride solution produces voltage differences between point A and point B during a phase of spreading depolarization or spreading repolarization, but not when all the cells are in a uniform state of polarization (i.e., not when all the cells are at rest or when all the cells are depolarized). No voltage difference is created between point C and point D. See text for a complete description.

right end and negative at its left end. Therefore a positive voltage would exist at point A compared with point B. Note, however, that the voltage at point C compared with point D would still be zero, because neither of these points is closer to the positive end of the dipole. The graphs in Fig. 20.2 summarize condition 2 by showing that $V_{A–B}$ is positive at this time, and $V_{C–D}$ is zero.

In *condition 3,* the entire muscle strip is depolarized; that is, all the cells are at the plateau of their action potential with a uniform negative charge outside of each cell. Therefore no voltage differences exist around the perimeter of the muscle strip. No dipole exists, so the recorded voltages (A–B and C–D) are both zero.

In *condition 4,* the muscle strip is repolarizing; cells at the left end have returned to a resting state, whereas cells at the right end are still at the plateau of their action potential. Under this condition, the outside of the muscle strip is charged negatively at its right end and positively at its left end. A dipole exists, with the voltage at point A being negative compared with point B. That is, $V_{A–B}$ is negative. The dipole does not create a voltage difference between C and D, so $V_{C–D}$ is still zero.

In *condition 5,* all the cells in the muscle strip have returned to a resting state (same as condition 1). So again, $V_{A–B}$ is zero and $V_{C–D}$ is zero.

Note that if the depolarization (in condition 2) had been spreading from right to left in the muscle strip (instead of from left to right), the voltage at point A compared with point B ($V_{A–B}$) would have been negative. Likewise, if the repolarization (in condition 4) had been spreading from right to left in the muscle strip, $V_{A–B}$ would have been positive during this repolarization. Table 20.1 summarizes these relationships.

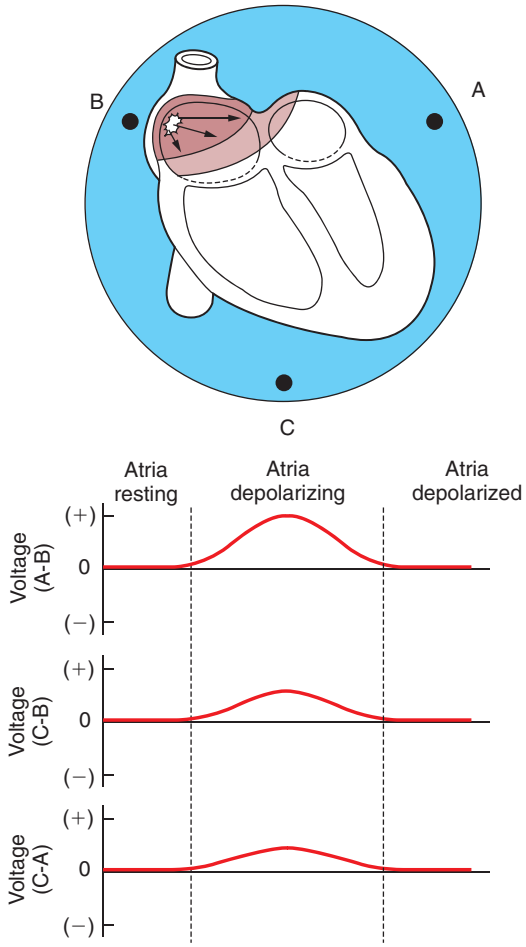
Fig. 20.3 takes the intuitive model of the ECG one step further by picturing the entire heart (rather than a strip of cardiac muscle) in the bowl of saline. The graphs below the drawing show the voltage differences that would be detected by electrodes at the perimeter of the bowl. The plots start at a time between cardiac contractions, when all the cells in the heart are at a resting membrane potential. Every cardiac cell is charged negatively on the inside of its membrane and positively on the outside. Therefore all around the entire heart, viewed as one large cell, the charge would be

Vet-ebooks.com

TABLE 20.1 Polarity of Voltage Created at Point A Compared With Point B (V_{A-B}) When a Strip of Cardiac Muscle Within a Conductive Medium is Depolarizing or Repolarizing[a]

	Depolarization	Repolarization
Approaching A	V_{A-B} is +	V_{A-B} is –
Going away from A	V_{A-B} is –	V_{A-B} is +

[a]The arrangement of muscle and electrodes is depicted in Fig. 20.2.



• **Fig. 20.3** A heart, placed in sodium chloride solution, would not create voltage differences among electrodes A, B, and C during the resting period between heartbeats. However, during depolarization of the atria (pictured at *top*), a positive voltage would be created at point A compared with point B. Atrial depolarization would also create positive voltages at point C compared with point B and at point C compared with point A. See text for a complete description.

positive, and no voltage differences would exist between any of the electrodes.

When pacemaker cells in the sinoatrial (SA) node depolarize to threshold level, they initiate an action potential that propagates from cell to cell, outward from the SA node. At the moment depicted in the top diagram of Fig. 20.3, most right atrial cells (shaded) are at the plateau of their action potential (i.e., negatively charged on their outside), whereas the cells in the left atrium and the cells in the inferior part of the right atrium (unshaded) are still at rest (i.e., positively charged on their outside). The arrows in Fig. 20.3 (*top*) indicate that the depolarizing atria create dipoles that are directed from the right atrium toward the left atrium (i.e., toward point A) and also downward in the atria (i.e., toward point C). The graphs in Fig. 20.3 (*bottom*) show that these depolarization dipoles create a voltage that is positive at point A compared with point B, and also a voltage that is positive at point C compared with point B. Atrial depolarization also creates a positive voltage at point C compared with point A, although the reason for this is admittedly not evident from the two-dimensional view of the atria depicted in Fig. 20.3. The graphs also show that, once the atria are completely depolarized (with every atrial cell at the plateau of its action potential), the voltage differences between all points return to zero.

Atrial Depolarization, Ventricular Depolarization, and Ventricular Repolarization Cause Characteristic Voltage Fluctuations in the Electrocardiogram

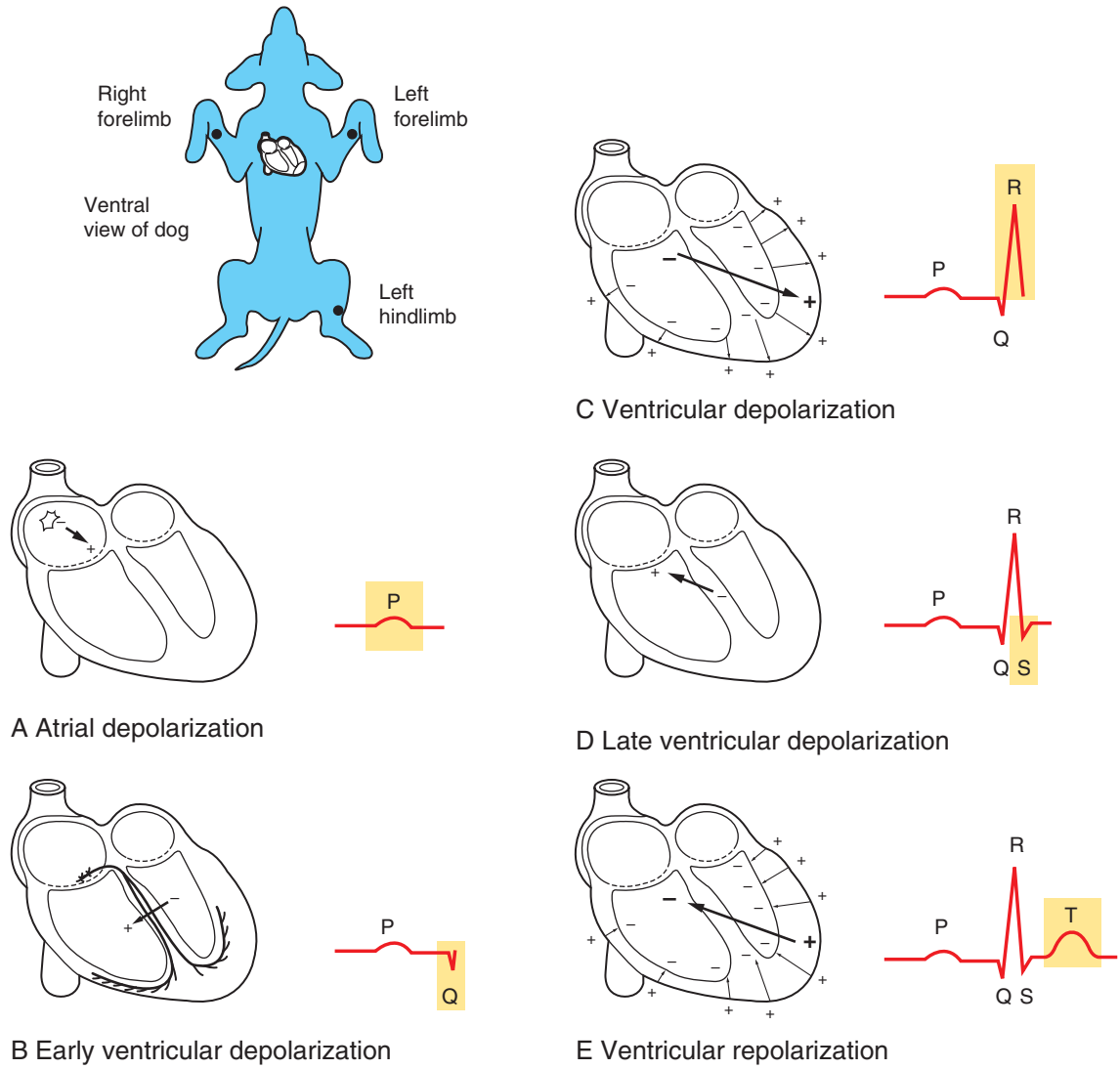
In Fig. 20.4, the heart is pictured in its normal position in the thorax of a dog. The extracellular fluids of the body contain NaCl (and other salts) in solution, so the body can be imagined as a substitute for the bowl of saline shown in the previous figures. The positions of the left forelimb, right forelimb, and left hind limb in Fig. 20.4 correspond with points A, B, and C in Fig. 20.3. Fig. 20.4A shows that, while atrial depolarization is in progress at the beginning of a heartbeat, there would be a positive voltage in the left forelimb compared with the right forelimb. This is simply a repetition of the idea illustrated in Fig. 20.3, the left forelimb being equivalent to point A and the right forelimb equivalent to point B.

The deflection in the ECG trace during atrial depolarization is called the *P wave*. At the end of atrial depolarization (i.e., at the end of the P wave), the ECG voltage returns to zero. At this moment during a normal cardiac cycle, the action potential is propagating slowly, from cell to cell, through the atrioventricular (AV) node and the first part of the AV bundle. However, these tissues have such a small mass that their depolarization generally does not create a voltage difference detectable at the body surface.

The next voltage differences that are detectable at the body surface are those associated with the depolarization of the ventricles. The first part of ventricular depolarization usually involves a depolarization that spreads from left to right (i.e., dog's left to dog's right) across the interventricular septum, as shown in Fig. 20.4B. This first phase of ventricular depolarization usually causes a small voltage difference (*Q wave*) between the left forelimb and the right forelimb, with the left forelimb being slightly negative compared with the right.

The next event in ventricular depolarization usually causes a large, positive voltage (*R wave*) at the left forelimb compared with the right, as depicted in Fig. 20.4C. To understand why this R wave is large and positive, recall that during ventricular depolarization, the left and right bundle branches conduct the spreading action potential to the ventricular apex. From there, Purkinje fibers carry the action potential rapidly up the inside walls of both ventricles. From there, the depolarization spreads from cell to cell, outward through the walls of both ventricles, as pictured by the small arrows in Fig. 20.4C. This outward-spreading action potential

Vet-ebooks.com



• **Fig. 20.4** Genesis of normal ECG. Upper left: Ventral view of heart in chest of dog. Panels A—E: Sequence of five time-lapse "snap shots" of the heart as a normal cardiac action potential is propagated through the atria and ventricles. Arrows indicate the orientation and relative strength of the electrical dipole created in the heart at each step in the sequence. The shifting electrical dipoles create a predictable sequence of voltage fluctuations (*waves*) detectable at the body surface. The red graphs depict the normal pattern of fluctuations in voltage in the left forelimb compared to the right forelimb (analogous to voltage at point B compared to point A in Fig. 20-3). By convention, each voltage "wave" in the ECG sequence is designated with a letter (P – T), as indicated. See text for more detail.

creates dipoles in each region of the ventricular wall. Therefore each small arrow in Fig. 20.4C can be considered a dipole. The positive end of each small dipole points outward in the ventricular wall because the inner portion of the ventricular wall depolarizes before the outer portion. The net electrical effect of depolarizations spreading outward through the walls of both ventricles is a large electrical dipole pointed diagonally downward (caudad) and toward the dog's left. This *net dipole* (vector sum of the small dipoles) is depicted by the bold arrow in Fig. 20.4C. Although the outward-pointing dipoles in the right ventricular wall would tend to cancel the outward-pointing dipoles in the left ventricular wall, the net dipole points toward the left for two reasons. First, the cardiac axis is tilted toward the left (i.e., the normal orientation of the heart is with the ventricular apex angled toward the left wall of the thorax). Second, the left ventricle is much more massive than the right ventricle, so the dipoles created by depolarizations spreading outward in the massive wall of the left ventricle dominate electrically over the dipoles created by depolarizations spreading outward in the thinner wall of the right ventricle. The net result is a large, positive voltage (R wave) in the left forelimb compared with the right. The R wave is the predominant feature of a normal ECG. Abnormalities in the magnitude or polarity of the R wave have great diagnostic significance, as explained later.

As the depolarizations finish spreading outward through the walls of both ventricles, the voltage in the left forelimb compared with the right forelimb returns to zero and then often becomes slightly negative for a few milliseconds (as pictured in Fig. 20.4D). The physical basis of this small, negative S wave is obscure. After the S wave, the voltage in the left forelimb compared with the right forelimb returns to zero and stays there for a time, because all the cells throughout both ventricles are uniformly at the plateau of their action potential; therefore no dipole exists.

Altogether, the process of *ventricular depolarization* produces a pattern of voltages in the ECG called the *QRS wave* (or *QRS complex*). The

Vet-ebooks.com

important feature to understand about the QRS complex is why its predominant component, the R wave, is normally large and positive.

Fig. 20.4E shows that repolarization of the ventricular muscle causes a voltage deflection in the ECG called the *T wave.* Whereas the wave of *depolarization* spreads outward through the walls of both ventricles, the pattern of *repolarization* is not as predictable. Fig. 20.4E illustrates one common pattern, in which the repolarization spreads inward through the walls of both ventricles; that is, the outermost portion of the ventricular wall was the last ventricular tissue to depolarize but the first to repolarize. The inward-going repolarization creates dipoles, as depicted by the small arrows in Fig. 20.4E, with their negative end pointed toward the inside surface of both ventricles. The net dipole during this repolarization has its negative end pointed upward (craniad) and toward the dog's right, as depicted by the bold arrow in Fig. 20.4E. This net dipole creates a positive voltage in the left forelimb compared with the right forelimb (*T wave*). The net dipole in Fig. 20.4E points toward the dog's right, simply because the left ventricular wall is so much more massive than the right ventricular wall. That is, the repolarization proceeding from outside to inside in the massive walls of the left ventricle creates larger voltages (stronger dipoles) than the repolarization proceeding from outside to inside in the thinner walls of the right ventricle.

In many normal dogs, ventricular repolarization proceeds in the same direction as the depolarization (from inside the ventricles to outside). This pattern of repolarization creates a negative voltage in the left forelimb compared with the right forelimb; that is, the T wave is negative. Whether positive or negative, T waves are caused by repolarization of the ventricles.

To summarize, the P wave is caused by atrial depolarization, the QRS complex by ventricular depolarization, and the T wave by ventricular repolarization. The pattern of ventricular repolarization varies from dog to dog, so the T wave may be positive or negative. Atrial repolarization does not cause an identifiable wave in the normal ECG, because atrial repolarization does not proceed in an orderly enough pattern or direction to create a significant net electrical dipole.

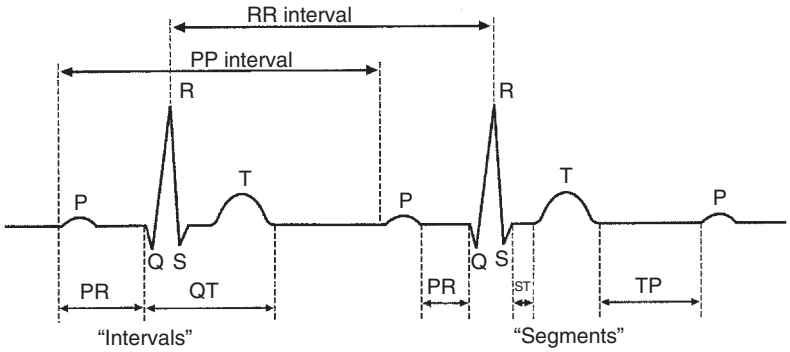
The Electrocardiogram Reveals the Timing of Electrical Events in the Heart

Because the predominant waves in an ECG correspond to specific electrical events in the heart, the time between these waves can be measured to determine the timing of events in the heart. Fig. 20.5 indicates the conventions used to define the important *intervals* and *segments* in the ECG. The *PR interval* corresponds to the time between the start of atrial depolarization (start of P wave) and the start of ventricular depolarization (start of QRS complex). The PR interval is typically about 0.13 second in a large, resting dog. During this time, the cardiac action potential is conducted slowly through the AV node. The duration of the QRS complex corresponds to the time it takes for the ventricles to depolarize, once the cardiac action potential emerges from the AV node and AV bundle. Typically this is less than 0.1 second. The *QT interval* (beginning of Q wave to end of T wave) corresponds to the time from the beginning of ventricular depolarization to the end of ventricular repolarization. This approximates the duration of an action potential in ventricular tissue. Typically the QT interval is about 0.2 second. The time between successive P waves (*PP interval*) corresponds to the time between atrial depolarizations (and thus atrial contractions). The PP interval can be used to calculate the number of atrial contractions per minute (the atrial rate), as illustrated in Fig. 20.5. Likewise, the time between successive R waves (*RR interval*) corresponds to the time between ventricular depolarizations (and thus ventricular contractions), so the RR interval can be used to calculate the ventricular rate. Of course, in a normal heart, the atrial rate equals the ventricular rate.

Six Standardized Electrocardiographic Leads Are Used in Veterinary Medicine

Fig. 20.6 shows actual ECG records obtained from a normal dog. To obtain these recordings, electrodes were placed on the left forelimb, right forelimb, and left hind limb. Electrodes on these limbs are usually envisioned as forming a triangle around the heart (just as electrodes at points A, B, and C form a triangle around the heart in Fig. 20.3). The various ECG tracings in Fig. 20.6 were obtained by interconnecting these electrodes in standardized combinations prescribed by Willem Einthoven, inventor of the ECG. As shown in Fig. 20.6B, the voltage in the left forelimb compared with the right forelimb is called *lead I.* Note that lead I corresponds to the voltage measurements discussed with Fig. 20.4. The same pattern of distinct P, R, and T waves is evident in the lead I tracing in Fig. 20.6, as was seen in Fig. 20.4 (although the T wave happens to be negative in Fig. 20.6).

In accordance with Einthoven's convention, the connections for the three standard limb leads are depicted in Fig. 20.6 in the

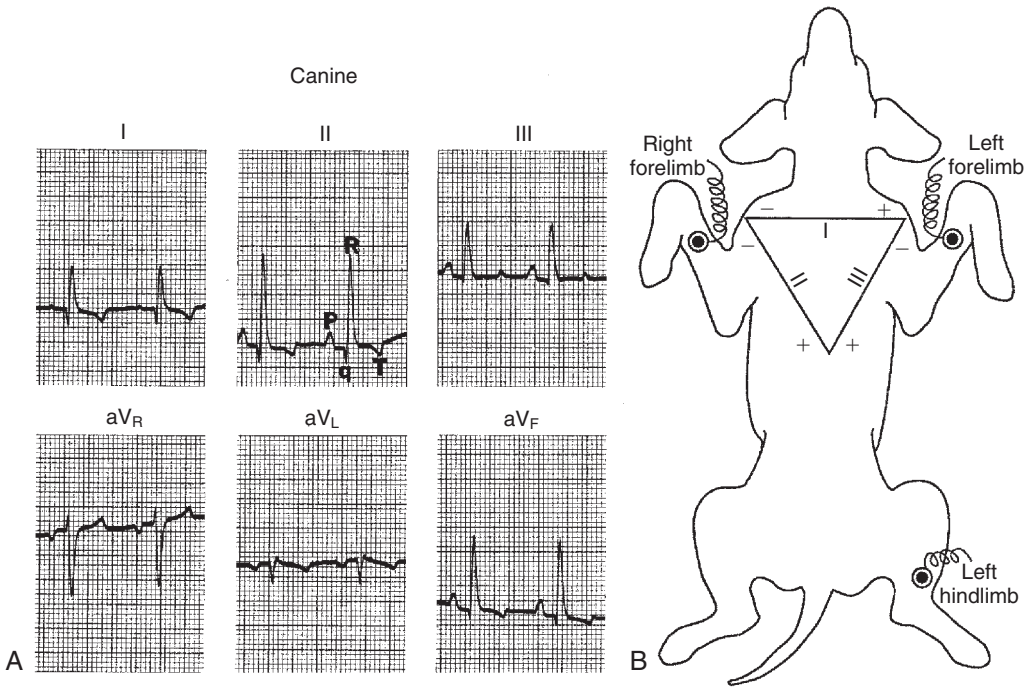


$$\text{Atrial rate (per minute)} = \frac{60{,}000 \text{ ms/min}}{\text{P-P Interval (in ms)}}$$

$$\text{Ventricular rate (per minute)} = \frac{60{,}000 \text{ ms/min}}{\text{R-R Interval (in ms)}}$$

• **Fig. 20.5** The time between various waves of the electrocardiogram corresponds to the timing of specific electrical events in the heart. See text for a complete description. The equations show how the atrial rate and the ventricular rate can be calculated from the PP and RR intervals, respectively. Of course, in a normally functioning heart, atrial rate = ventricular rate = heart rate.

Vet-ebooks.com



• **Fig. 20.6** (A) Six-lead electrocardiogram (ECG) from a normal dog. P, Q, R, and T waves (visible in all six leads) are labeled in lead II. There are no distinct S waves in these ECG recordings, and the T waves happen to be negative in leads I, II, aV_L, and aV_F. These are not abnormal signs. (B) Einthoven's triangle (superimposed on ventral view of dog) depicts the standard conventions for interconnecting the three limb electrodes to obtain lead I, lead II, and lead III ECGs. See text for additional explanation. (A, From Tilley LP. *Essentials of Canine and Feline Electrocardiography: Interpretation and Treatment*. 2nd ed. Philadelphia: Lea & Febiger; 1985.)

form of a triangle (*Einthoven's triangle*). The triangle indicates that, to make a lead I ECG, the voltage is recorded in the left forelimb (labeled the + *electrode*) compared with the right forelimb (called the − *electrode*). Similarly, the diagram indicates that *lead II* is the voltage measured in the left hind limb compared with the right forelimb, and *lead III* is defined as the voltage in the left hind limb compared with the left forelimb. It is important to remember that the + and − signs on Einthoven's triangle are simply notations about how to hook up the electrodes. They indicate, for example, that lead I is obtained by measuring the voltage in the left forelimb compared with the right forelimb (not vice versa). The + and − signs on the triangle do not necessarily correspond to the orientation of the dipoles created in the heart.

As illustrated in Fig. 20.6A, the major ECG events (P, R, and T waves) are normally evident whether one is looking at the recording from lead I, II, or III. These *standard limb leads* simply provide different angles for viewing the electrical dipoles created by the heart muscle as it depolarizes and repolarizes. Three additional electrical views are provided by the *augmented unipolar limb leads* (aV_R, aV_L, and aV_F). Lead aV_R measures the voltage from the right forelimb electrode compared with the average voltage from the other two limb electrodes. Similarly, aV_L and aV_F measure the voltages from the left forelimb and left hind limb electrodes respectively, compared with the average voltage from the other two electrodes.

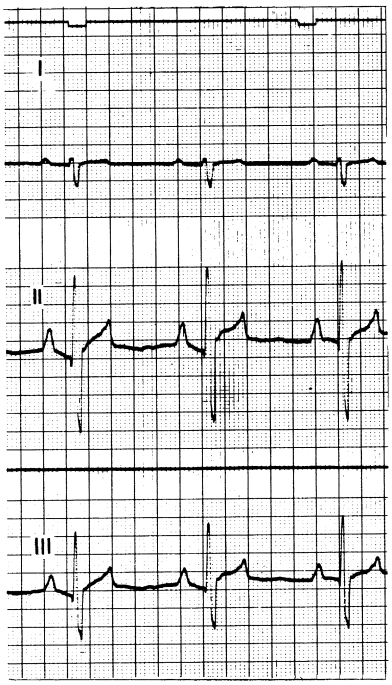
Leads I, II, and III are used routinely in veterinary electrocardiography of small animals. Recordings from the augmented unipolar limb leads (aV_L, aV_R, and aV_F) are often included as well. Six additional leads are sometimes recorded by placing ECG electrodes at six standardized sites on the thorax. These *precordial* (*chest*) *leads* are used much more often in human medicine than in veterinary medicine. They are helpful in evaluation of very specific cardiac electrical dysfunctions.

The standardized vertical calibration on an ECG is that 10 mm (two major divisions) equals 1 millivolt (mV). There are two standard chart speeds: 25 millimeters per second (mm/s), whereby five major divisions on the horizontal (time) axis equal 1 second; or 50 mm/s, whereby 10 major divisions on the horizontal axis equal 1 second. Using the faster chart speed (50 mm/s) helps to spread out the ECG events in an animal with a rapid heart rate (e.g., a cat). "Chart speed" is a convention derived from older, analog, paper-readout (strip chart) ECG machines. Although ECG is now more commonly captured and stored digitally, the chart speed convention is still used to set the resolution of the digital display. Furthermore, many of these digital units can produce a permanent paper printout of their data that looks just like the older strip chart.

Abnormal Voltages in the Electrocardiogram Are Indicative of Cardiac Structural or Electrical Abnormalities

The ECG in Fig. 20.7 was obtained from a dog with right ventricular hypertrophy. Note that the sequence of waves in the ECG appears to be normal; that is, each heartbeat begins with an upward-going P wave, which is followed by a QRS complex and a T wave (which happens to be positive in this dog). The atrial and ventricular rates are equal, at about 100 beats per minute (beats/min). An abnormality is evident, however, because the predominant polarity of the QRS complex recorded from lead I is negative. As mentioned

Vet-ebooks.com



• **Fig. 20.7** Electrocardiogram (ECG) from a dog with right ventricular hypertrophy. The chart speed is 50 mm/s; therefore 10 major grid divisions on the horizontal axis equal 1 second. One-second timing marks are visible as small, downward deflections at the very top. Both the PP and the RR intervals are 0.6 second, so both atrial and ventricular rates are 100 per minute. The salient abnormalities are (1) predominantly negative QRS complexes recorded from lead I and (2) large-amplitude, bidirectional QRS complexes recorded from leads II and III. (From Ettinger SJ. *Textbook of Veterinary Internal Medicine*. 3rd ed. Philadelphia: Saunders; 1989.)

earlier, the QRS complex is caused by ventricular depolarization, and its dominant feature is normally a large, positive R wave. As explained earlier, the R wave is normally positive as recorded from lead I, because the cardiac axis is normally angled to the left side of the thorax and because the left ventricular wall is much more massive than the right ventricular wall. Both these features have the effect of making the net dipole of ventricular depolarization point toward the left forelimb (as shown in Fig. 20.4C). Therefore reversal of this polarity suggests that the cardiac axis has shifted to the right, that the mass of the right ventricle has dramatically increased, or both. The abnormally high voltages of the QRS complex recorded from leads II and III are indicative of ventricular hypertrophy. The pronounced negative components in the QRS complexes recorded from leads II and III suggest that during part of ventricular depolarization, the predominant direction of depolarization is away from the left hind limb. This is consistent with a cardiac axis shifted to the right and a massive right ventricle. Substantial right ventricular hypertrophy is a common consequence of cardiac defects that increase the pressure that must be generated within the right ventricle during its contractions. Examples include pulmonic stenosis, patent ductus arteriosus, and ventricular septal defect (see Chapter 21).

Sometimes, ECG voltages are abnormally low. One common cause of low-voltage ECG waves is an accumulation of fluid within the pericardium. In a sense, this pericardial fluid creates a short circuit for the ionic currents that would ordinarily flow outward toward the body surface. Therefore voltages smaller than normal are created at the body surface.

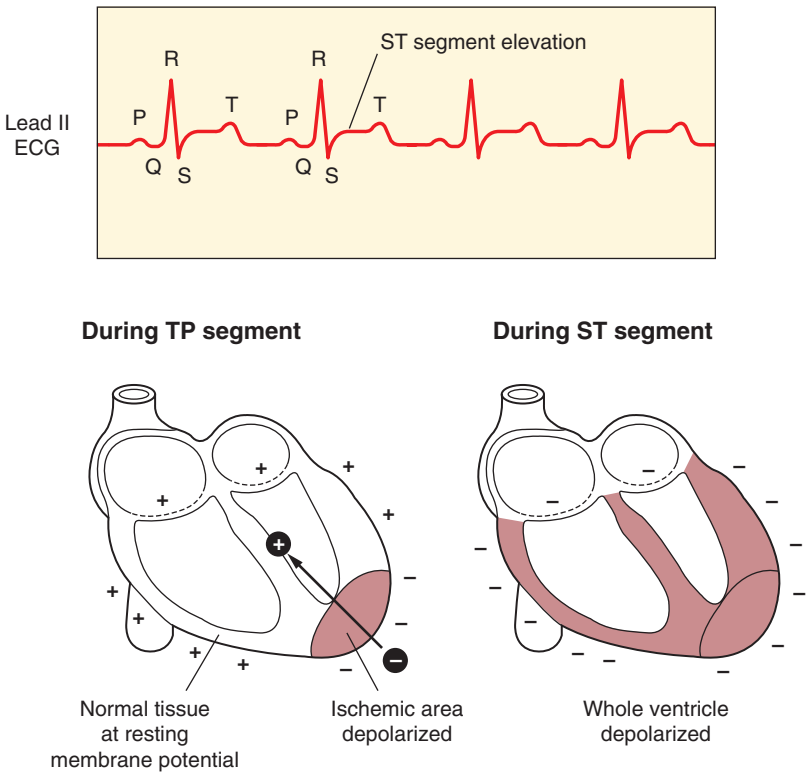
An upward or downward shift of the ST segment, compared with the rest of the ECG, is often indicative of an area of ischemic or infarcted ventricular muscle. Typically, ischemic or infarcted ventricular muscle cells cannot maintain a normal, negative resting membrane potential; these cells are always more or less depolarized. Therefore between ventricular contractions, when normal ventricular cells are at a normal resting membrane potential, a voltage difference exists between the normal and ischemic (or infarcted) ventricular cells. This voltage difference creates an electrical dipole between normal, resting ventricular muscle and ischemic (or infarcted) ventricular muscle. Fig. 20.8 (*bottom left*) shows the orientation of this dipole for the case of an ischemic area in the inferior (caudal) part of the ventricles. The dipole creates a negative voltage in lead II during ventricular rest (i.e., during the TP segment). When an action potential enters this ventricle, the normal ventricular tissue becomes depolarized, and a QRS complex is observed. The ischemic area cannot form action potentials; it simply remains depolarized. As a result, during the ST segment, the entire ventricle, normal and ischemic, is depolarized (Fig. 20.8, *bottom right*). During the ST segment, there is no voltage difference (no dipole) between the injured area and the normal area. With no dipole present, the ECG voltage during the ST segment is close to a true zero level. However, the ST segment is elevated in relation to the more negative voltage during the TP segment (ventricular rest). Thus *ST segment elevation* (which is actually "TP segment depression") is indicative of an ischemic or infarcted area in the inferior (caudal) part of the ventricle. Ischemia or infarction in the anterior (cranial) ventricular area would cause *ST segment depression.*

Making a diagnosis of a structural abnormality in the heart solely on the basis of abnormal ECG voltage is risky. Theoretically, if the structural and electrical properties of a particular heart are known in detail, the appearance of the ECG can be predicted with certainty. However, the reverse situation is not strictly true. Several different cardiac abnormalities may result in similar voltage abnormalities. Thus an abnormal ECG cannot be ascribed with certainty to a particular cardiac defect. That is why other clinical data (e.g., thoracic radiographs or echocardiographs) are used in conjunction with ECG observations in diagnosing specific structural abnormalities in the heart.

Electrical Dysfunctions in the Heart Cause Abnormal Patterns of Electrocardiogram Waves

Fig. 20.9 is an ECG from a dog with *premature ventricular contractions* (PVCs). This lead I recording begins with five normal beats (each QRS complex is preceded by a P wave and followed by a T wave). The P waves are evenly spaced with a PP interval of 0.5 second (so heart rate is 120 beats/min). After five normal beats, a large-voltage complex of abnormal shape occurs without a preceding P wave. This is indicative of a premature ventricular depolarization (atrial depolarization could not produce such large voltage deflections). The abnormal shape and long duration of the complex indicate that the premature depolarization did not spread through the ventricles via the normal, rapidly conducting bundle branches and Purkinje fibers. In other words, the ectopic site that originated the premature depolarization was not within the AV node, bundle, or bundle branches. The predominant voltage in the abnormal complex is positive in lead I, indicating that the premature ventricular depolarization propagated predominantly in a right-to-left direction through the dog's ventricles. The premature depolarization likely originated in nonspecialized ventricular tissue in the right side of the heart. Given that the depolarization proceeded through

Vet-ebooks.com



• **Fig. 20.8** Voltage recorded during the ST segment is elevated compared with the baseline (TP segment) in this lead II electrocardiogram (ECG) from a dog with an inferior (caudal) ventricular infarction. The drawings show why an ischemic or infarcted area of ventricle creates a net electrical dipole in the resting ventricle (during TP segment) but not in the depolarized ventricle (during ST segment).



• **Fig. 20.9** Lead I electrocardiogram (ECG) of a dog showing five normal beats (normal P-QRS-T pattern) followed by one premature ventricular beat. A sixth P wave would be expected at the time marked by the *arrow.* This P wave is obscured by the large voltages associated with the premature ventricular beat. Also, the refractory period associated with the premature beat prevented the sixth normal ventricular beat from occurring; this creates a long pause (called the *compensatory pause*) between the premature beat and the next regular beat. In this and the remaining ECG examples, chart speed is 50 mm/s (10 major grid divisions equal 1 second). (From Ettinger SJ. *Textbook of Veterinary Internal Medicine.* 3rd ed. Philadelphia: Saunders; 1989.)

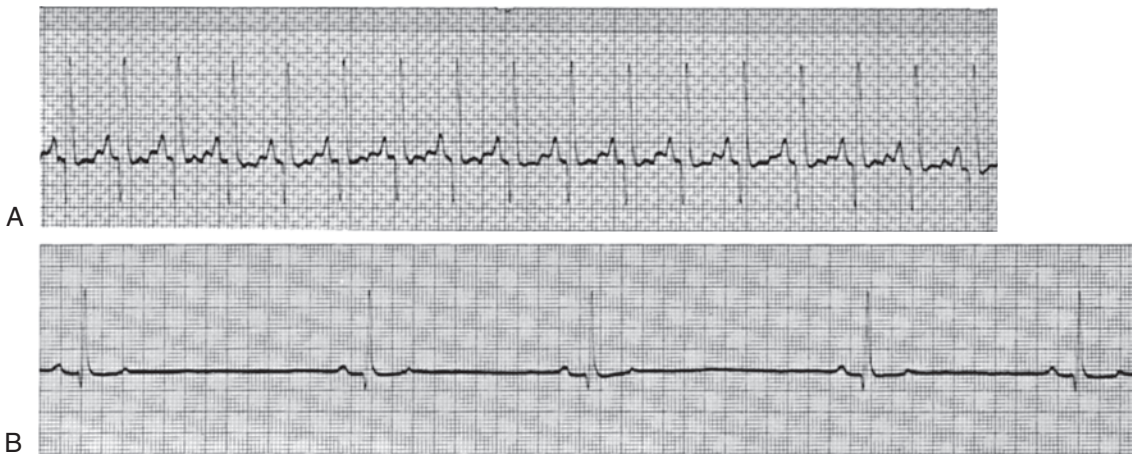
the ventricles with abnormal direction and speed, it is not surprising that the pattern of repolarization was also abnormal (causing the T wave to have an unusual amplitude).

If a premature ventricular depolarization does originate from an ectopic pacemaker within the AV node, bundle, or bundle branches, the pattern and timing of ventricular depolarization and of ventricular repolarization would be normal; that is, the QRS complex and the T wave of the premature beat would look like the normal QRS and T waves. The QRS-T sequence would simply occur earlier than expected and would not be preceded by a P wave.

Sometimes, premature contractions are initiated by ectopic pacemakers in the atria (*premature atrial contractions,* PACs). If an early atrial depolarization is conducted to the ventricles (i.e., if the AV node is not still refractory from the preceding atrial depolarization), the resulting ventricular depolarization and repolarization would follow normal ventricular pathways. Therefore the ECG would show an earlier-than-expected P wave, followed by QRS-T sequence of normal size and shape.

Fig. 20.10 shows additional examples of cardiac electrical dysfunctions, recorded from resting dogs. In the ECG in Fig. 20.10A, the R waves are evenly spaced and indicate a ventricular rate of 235 beats/min. This is fast for a resting dog. However, the pattern of ECG waves appears to be normal; each QRS complex is preceded by a clear, positive P wave and is followed by a positive T wave (which overlaps the next P wave). The most likely diagnosis

Vet-ebooks.com



• **Fig. 20.10** Sinus tachycardia (A) and sinus bradycardia (B) are evident in these otherwise-normal electrocardiograms (ECGs) from two resting dogs. Chart speed is 50 mm/s. (From Ettinger SJ. *Textbook of Veterinary Internal Medicine*. 3rd ed. Philadelphia: Saunders; 1989.)

is *sinus tachycardia* (rapid heart rate initiated by SA node pacemakers). Fig. 20.10B shows the opposite extreme. The pattern of ECG waves is normal, but the heart rate is only 55 beats/min. The diagnosis is *sinus bradycardia* (the SA node is the pacemaker, but its rate is abnormally slow).

The ECG provides an easy way to diagnose AV node block. The ECG in Fig. 20.11A looks normal, except that the PR intervals are abnormally long, which is indicative of abnormally slow conduction of the action potential through the AV node and AV bundle, and thus *first-degree AV node block.* In Fig. 20.11B, the P wave spacing indicates an atrial rate of 123 beats/min. Four of the P waves are followed by tall (but faintly visible) QRS complexes and large, negative T waves, but the other seven P waves are not followed by QRS-T sequences. Apparently some, but not all, atrial depolarizations are conducted through the AV node, which indicates a condition of *second-degree AV block.* The condition is not life threatening unless there are so many missed ventricular beats that cardiac output falls to dangerously low levels.

Fig. 20.11C shows *third-degree* (*complete*) *AV node block* (and, incidentally, ST segment depression). Two large QRS complexes are visible, each followed by a negative T wave. The RR interval is about 2.9 seconds, indicating that the ventricular rate is only 21 beats/min. The QRS complexes are not immediately preceded by P waves. Small, evenly spaced, positive P waves are present, indicating a constant atrial rate of 142 beats/min, but there is no synchronization between the P waves and the QRS complexes. Atrial action potentials are apparently being blocked at the AV node. The ventricles are beating slowly in response to an auxiliary (emergency) pacemaker further along in the AV node.

Fig. 20.12A shows an ECG record of a dog that is drifting in and out of ventricular tachycardia. The first five waves are abnormally shaped ventricular complexes, indicative of an ectopic ventricular pacemaker located outside the normal ventricular conduction system. P waves (if present) are being obscured by the large ventricular complexes. Then there are three normal-appearing P-QRS-T sequences, which suggests that a normal rhythm is being established. However, the ectopic ventricular pacemaker usurps control again, and ventricular tachycardia returns.

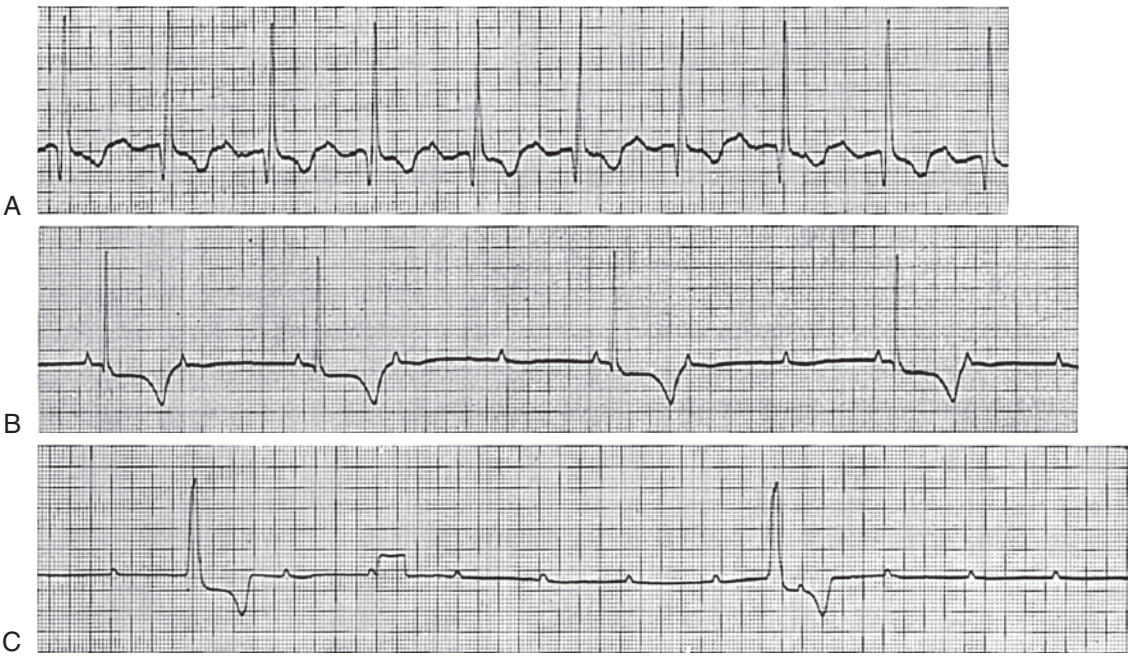
Ventricular tachycardia degenerates frequently into ventricular fibrillation. The ECG in Fig. 20.12B indicates ventricular fibrillation. The record shows fairly large, irregular voltage fluctuations with no discernible pattern. The atria may or may not be fibrillating; that is, regularly occurring P waves may be present but obscured by the random electrical activity in the ventricles. Fibrillating ventricles are just quivering and do not pump significant cardiac output, even if the atria continue to contract in a synchronized manner.

Atrial fibrillation, as with ventricular fibrillation, typically produces random voltage dipoles. However, because the atrial muscle mass is relatively small, the ECG voltages generated by atrial fibrillation are always much smaller than those seen in Fig. 20.12B. An ECG from an animal with atrial fibrillation would typically show normally shaped QRS-T sequences against a background of low-amplitude voltage fluctuations created by the fibrillating atria. In such a case, atrial tissues bombard the AV node with frequent but randomly spaced action potentials. Some of these action potentials are conducted to the ventricles, and others are blocked (because of the long refractory period of the AV node). The QRS-T sequences would typically have normal shape but irregular spacing in time.

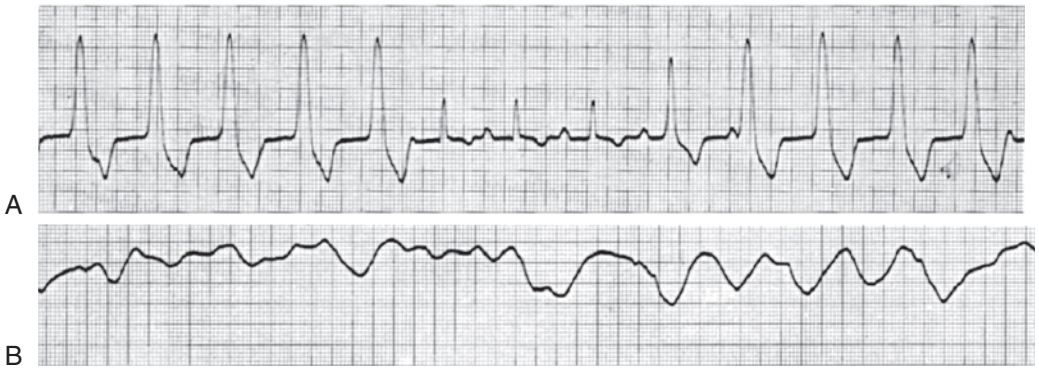
In Large Animals, There Is Considerable Variability in the Polarity and Amplitude of the Electrocardiogram Waves

The appearance of the normal ECG waves varies more, from animal to animal, among horses and cattle than among dogs and cats. For example, healthy cattle are likely to have QRS complexes (in any particular ECG lead) that are rather different in magnitude, duration, and shape between individuals. This variability arises from the less-consistent pathways followed by cardiac depolarizations in the atria and ventricles of large animals compared with small animals. As a consequence, the ECG is not very useful for detecting cardiac structural abnormalities (e.g., ventricular hypertrophy) in large animals. Nevertheless, there is consistency in the basic sequence of electrical events in the hearts of normal animals, whether large or small. Each normal heartbeat begins with a depolarization of the SA node, and the consequent sequence of events (depolarization of the atria, depolarization of the ventricles, and repolarization of the ventricles) produces waves of voltage that are evident on

Vet-ebooks.com



• **Fig. 20.11** (A) Example of first-degree atrioventricular (AV) node block (abnormally slow AV conduction). Each QRS complex is preceded by a positive P wave and followed by a negative T wave, which is normal. However, the PR interval is 0.2 second (normal for a dog is less than 0.14 second). (B) Example of second-degree AV node block (sporadic AV conduction). Each of the four QRS complexes in this recording is preceded by a positive P wave, and followed by a broad, negative T wave. Where P waves are followed by QRS-T complexes, the PR interval is normal. However, several P waves are not followed by a QRS-T complex; as a result, there are two or three atrial beats for every ventricular beat. (C) Example of third-degree (complete) AV node block. Regularly spaced P waves are evident (although two of them are obscured by the two large QRS-T complexes). The QRS-T complexes are not immediately preceded by P waves. ST segment depression is also evident, but this is irrelevant to the diagnosis of AV block. The rectangular deflection one-third of the way through the record is a voltage calibration signal (1 mV). Chart speed is 50 mm/s. (From Ettinger SJ. *Textbook of Veterinary Internal Medicine*. 3rd ed. Philadelphia: Saunders; 1989.)



• **Fig. 20.12** (A) Example of ventricular tachycardia, which reverts briefly to a sinus rhythm. The ventricular rate is about 165 beats/min. This pattern would be typical for a dog with an ectopic ventricular pacemaker functioning at almost the same rate as the sinoatrial (SA) node pacemaker; some ventricular beats would be initiated by the ectopic pacemaker, and others would be initiated in the normal way through the atrioventricular (AV) node. (B) Example of ventricular fibrillation. The random voltage fluctuations generated by the fibrillating ventricles would obscure any P waves that might be present, so it is not possible to determine whether the atria are beating normally or are also fibrillating. Chart speed is 50 mm/s. (From Ettinger SJ. *Textbook of Veterinary Internal Medicine*. 3rd ed. Philadelphia: Saunders; 1989.)

Vet-ebooks.com

an ECG. Therefore the ECG is very useful in large animals for detecting and characterizing cardiac arrhythmias. Standardizing the placement of electrodes for particular ECG leads is usually not necessary for this purpose. Any ECG lead or electrode placement that yields clearly discernable P waves, QRS complexes, and T waves will suffice.

Sophisticated techniques are widely used in the analysis of ECGs both in human medicine and in many veterinary clinics. The purpose in this chapter is to introduce only enough complexity to establish a conceptual model for thinking about the ECG and to illustrate the usefulness of that model in the clinical diagnosis of cardiac electrical dysfunctions.

CLINICAL CORRELATIONS

Dilative Cardiomyopathy With Paroxysmal Atrial Tachycardia

History

An owner brings his 5-year-old male Saint Bernard to you because of a distended abdomen, weakness, coughing, and difficulty breathing. The owner believes these signs developed gradually over several weeks; however, before the last few weeks, there were occasional episodes when the dog suddenly seemed weak and very listless. The episodes lasted from a few minutes to about an hour.

Clinical Examination

Palpation reveals that the dog has muscle wasting and marked ascites (fluid in the abdominal cavity). The jugular veins are distended. The arterial pulse is rapid and irregular; there are frequent pulse deficits ("missing" beats). Thoracic radiography reveals an enlarged heart and an accumulation of fluid near the lung hilus.

You record the dog's ECG for several minutes. You note that P waves usually occur at a rate of 160 to 170 per minute and that each P wave is followed by a QRS-T complex. However, the ECG also shows frequent episodes when there are 210 to 230 P waves per minute. During these episodes, most P waves are followed by QRS-T complexes, but others are not. As a result, the QRS-T complexes occur irregularly, with about 180 per minute.

Echocardiography reveals severe dilation of all four cardiac chambers, particularly the atria. Even though the ventricles are enlarged, the ventricular walls are thinner than normal, a condition called *eccentric hypertrophy.* Ventricular contractions are weak.

Comment

The ECG indicates that this dog has atrial tachycardia. The information presented does not establish whether the atrial pacemaker is located in the SA node or somewhere else in the atria. It is likely that one atrial pacemaker area is initiating depolarizations at a rate of 160 to 170 per minute and that another atrial area intermittently preempts the first pacemaker by initiating depolarizations at the more rapid rate of 210 to 230 per minute. When the atrial rate is 160 to 170 per minute, the AV node conducts every atrial action potential to the ventricles, so that the ventricles also contract 160 to 170 times per minute. However, when the atrial rate is 210 to 230 per minute, some of the atrial action potentials arrive at the AV node when the nodal cells are still refractory from the preceding action potential. These atrial action potentials are not conducted into the ventricles, which is why there are only about 180 ventricular contractions per minute. This is a case in which a second-degree AV node block, created by the relatively long refractory period of AV node cells, is beneficial, because it prevents the ventricles from beating too rapidly. The problem, when an arrhythmia triggers very frequent ventricular contractions, is that the time available between contractions becomes too short for adequate ventricular refilling. As ventricular rate increases, the volume of blood pumped with each beat (stroke volume) decreases, and so does cardiac output. At ventricular rates above 180 per minute, cardiac output could fall to such a low level that the dog would collapse.

This dog's primary problem is probably a chronic, progressive weakening of his heart muscle (*cardiomyopathy*). All the clinical signs, including atrial tachycardia, can be attributed to a primary cardiomyopathy. Dilative cardiomyopathy is common in giant-breed dogs, especially males, and often (as in this case) there is no discernible cause.

Even though the cause of the cardiomyopathy could not be determined from the evidence available in this case, the sequence of dysfunctions that resulted from the cardiomyopathy can be inferred with near-certainty. Ventricular weakness caused heart failure; the cardiac output fell below normal, especially during exercise. The dog's body attempted to compensate for the heart failure by increasing blood volume, which increased both venous and atrial pressures far above normal. The elevated atrial pressure had the beneficial effect of "supercharging" the ventricles with an extra volume of blood before each contraction, which partially returned *stroke volume* (the volume of blood pumped by a ventricle with each heartbeat) toward normal. However, the excessive volume and pressure of blood in the veins caused pulmonary edema (which led to coughing and difficulty breathing) and systemic edema (which led to fluid in the abdomen). Also, distention of the atria made the atrial cells more excitable electrically, which resulted in the formation of ectopic pacemakers and the onset of atrial tachycardia. The tachycardia limited the ventricular refilling time, causing further compromise in cardiac output. A vicious cycle began in which decreased cardiac output caused further venous congestion and atrial distention, which aggravated the arrhythmia, and so forth. The atrial tachycardia will likely progress to atrial fibrillation. The prognosis is poor without treatment.

This case of heart failure provides a good preview for the next several chapters, which deal in detail with the physiological mechanisms of cardiac and vascular control in both normal and heart failure states.

Treatment

A diuretic drug (e.g., furosemide) is administered to promote an increase in urine formation. The goal is to reduce the blood volume and venous and atrial pressures, thereby reducing the signs resulting from congestion and edema. Sometimes the paroxysmal atrial tachycardia resolves after diuretic-induced reductions in atrial size. If it does not, antiarrhythmic drugs (e.g., quinidine or lidocaine, and/or a cardiac glycoside such as digitalis) can be used to try to reduce the electrical excitability of atrial tissue.

Race Horse Fades During Workout

History

A trainer calls you about a 3-year-old thoroughbred stallion that they started training at their track 2 weeks ago. The colt starts strong but then appears to fade and sometimes stumbles toward the end of hard work. According to the previous trainer and owner, the horse had no prior health issues. He is not on any medications, and he is current on vaccinations and Coggins (a blood test for Equine Infectious Anemia).

Clinical Examination

When examined at rest, the colt's temperature and respiratory rate are normal. Pulse rate is 40 per minute, but both the strength of the pulse and the interval between pulses are quite variable. Upon cardiac auscultation, you hear an irregularly irregular heart rhythm. That is, there is no predictable pattern to the timing of heartbeats. When you simultaneously auscultate the heart sounds while palpating the arterial pulse, you notice that the pulses are weaker when heartbeats occur close together in time, whereas pulses are stronger when there is a longer interval between beats. In other respects, the physical examination is normal. You hear no cardiac murmurs (abnormal "extra" heart sounds, which would indicate defective heart valves). Palpation

Continued

Vet-ebooks.com



• **Fig. 20.13** Electrocardiogram (ECG) from a horse with atrial fibrillation. P waves are absent. The fibrillating atria create random, frequent, low-amplitude voltage fluctuations (several of these labeled *f*). QRS-T complexes (two of these are labeled) have normal morphology, but RR intervals are irregularly irregular. The line bar indicates 1 second. See text for additional explanation. (From Verheyen T, Decloedt A, De Clercq D, et al. Electrocardiography in horses—Part 2: how to read the equine ECG, *Vlaams Diergeneeskundig Tijdschrift*. 2010;79:337–344.)

CLINICAL CORRELATIONS—cont'd

of muscles and joints is within normal limits. Hoof tests are normal. You have someone walk, trot, and lunge the horse for you and you do not see any lameness or neurological issues. Lung sounds (auscultated with the aid of a rebreathing bag) are normal.

Comment

Based on the history of fading during exercise and the irregularly irregular heartbeat, atrial fibrillation is the most likely differential diagnosis. ECG is used to confirm atrial fibrillation (see Fig. 20.13). Distinct, well-spaced P waves are absent. Instead, fibrillating atria create random, high-frequency, low-amplitude voltage fluctuations on the ECG, which are called *f waves* (fibrillation waves). With atrial fibrillation, QRS-T complexes usually appear normal in shape. However, they occur irregularly for two reasons. First, the arrival of strong atrial depolarizations at the AV node occurs randomly in time. Second, atrial depolarizations arriving at the AV node during the node's refractory period are not propagated to the ventricles. Fibrillating atria do not contract synchronously or effectively, but simply quiver. The variable strength of the arterial pulse results from the variable time available for ventricular filling. The shorter filling time between closely spaced ventricular contractions leads to lower stroke volume on those beats and therefore a weaker arterial pulsation. The main reason why atrial fibrillation limits exercise performance is that the lack of well-timed, effective atrial contractions limits ventricular filling and therefore limits stroke volume and cardiac output.

Treatment

Based on this colt's age and lack of other medical issues, an underlying cardiac disease is unlikely. However, a cardiac ultrasound is recommended to confirm that there are no underlying structural or functional cardiac defects. Blood work should also be done to confirm no underlying plasma electrolyte abnormalities or other health issues. The goal of treatment is conversion to normal sinus rhythm. This is most often accomplished by administering quinidine. Quinidine is class IA antiarrhythmic drug that basically acts to block some of the fast, voltage-gated Na^+ channels. In effect, quinidine depresses myocardial excitability and slows conduction velocity. As a result, it prolongs refractive period and suppresses reentrant arrhythmias. Na^+ channel blockers are toxic, so a test dose of quinidine is given to the horse to make sure there are no significant side effects. The horse is then given another dose 2 hours later and monitored for conversion to normal cardiac rhythm. If conversion does not occur, up to four additional doses are given, while monitoring for side effects. If the horse converts, he will be rested for 48 to 72 hours, and then he can go back into light work.

Practice Questions

1. Which of the following intervals on an ECG most closely corresponds to the time required for propagation of a cardiac action potential through the AV node?
- **a.** RR interval
- **b.** PR interval
- **c.** ST interval
- **d.** PP interval
- **e.** QT interval

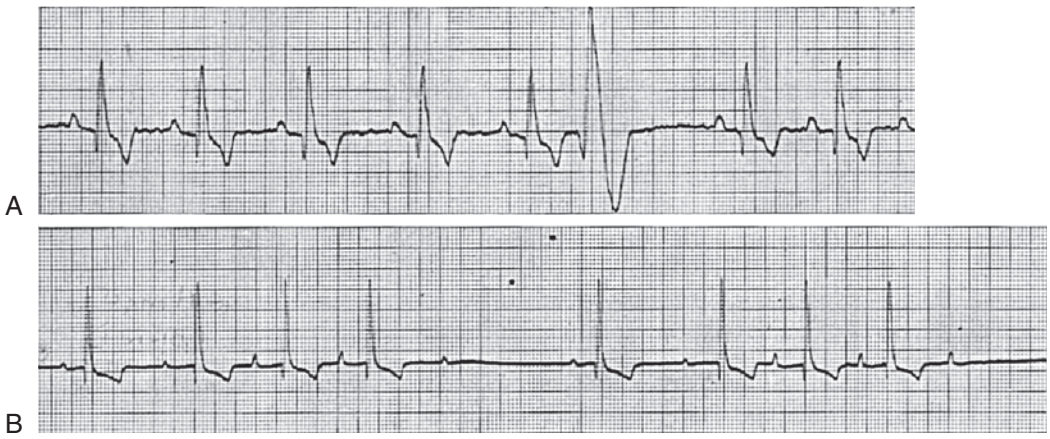
2. The T wave in a normal lead I ECG is:
- **a.** Always negative.
- **b.** Always positive if the R wave is positive.
- **c.** Also known as the *pacemaker potential.*
- **d.** Caused by the delay between atrial and ventricular depolarization.
- **e.** Caused by ventricular repolarization.

3. What would the lead I ECG look like if an ectopic pacemaker in the free wall of the left atrium subsumed the role of the SA node (i.e., "took over" the initiation of atrial action potentials)?
- **a.** The ECG would appear normal.
- **b.** The order of waves would be reversed (i.e. T-QRS-P, instead of P-QRS-T).
- **c.** The P waves and T waves would appear normal, but there would be no R waves.
- **d.** The P wave would be negative and the R wave would be positive.
- **e.** The P wave would be negative and the R wave would be negative.

4. In which of the following arrhythmias will the ECG characteristically show the same number of P waves and QRS complexes?
- **a.** Complete (third-degree) AV block
- **b.** First-degree AV block
- **c.** Ventricular tachycardia
- **d.** Atrial flutter
- **e.** All the above are correct.

5. The ECG in Fig. 20.14A indicates:
- **a.** Sinus bradycardia
- **b.** Right ventricular hypertrophy
- **c.** ST segment elevation
- **d.** Premature ventricular contraction
- **e.** Atrial fibrillation

Vet-ebooks.com



• **Fig. 20.14** Lead I electrocardiogram recordings from two dogs. A is the basis for Practice Question 5. B is the basis for Practice Questions 6 and 7. Chart speed is 50 mm/s. (From Ettinger SJ. *Textbook of Veterinary Internal Medicine*. 3rd ed. Philadelphia: Saunders; 1989.)

6. The ECG in Fig. 20.14B indicates:
- **a.** Second-degree AV block
- **b.** Third-degree AV block
- **c.** Sinus bradycardia
- **d.** Ventricular tachycardia
- **e.** ST segment elevation

7. Which of the following is the best estimate for the atrial rate in the dog whose ECG is depicted in Fig. 20.14B?
- **a.** 30/min
- **b.** 60/min
- **c.** 90/min
- **d.** 120/min
- **e.** 150/min

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.
Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. Elsevier; 2017.
Katz AM. *Physiology of the Heart*. 5th ed. Baltimore: Lippincott, Williams & Wilkins; 2010.
Marr C, Bowen M, eds. *Cardiology of the Horse*. 2nd ed. Philadelphia: Saunders; 2011.
Martin M. *Small Animal ECGs: An Introductory Guide*. 3rd ed. Wiley-Blackwell; 2015.
McMichael M, Fries R. *Life-Threatening Cardiac Emergencies for the Small Animal Practitioner (Rapid Reference)*. Wiley-Blackwell; 2016.
Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.
Reed S, Bayly W, Sellon D. *Equine Internal Medicine*. 4th ed. Elsevier; 2017.
Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2015.
Verheyen T, Decloedt A, De Clercq D, et al. Electrocardiography in horses – part 2: how to read the equine ECG. *Vlaams Diergeneeskd Tijdschr*. 2010;79:337–344.

Vet-ebooks.com

21
The Heart As a Pump

ROBERT B. STEPHENSON

KEY POINTS

1. Each heartbeat consists of ventricular systole and ventricular diastole.
2. Cardiac output equals heart rate multiplied by stroke volume.
3. Increases in end-diastolic ventricular volume cause increases in stroke volume.
4. End-diastolic ventricular volume is influenced by ventricular preload, ventricular compliance, and diastolic filling time.
5. Increases in ventricular contractility cause decreases in ventricular end-systolic volume.
6. Increasing the heart rate does not increase cardiac output substantially unless stroke volume is maintained.
7. Murmurs are abnormal heart sounds caused by turbulent flow through cardiac defects.
8. Some cardiac defects increase the heart's workload, which leads to cardiac hypertrophy.
9. The pathophysiological consequences of cardiac defects are direct results of the abnormal pressures, volumes, and workloads created in the cardiac chambers.

Each Heartbeat Consists of Ventricular Systole and Ventricular Diastole

The heart is actually two pumps (two ventricles) that work together, side by side, in a synchronized cycle, first relaxing and filling with blood and then contracting and ejecting some blood. In each *cardiac cycle* (heartbeat), the left ventricle takes in a volume of blood from the pulmonary veins and left atrium, then ejects it into the aorta. The right ventricle takes in a similar volume of blood from the systemic veins and right atrium, then ejects it into the pulmonary artery.

Fig. 21.1 shows the events of a single cardiac cycle. A normal electrocardiogram (ECG) tracing is presented at the top of the figure. Atrial contraction (*atrial systole*) is initiated by atrial depolarization, which is indicated by the P wave. Ventricular contraction (*ventricular systole*) is initiated by ventricular depolarization, which is indicated by the QRS complex. Blood is ejected from the ventricles during ventricular systole. Each ventricular systole is followed by ventricular relaxation (*ventricular diastole*), during which the ventricles refill with blood before the next ventricular systole. Note that ventricular diastole corresponds to the period between a T wave and the subsequent QRS complex (rightmost arrow and its continuation on the left, Fig. 21.1 *top*), when ventricular cells are at resting membrane potential.

The ventricles do not empty completely during systole. As shown in the graph of ventricular volume (see Fig. 21.1, *second from top*), each ventricle of a large dog contains about 60 mL of blood at the end of diastole. This is called *end-diastolic volume.* During systole, about 30 mL of this blood is ejected from each ventricle, but 30 mL remains. This remaining blood volume is called *end-systolic volume.* The volume of blood ejected from one ventricle in one beat is called *stroke volume,* expressed as follows:

$$\text{Stroke volume} = \text{end-diastolic volume} - \text{end-systolic volume}$$

The fraction of end-diastolic volume that is ejected during ventricular systole is called the *ejection fraction,* as follows:

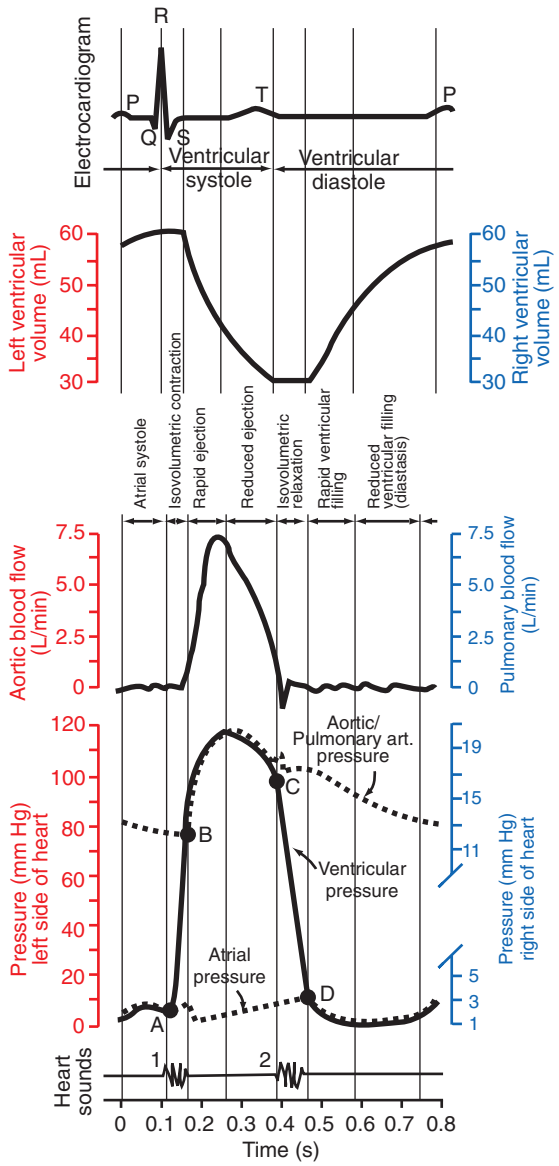
$$\text{Ejection fraction} = \frac{\text{Stroke volume}}{\text{End-diastolic volume}}$$

In the example depicted in Fig. 21.1, the ejection fraction is 50%. Values between 50% and 65% are typical for resting dogs.

As shown in Fig. 21.1, ventricular pressure is low (less than 10 mm Hg) during ventricular filling. Then, as ventricular systole begins (point *A* in the figure), the powerful contraction of the ventricular muscle causes the ventricular pressure to increase rapidly. This increase in ventricular pressure causes a momentary backflow of blood from the ventricles to the atria, which closes the *atrioventricular* (*AV*) *valves* (i.e., closes the *mitral valve* in the left heart and the *tricuspid valve* in the right heart). However, blood is not immediately ejected from the left ventricle into the aorta at the beginning of systole because the *aortic valve* remains closed until the left ventricular pressure exceeds the aortic pressure. Likewise, no blood is ejected from the right ventricle into the pulmonary artery until right ventricular pressure exceeds pulmonary arterial pressure. Therefore ventricular volume remains unchanged during this first phase of systole, which is aptly named *isovolumetric contraction.*

When left ventricular pressure does rise above aortic pressure (point *B* in Fig. 21.1), the aortic valve is pushed open, and there is a *rapid ejection* of blood into the aorta. Rapid ejection is followed by a phase of *reduced ejection* of blood, as both ventricular pressure and aortic pressure pass their peak (*systolic*) values and begin to decrease. During the period of reduced ejection, the ventricular pressure actually falls a bit below the aortic pressure, but ejection continues for a few moments because the blood flowing out of the ventricle is carried along by the momentum imparted to it during rapid ejection. As the ventricular pressure continues to decrease, ejection comes to an end. A momentary backflow of blood from the aorta into the left ventricle closes the aortic valve (point *C* in the figure). The closure of the aortic valve demarcates the end of ventricular systole and the beginning of ventricular diastole.

Vet-ebooks.com



• **Fig. 21.1** Events and terminology associated with one cardiac cycle (one heartbeat) in a normal dog at rest. The cycle is depicted as beginning with the occurrence of a P wave in the electrocardiogram and ending, 0.75 s later, when the subsequent P wave occurs. Vertical scales on the left side of the graphs (*red*) are for the left side of the heart. Vertical scales on the right side of the graphs (*blue*) are for the right side of the heart. In the graph of *ventricular pressure,* point *A* indicates closure of the mitral and tricuspid valves (the atrioventricular valves); point *B* indicates opening of the aortic and pulmonic valves; point *C* indicates closure of the aortic and pulmonic valves; and point *D* indicates opening of the mitral and tricuspid valves. See text for details.

During the first phase of ventricular diastole, the ventricular muscle relaxes, which causes left ventricular pressure to decline rapidly from a value near aortic pressure to a value near left atrial pressure. However, no filling of the left ventricle can occur because the mitral valve remains closed until left ventricular pressure drops below left atrial pressure. This first phase of ventricular diastole is called *isovolumetric relaxation* because there is neither filling nor emptying of the ventricle.

When left ventricular pressure does fall below left atrial pressure, the mitral valve is pushed open, as blood begins to flow from the left atrium into the left ventricle. First, there is a period of *rapid ventricular filling,* which is followed by a phase of *reduced ventricular filling* (*diastasis*). Diastasis persists until the sinoatrial node cells initiate an atrial action potential, which initiates another atrial contraction (*atrial systole*).

In a resting dog, as depicted in Fig. 21.1, ventricular volume nearly reaches its end-diastolic level even before the atria contract. Typically, 80% to 90% of ventricular filling occurs before atrial systole; atrial contraction simply "tops up" the almost-full ventricles. An important clinical consequence of this fact is that the ventricles in a resting animal can pump a nearly normal stroke volume even in the absence of properly timed atrial contractions (e.g., during atrial fibrillation). In contrast, during exercise, properly timed atrial contractions do make an important contribution to ventricular filling, because the rapid heart rate in exercise leaves a shorter time for diastolic filling. For this reason, animals with atrial fibrillation typically exhibit exercise intolerance. Atrial systole also becomes important for ventricular filling in patients with certain valve defects, such as narrowing of the mitral valve (*mitral stenosis*).

At the end of atrial systole, the atria begin to relax, and atrial pressure drops slightly. Then, as the ventricles begin to contract, there is a momentary backflow of blood from the left ventricle to the left atrium. The backflow closes the mitral valve, which marks the end of ventricular diastole and the beginning of another left ventricular systole.

By definition, the cardiac cycle is divided into ventricular systole and ventricular diastole. Closure of the mitral valve marks the beginning of ventricular systole. Closure of the aortic valve marks the beginning of ventricular diastole. Note that atrial systole takes place during ventricular diastole. Note also that the words *systole* and *diastole,* if used without a modifier, are presumed to refer to ventricular systole and ventricular diastole.

The preceding paragraphs discussed pressure changes in the left atrium, left ventricle, and aorta. However, corresponding events also take place on the right side of the heart; simply substitute "pulmonary artery" for "aorta," "pulmonic valve" for "aortic valve," and "tricuspid valve" for "mitral valve." As indicated in Fig. 21.1, the ventricular volumes are similar for the left and right sides, and so are the blood flow rates. The pressures, however, differ greatly on the two sides. Systolic (peak) pressure in the right ventricle and pulmonary artery is only about 20 mm Hg, whereas systolic pressure on the left side of the heart reaches 120 mm Hg. This explains why there are different scales on the pressure axes in Fig. 21.1 for the left and right sides of the heart.

Fig. 21.1 (*bottom*) shows the timing of the two major heart sounds. The *first heart sound* (S1) is associated with the closure of the AV valves (the mitral and tricuspid valves). The actual closure of the valves does not make this sound; indeed, the valve leaflets are so light and thin that their closing would be almost silent. However, there is a momentary backflow of blood from the ventricles to the atria at the beginning of ventricular systole. This backflow of blood is brought to a sudden stop against the closing valve leaflets. The resulting brief vibrations of the valve leaflets and ventricular wall creates the heart sound.

The *second heart sound* (S2) is associated with closure of the aortic valve on the left side of the heart and the pulmonic valve on the right side of the heart. S2 is usually briefer, sharper, and higher pitched than S1. Again, what makes the sound is not the valve leaflets closing, but rather the reverberation produced when the momentary backflow of blood into the ventricles is brought to a sudden stop by closure of the valves. The closures of the aortic and pulmonic valves are normally simultaneous. Under certain

Vet-ebooks.com

circumstances, however, the two valves close at slightly different times, and the second heart sound is heard as two distinct sounds in quick succession; this condition is called a *split second heart sound.*

AV valves close at the beginning of ventricular systole, and the aortic and pulmonic valves close at the end of ventricular systole. Therefore ventricular systole is sometimes defined as the part of the cardiac cycle between the first heart sound and the second heart sound.

Two additional heart sounds can commonly be heard in large animals (and occasionally in dogs). In early diastole, the rush of blood into the ventricles can create sufficient turbulence to vibrate the ventricular walls and create a *third heart sound* (S3). A *fourth heart sound* (S4), if audible, occurs at the very end of diastole, as atrial systole causes a sudden rush of blood into the ventricles.

Cardiac Output Equals Heart Rate Multiplied by Stroke Volume

The number of heartbeats per minute is called the *heart rate.* During each heartbeat, the left ventricle ejects a stroke volume of blood into the aorta, and the right ventricle ejects a stroke volume of blood into the pulmonary artery. Therefore *cardiac output* (the total volume of blood pumped by one ventricle in 1 minute) is determined as follows:

$$\text{Cardiac output} = \text{Stroke volume} \times \text{Heart rate}$$

This relationship emphasizes that cardiac output can be increased only if stroke volume increases, heart rate increases, or both increase. Therefore to understand how the body controls cardiac output, one must understand how the body controls stroke volume and heart rate. Fig. 21.2 summarizes the factors that affect stroke volume and heart rate. These factors are described in detail in the following sections.

Increases in End-Diastolic Ventricular Volume Cause Increases in Stroke Volume

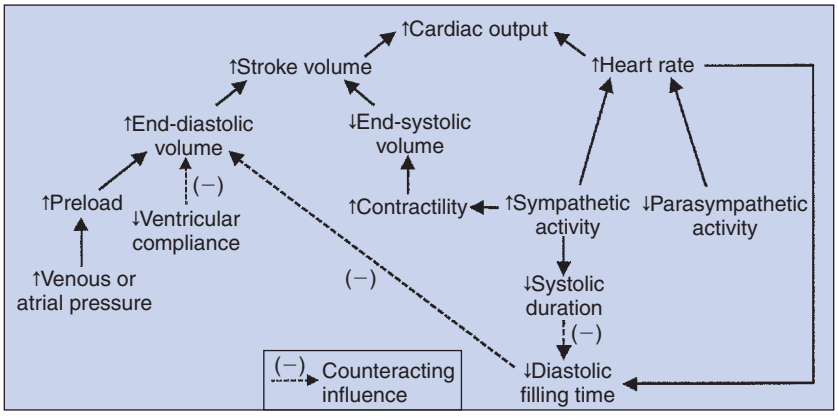
Stroke volume equals end-diastolic volume minus end-systolic volume. Therefore as shown in Fig. 21.2, stroke volume can increase only if end-diastolic volume increases (i.e., ventricles fill fuller during diastole) or if end-systolic volume decreases (i.e., ventricles empty more completely during systole), or both.

The effect of increasing end-diastolic ventricular volume (EDV) on stroke volume is plotted in Fig. 21.3A. The detailed physiological mechanisms underlying this relationship are complex. Basically, however, greater ventricular filling during diastole places the ventricular muscle fibers in a more favorable geometry for the ejection of blood during the next systole. Also, stretching the ventricular muscle fibers during diastole causes a greater amount of calcium (Ca^{2+}) to be released from the sarcoplasmic reticulum during the subsequent systolic contraction, and this enhances the force of contraction. Resting conditions in a normal animal are somewhere around the middle of this *ventricular function curve.* Therefore over a considerable range, increases or decreases in ventricular end-diastolic volume result in approximately proportional increases or decreases in stroke volume.

End-Diastolic Ventricular Volume Is Influenced by Ventricular Preload, Ventricular Compliance, and Diastolic Filling Time

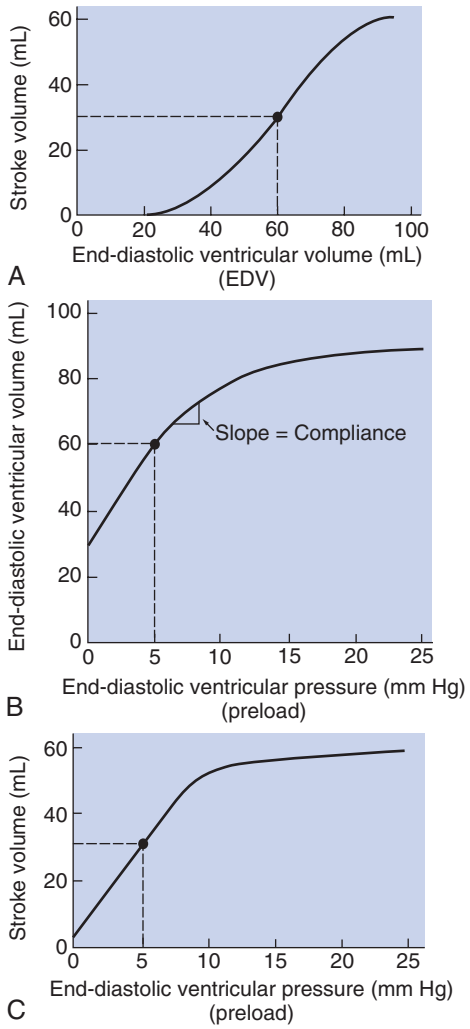
Ventricular preload is the pressure within a ventricle during diastolic filling. Because ventricular pressure changes throughout filling (see Fig. 21.1), the value of ventricular pressure at the end of diastole is usually accepted as a singular measure of preload. As depicted in Fig. 21.1, typical values of preload (*end-diastolic ventricular pressure*) are about 5 mm Hg for the left ventricle and about 3 mm Hg for the right ventricle. In a normal heart, ventricular pressure is essentially equal to atrial pressure during ventricular filling, because the AV valves are open widely. Also, because there are no valves between the veins and the atria, the atrial pressure is almost identical to the pressure within the nearby veins. Thus pulmonary venous pressure, left atrial pressure, and left ventricular end-diastolic pressure are all essentially equivalent measures of left ventricular preload. Similarly, right ventricular end-diastolic pressure, right atrial pressure, and vena caval pressure are all essentially equivalent measures of right ventricular preload. In the clinic, right ventricular preload is measured by introducing a catheter into a peripheral vein (e.g., the jugular vein) and advancing it into the cranial vena cava (precava) or right atrium. Such a catheter is called a *central venous catheter,* and the pressure measured at its tip is called *central venous pressure* (*CVP*). Left ventricular preload is more difficult to measure clinically because there is no easy way to place a catheter tip into the left atrium or pulmonary veins.

Fig. 21.3B shows that increases in ventricular preload bring about increases in end-diastolic ventricular volume. The graph



• **Fig. 21.2** Summary of the control of cardiac output. The relationships shown here are described in detail in the text.

Vet-ebooks.com



• **Fig. 21.3** (A) Increase in end-diastolic ventricular volume causes increased stroke volume. (B) Increase in end-diastolic ventricular pressure (preload) causes increased end-diastolic ventricular volume. The slope of this relationship corresponds to ventricular compliance (details in text). (C) Combines the relationships of *A* and *B* to show that an increase in ventricular preload causes increased stroke volume. An upper limit is reached in each relationship (*A* to *C*) primarily because, at high levels of end-diastolic ventricular volume, the ventricular walls become stretched to their elastic limit. The numerical data are for the left ventricle of a large dog. The *points* and *dashed lines* indicate normal values for the resting state.

depicts a left ventricle that has a natural volume of 30 mL in a relaxed, nonpressurized state (i.e., when the preload equals 0 mm Hg). Increases in preload cause increased ventricular filling. A preload of 5 mm Hg brings about the normal left ventricular end-diastolic volume of 60 mL. However, ventricular tissue reaches its elastic limit when the ventricular volume approaches 90 mL, so further increases in the preload do not cause much additional ventricular filling.

Because increases in ventricular preload cause increases in end-diastolic volume (see Fig. 21.3B), and increases in end-diastolic volume cause increases in stroke volume (see Fig. 21.3A), it follows that increases in preload cause increases in stroke volume (see Fig. 21.3C). Each of these relationships reaches an upper limit. Several factors are involved, but the main one (already mentioned) is that the ventricular walls become stretched to their elastic limit at high levels of end-diastolic ventricular volume. In a resting dog, the normal values of ventricular preload, end-diastolic volume, and stroke volume are about midway between their minimum and maximum values (see Fig. 21.3). Therefore a decrease below normal in preload will cause a decrease in both end-diastolic ventricular volume and stroke volume. This happens, for example, in response to hemorrhage (see Chapter 26).

The relationships among ventricular preload, end-diastolic volume, and stroke volume were first studied in detail by Ernest Henry Starling. The observation that changes in preload cause corresponding changes in end-diastolic ventricular volume and stroke volume is called *Starling's law of the heart.* The Starling mechanism is critical for moment-to-moment adjustments of stroke volume. For example, if the right ventricle begins, for any reason, to pump an increased stroke volume, the resulting additional pulmonary blood flow causes an increase in the pulmonary venous pressure, which increases left atrial pressure, which in turn increases left ventricular preload, which increases the filling of the left ventricle during diastole. The resulting increase in left ventricular end-diastolic volume leads to a greater stroke volume from the left ventricle. Thus an increase in right ventricular stroke volume quickly results in a corresponding increase in left ventricular stroke volume. The reverse is also true.

The sequence just described has a potential for developing into a vicious circle, with runaway increases in stroke volume. Other control mechanisms prevent this from happening, as discussed in Chapter 25. The point here is that the Starling mechanism keeps the stroke volumes of the left and right ventricles balanced. If this equality were not maintained (and one ventricle pumped more blood than the other for several minutes), a large part of the body's blood volume would accumulate either in the lungs or in the systemic circulation.

An alternate name for Starling's law of the heart is *heterometric autoregulation.* This name implies that the ventricles self-control (*autoregulate*) stroke volume in response to changes or differences (*hetero*) in initial volumes (*metric*). That is, *heterometric* refers to changes or differences in end-diastolic ventricular volumes between the two sides of the heart.

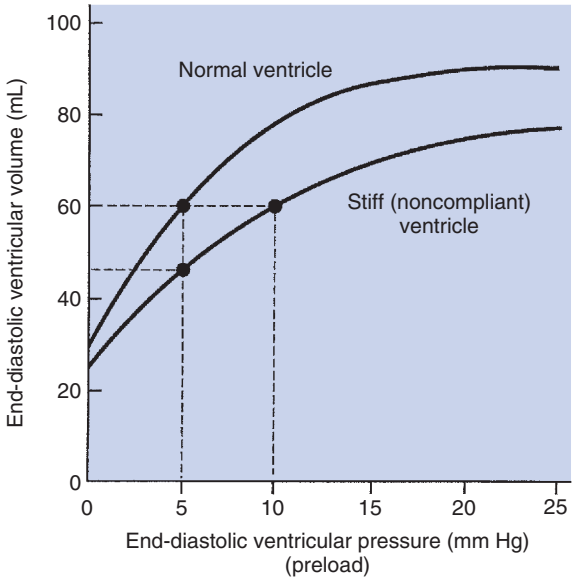
End-diastolic ventricular volume is determined not only by preload but also by *ventricular compliance.* Compliance is a measure of the ease with which the ventricular walls stretch to accommodate incoming blood during diastole. A compliant ventricle is one that yields easily to preload pressure and readily fills with blood during diastole. Compliance is more rigorously defined as follows:

$$\text{Compliance} = \text{Change in volume} \div \text{Change in pressure}$$

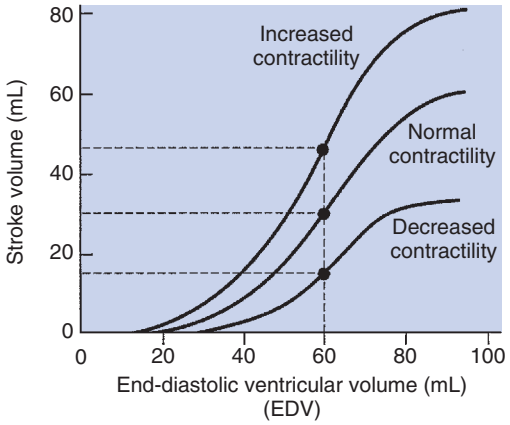
Ventricular compliance therefore corresponds to the slope of a ventricular volume versus pressure graph, such as the one shown in Fig. 21.3B. This figure illustrates that a normal ventricle is very compliant over the range of ventricular volumes up to and including the normal end-diastolic ventricular volume. Within this range, small changes in preload result in substantial changes in end-diastolic ventricular volume. At preloads higher than about 10 mm Hg, however, the ventricle becomes less compliant (stiffer). Inelastic connective tissue in the ventricular walls prevents ventricular volume from increasing above about 90 mL.

Myocardial ischemia, certain cardiac diseases, or advancing age can cause the ventricular walls to become stiffer (less compliant). Fig. 21.4 shows a comparison of volume versus pressure curves for a normal ventricle and for a noncompliant ventricle. In the noncompliant ventricle, there is a smaller increase in ventricular volume for any given increase in ventricular preload. As a

Vet-ebooks.com



• **Fig. 21.4** A stiff (noncompliant) ventricle requires a higher filling pressure (higher preload) to reach a normal degree of filling (normal end-diastolic ventricular volume). The noncompliant ventricle also reaches its elastic limit at a smaller ventricular volume.



• **Fig. 21.5** Increase in cardiac contractility is identifiable graphically as a leftward and upward shift of the ventricular function curve. Increase in contractility means that there will be a larger stroke volume for any given end-diastolic volume. Conversely, decrease in contractility (rightward and downward shift) means that there will be a smaller stroke volume for any given end-diastolic volume. With normal contractility and a normal end-diastolic volume of 60 mL, the end-systolic volume is 30 mL, and so the stroke volume is 30 mL (*middle dot*). Increased contractility (with no change in end-diastolic volume) results in decreased end-systolic volume. For example, if end-systolic volume is reduced to 15 mL, the stroke volume increases to 45 mL (*upper dot*).

consequence, a larger-than-normal preload is needed to obtain a normal end-diastolic ventricular volume and a normal stroke volume. An elevated preload necessitates elevated atrial and venous pressure, which can cause edema to develop in tissues "upstream" from the stiff ventricle. Thus stiffening of the left ventricle leads to elevated pressure in pulmonary veins and pulmonary edema; stiffening of the right ventricle leads to elevated pressure in the systemic veins and systemic edema (details in Chapters 23 and 26).

In addition to preload and compliance, a third factor that affects ventricular end-diastolic volume is the length of time available for ventricular filling during diastole. Heart rate is the main determinant of *diastolic filling time.* At a normal resting heart rate, there is ample time for ventricular filling during diastole; in fact, ventricular filling is almost complete even before atrial systole occurs (review Fig. 21.1). As heart rate increases, however, the duration of diastole decreases, and this tends to decrease diastolic ventricular volume. This limitation on ventricular filling would dramatically reduce stroke volume when heart rate is high if not for an additional, compensating influence brought about by the sympathetic nervous system, as discussed later.

Fig. 21.2 (*left side*) provides a useful summary of the preceding discussion. End-diastolic ventricular volume is determined by ventricular preload, ventricular compliance, and diastolic filling time. An elevated preload increases ventricular filling. Decreased ventricular compliance or decreased diastolic filling time can limit ventricular filling.

Increases in Ventricular Contractility Cause Decreases in Ventricular End-Systolic Volume

Contractility refers to the pumping ability of a ventricle. With increased contractility, there is a more complete emptying of the ventricle during systole and therefore a decreased end-systolic volume. An increase in contractility brings about an increase in stroke volume without requiring an increase in end-diastolic volume (see Fig. 21.2, *middle*). Fig. 21.5 shows graphically that an increase in contractility brings about an increased stroke volume for any given end-diastolic volume.

Sympathetic nerve activity increases ventricular contractility through the action of the neurotransmitter norepinephrine, which activates β-adrenergic receptors on ventricular muscle cells. As discussed in Chapter 19, activation of β-adrenergic receptors leads to an increased influx of extracellular Ca^{2+} into cardiac cells during an action potential (and to several other effects); the overall result is that cardiac contractions are stronger, quicker to develop, and shorter. Epinephrine and norepinephrine released from the adrenal medulla and circulating in the blood can likewise activate β-adrenergic receptors and increase contractility, as can β-adrenergic agonist drugs (e.g., isoproterenol). The cardiac glycosides (e.g., digitalis) are another class of drugs that increases cardiac contractility, again by increasing the cytosolic Ca^{2+} concentration during an action potential.

If cardiac contractility becomes depressed, there is less-than-normal ventricular emptying during systole. End-systolic volume increases and (as shown in Fig. 21.5) stroke volume decreases. A decrease in sympathetic activity causes a decrease in cardiac contractility, as do β-adrenergic antagonist drugs (e.g., propranolol or atenolol), which block the β-adrenergic receptors on cardiac muscle cells. As with β-adrenergic antagonists, calcium channel–blocking drugs also decrease cardiac contractility by making less Ca^{2+} available for the activation of the contractile proteins. Barbiturates, opioids, and some general anesthetics depress cardiac contractility as well; this must be kept in mind, particularly when administering such drugs to a patient who may already have compromised cardiac function. A decrease in cardiac contractility causes a decrease in stroke volume and therefore cardiac output. Consequently, the patient's blood pressure may fall to dangerously low levels.

A decreased cardiac contractility is the hallmark of the clinical condition called *myocardial failure,* which can result from coronary artery disease, myocardial ischemia, myocardial infarction, myocarditis, toxins, or electrolyte imbalances. Myocardial failure impairs

Vet-ebooks.com

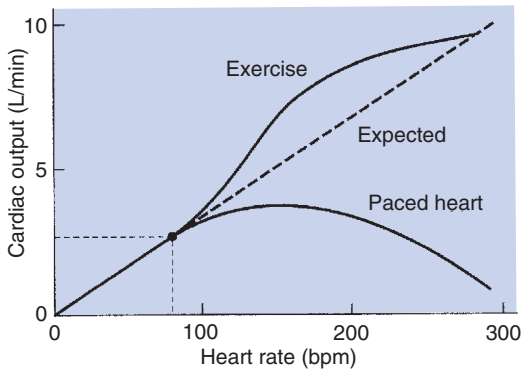
the pumping ability of one or both ventricles. Myocardial failure is one subtype of the more general dysfunction called *heart failure* (*pump failure*), which is the term that encompasses any dysfunction in the heart that compromises the heart's ability to supply the systemic organs with the blood flow they need to support their metabolism. Heart failure can also be caused by structural abnormalities in the heart (e.g., defective cardiac valves). Valvular heart disease will be described subsequently.

Although ventricular contractility is usually the predominant factor affecting ventricular end-systolic volume, the effect of arterial blood pressure must also be considered. A substantial increase in arterial blood pressure impairs ventricular ejection because the left ventricular pressure during systole must exceed aortic pressure before ejection of blood from the ventricle can occur. The pressure that a ventricle must generate in order to eject blood is called *cardiac afterload.* The higher the afterload, the more difficult it is for the ventricle to eject blood. If arterial pressure is excessively high, ventricular ejection is impaired, end-systolic volume increases, and stroke volume decreases. This effect is minor for a normal heart that is operating within the normal range of arterial pressure. However, high afterload can significantly limit stroke volume for a heart that is in failure.

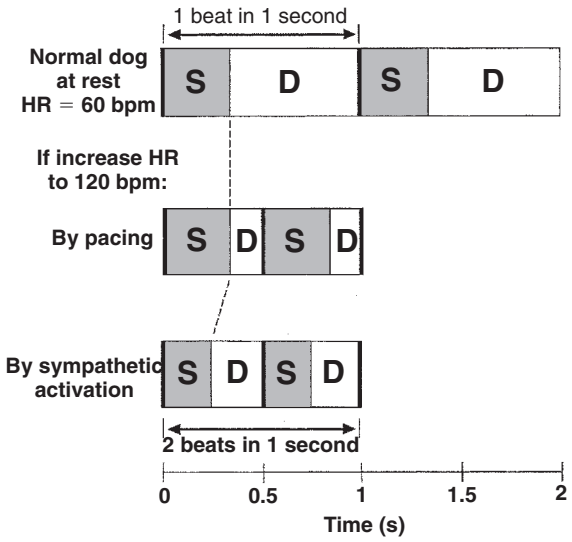
Increasing the Heart Rate Does Not Increase Cardiac Output Substantially Unless Stroke Volume Is Maintained

Because cardiac output is equal to stroke volume multiplied by heart rate, cardiac output might be expected to be proportional to heart rate; that is, doubling the heart rate would be expected to double cardiac output (Fig. 21.6, *dashed line*). However, if the heart rate is experimentally increased above its normal level with an artificial electrical pacemaker, cardiac output increases somewhat but not in proportion to the increase in heart rate. The reason, as mentioned earlier, is that increasing heart rate reduces diastolic filling time. The resulting reduction in end-diastolic volume reduces stroke volume, so cardiac output does not increase in proportion to heart rate (see Fig. 21.6, *lower solid line*). In fact, at heart rates greater than about 160 beats/min, stroke volume decreases so much that cardiac output actually declines with further increases in heart rate. This problem was encountered when early versions of artificial cardiac pacemakers malfunctioned in ways that caused high ventricular rates. Decreases in stroke volume at high heart rates are also encountered in certain cardiac arrhythmias. In *paroxysmal atrial tachycardia,* for example, a rapid heart rate is originated by an ectopic atrial pacemaker. The tachycardia occurs typically in bursts or paroxysms. During periods of very high heart rate, diastolic filling becomes so limited that cardiac output falls below normal. This causes the blood pressure to fall so low that the patient becomes lethargic and may even faint.

As previously noted, increasing heart rate by means of an artificial pacemaker does not cause the expected, proportional increase in cardiac output. In contrast, increases in heart rate in the course of normal daily activities are accompanied by substantial increases in cardiac output. An example is the increase in cardiac output that normally accompanies exercise. As shown in Fig. 21.6 (*upper solid line*), the actual increase in cardiac output during progressively more intense exercise is even greater than would be expected on the basis of the associated increase in heart rate. The reason that cardiac output increases so much during exercise is that stroke volume also increases. During exercise, sympathetic activation increases both heart rate and cardiac contractility, so the ventricles contract more times per minute and also empty more completely with each beat. In addition, sympathetic activation shortens the duration of systole, which helps to restore some of the diastolic filling time that would otherwise be lost at high heart rates. In summary, under sympathetic action, the heart not only contracts more frequently (increased rate) and more forcefully (increased contractility) but also contracts and relaxes more quickly (helping to preserve diastolic filling time). Fig. 21.7 illustrates how the



• **Fig. 21.6** Point indicates normal values of heart rate (80 beats/min) and cardiac output (2.4 L/min) for a dog at rest. *Dashed line* shows that cardiac output would increase in proportion to increases in heart rate (if stroke volume remains constant). However, if the heart is artificially paced to higher and higher rates, the observed increase in cardiac output is less than expected because stroke volume decreases (*lower solid line*). In contrast, when a dog increases its own heart rate through sympathetic activation (e.g., during exercise), cardiac output increases even more than expected because stroke volume increases (*upper solid line*).



• **Fig. 21.7** How shortening of systole (by sympathetic activation) helps to preserve diastolic filling time. *Top,* In this example, a large dog is resting very quietly with a heart rate (*HR*) of 60 beats per minute (*bpm*), so each beat occupies 1 second. Systole lasts about 1/3 second, which leaves 2/3 second (plenty of time) for diastolic filling. *Middle,* If HR is increased to 120 bpm by an artificial pacemaker, each beat occupies only 1/2 second. If the duration of systole is unchanged, diastolic duration (filling time) is greatly reduced. *Bottom,* if HR is increased to 120 bpm by sympathetic activation, systole becomes shorter, which restores part of the lost diastolic filling time. Thus sympathetic activation is said to help preserve diastolic filling time.

Vet-ebooks.com

TABLE 21.1 Typical Cardiac Changes During Vigorous Exercise in a Large Dog

Measurement	Rest	Exercise
Ventricular end-diastolic volume (mL)	60	55
Ventricular end-systolic volume (mL)	30	15
Stroke volume (mL)	30	40
Ejection fraction (%)	50	73
Heart rate (beats/min)	80	240
Cardiac output (L/min)	2.4	9.6

TABLE 21.2 Cardiac Valve Defects and Resulting Murmurs

	Nature of Defect	
Site of Defect	Incompetence (Insufficiency)—Allows Regurgitation	Stenosis (Narrow Valve Opening)—Creates Restriction
Atrioventricular valves	Systolic murmur	Diastolic murmur
Aortic or pulmonic valves	Diastolic murmur	Systolic murmur

shortening of systole helps to preserve diastolic filling time. Overall, sympathetic activation (especially when coupled with a decrease in parasympathetic activity) can dramatically increase cardiac output (Table 21.1).

It is useful at this point to review the control of cardiac output, as summarized in Fig. 21.2. Cardiac output is determined by stroke volume and heart rate. Stroke volume is determined by end-diastolic volume and end-systolic volume. End-diastolic volume depends on preload, ventricular compliance, and diastolic filling time. End-systolic volume depends on contractility and, to a lesser extent, on afterload (not shown in Fig. 21.2). Sympathetic activation increases contractility. Heart rate is increased by sympathetic activation and/or parasympathetic suppression. Increased heart rate acts directly to increase cardiac output, but it also decreases diastolic filling time, which reduces end-diastolic volume and compromises the increased cardiac output. Sympathetic activation shortens systolic duration, which helps to preserve diastolic filling time. The overall effects of sympathetic activation on the heart are made evident by noting the dramatic increase in cardiac output that occurs when a normal dog transitions from rest to vigorous exercise (see Table 21.1).

Murmurs Are Abnormal Heart Sounds Caused by Turbulent Flow Through Cardiac Defects

Cardiac murmurs are abnormal heart sounds, and they often indicate the presence of cardiac abnormalities. Some murmurs are exaggerations of normal heart sounds; others are additional ("extra") heart sounds. Murmurs are caused by turbulent blood flow. The underlying physical principle is that *laminar, smooth,* or *streamlined flow* of blood or other fluid is quiet, whereas *turbulent flow* is noisy. An analogy is that a river does not make any sound as it flows smoothly through a broad, relatively flat channel. If the same river enters a channel that is restricted or drops steeply, a rapid or cataract forms. The flow becomes turbulent, and the turbulent flow makes noise.

The flow of blood through the heart and blood vessels is normally smooth, and therefore quiet, during the majority of the cardiac cycle. A moment of turbulent flow normally occurs at the beginning of ventricular contraction, on closure of the AV valves. A second moment of turbulent flow occurs at the end of ventricular systole, when the aortic and pulmonic valves close. The moments of turbulence and vibration associated with valve closure create the first and second heart sounds (S1 and S2) as discussed previously and as illustrated in Fig. 21.1. On occasion (particularly in large animals), normal third and fourth heart sounds are faintly audible with the stethoscope, during rapid ventricular filling (S3) or during atrial systole (S4). In comparison with S3 and S4, clinically important murmurs are usually louder and persist through a greater portion of the cardiac cycle. Some murmurs are even louder than the normal first and second heart sounds.

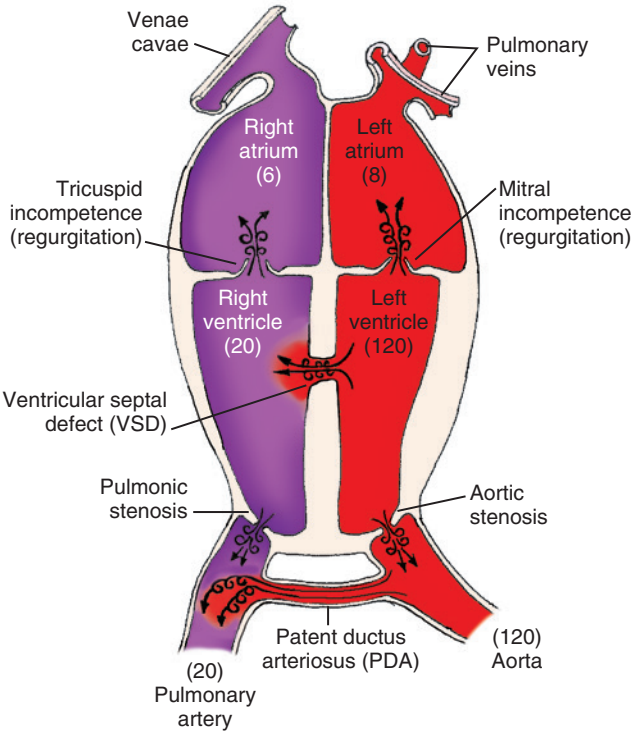
Table 21.2 lists cardiac valve defects that cause additional instances of turbulent flow and therefore murmurs. The table also indicates the timing of the murmurs in relation to the cardiac cycle. *Systolic murmurs* occur during ventricular systole; *diastolic murmurs* occur during ventricular diastole. *Continuous murmurs* occur throughout both systole and diastole. The timing of each murmur is easy to understand if two basic principles are kept in mind: murmurs are caused by turbulent blood flow, and blood flows in response to pressure differences. In other words, turbulent (noisy) flow through a cardiac defect occurs only if there is a substantial pressure difference from one side of the defect to the other.

Fig. 21.8 indicates how these principles can be used to account for systolic murmurs. The numbers in the figure indicate the maximum pressures that normally exist in each cardiac chamber during ventricular systole. Note, for example, that the pressure in the left ventricle is normally much higher than the pressure in the left atrium during ventricular systole. The mitral valve is normally closed during ventricular systole, so no blood flows backward from the ventricle to the left atrium. If the mitral valve fails to close completely during ventricular systole, the large pressure difference between the left ventricle and the left atrium causes a rapid, backward flow of blood through the partially closed valve. This turbulent backflow creates a systolic murmur. A mitral valve that fails to close completely is said to be *insufficient* or *incompetent.* The backflow across the valve is called *regurgitation.* Mitral regurgitation is present in about 8% of dogs over 5 years of age.

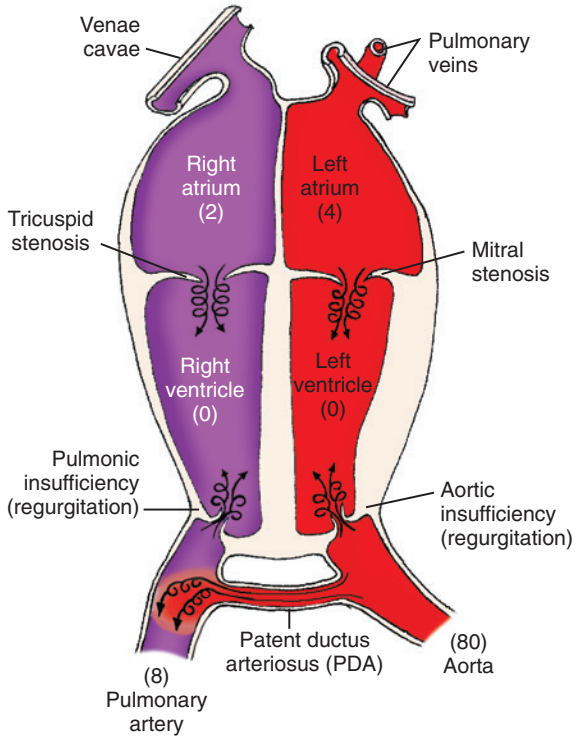
A *ventricular septal defect* (VSD) is a hole or cleft in the interventricular septum. Blood flows through a VSD from the left ventricle to the right ventricle during ventricular systole because systolic pressure is much higher in the left ventricle than in the right ventricle. Typically, the flow of blood through a VSD is turbulent, and a systolic murmur is created.

Systolic turbulence is also created if the aortic valve does not open widely enough. Blood ejected from the ventricle accelerates to a high velocity as it squeezes through the restricted aortic opening, and turbulence occurs. A valve that fails to open widely is called *stenotic;* the defect of *aortic stenosis* produces a systolic murmur. Likewise, *pulmonic stenosis* causes a systolic murmur. Aortic and pulmonic stenosis are common congenital defects in dogs.

Vet-ebooks.com



• **Fig. 21.8** Schematic view of the heart showing cardiac defects that cause systolic murmurs. The *numbers in parentheses* indicate normal maximum pressures (mm Hg) during ventricular systole. The *swirled arrows* indicate the sites of turbulent (noisy) flow. See text for details.



• **Fig. 21.9** Cardiac defects that cause diastolic murmurs. The *numbers in parentheses* indicate normal minimum pressures (mm Hg) during ventricular diastole. The *swirled arrows* indicate the sites of turbulent (noisy) flow. See text for details.

Patent ductus arteriosus (*PDA*) is persistence after birth of the opening between the aorta and the pulmonary artery (see Chapter 51). A PDA characteristically produces a murmur that persists throughout the cardiac cycle (*continuous murmur*). Aortic pressure is higher than pulmonary arterial pressure during both systole and diastole. Therefore some blood flows through the PDA, from aorta to pulmonary artery, during both systole and diastole, and this flow is typically turbulent. The resulting, continuous murmur is sometimes called a *machinery murmur* because it sounds like the rumble of machinery. PDA is common in young dogs, especially females.

The site on the thorax where a particular murmur can most clearly be heard (*auscultated*) is indicative of the particular location and type of defect that causes the murmur. For example, the murmur of PDA is characteristically heard best over the left heart base. Occasionally, the turbulence caused by a cardiac defect will be so extreme as to cause a palpable thoracic vibration (*thrill*).

Animals sometimes have abnormal openings between peripheral arteries and peripheral veins. These *arteriovenous fistulae* carry turbulent blood flow during both systole and diastole and therefore create continuous murmurs. The murmur of an arteriovenous fistula is most audible at the body surface close to the point of the fistula.

Fig. 21.9 shows defects that characteristically produce turbulent flow, and therefore murmurs, during diastole. For example, a normal mitral valve opens widely during ventricular diastole, which creates a low-resistance pathway for blood to flow smoothly from the left atrium into the left ventricle. However, if the mitral valve fails to open widely (*mitral stenosis*), ventricular filling must occur through a stenotic (narrow) valve. This creates turbulent flow and a diastolic murmur. Mitral stenosis is a common murmur among humans who have developed calcification of the mitral valve as a result of rheumatic heart disease.

During diastole, the pressure in the aorta is much higher than the pressure in the left ventricle. However, the aortic valve is closed during diastole, which normally prevents backflow of blood from the aorta into the left ventricle. If the aortic valve does not close tightly during diastole, a turbulent flow of blood regurgitates from the aorta into the left ventricle. Therefore *aortic regurgitation* produces a diastolic murmur. The defect is called *aortic incompetence* or *aortic insufficiency.* Aortic regurgitation is common in horses but not in dogs.

Diastolic murmurs can also be produced by defects on the right side of the heart. Pulmonic regurgitation or tricuspid stenosis would produce a diastolic murmur, although these abnormalities are uncommon, at least as congenital defects. However, a heavy infestation of heartworms in the right heart can create a functional stenosis at the tricuspid valve and a diastolic murmur.

Cardiac murmurs themselves are not harmful. They are clinically important, however, because the defects that cause the murmurs also have pathophysiological consequences. Cardiac defects typically lead to one or more of the following consequences: (1) abnormally high or low blood flow to a region of the body, (2) abnormally high or low blood pressure in a region of the body, and (3) *cardiac hypertrophy* (enlargement of cardiac muscle).

It is not difficult to understand why cardiac defects lead to abnormal blood flows or abnormal blood pressures. For example, in the presence of a VSD, the right ventricle receives blood from both the right atrium and the left ventricle, which leads to an abnormally high blood flow through the pulmonary circulation. In the presence of aortic stenosis, the left ventricle must generate

Vet-ebooks.com

an abnormally high systolic pressure to eject blood through the narrow valve opening. In an animal with mitral stenosis, blood dams up (and excessive pressure builds up) in the left atrium and pulmonary veins. This elevated pressure has the beneficial effect of helping to force blood flow through the stenotic mitral valve and into the left ventricle. However, a common adverse consequence is the development of pulmonary edema.

It is more difficult to understand why some cardiac defects lead to cardiac hypertrophy. The underlying principle is that some cardiac defects increase the workload of one or both ventricles, and an increase in the workload of cardiac muscle leads to hypertrophy. Developing this concept more fully requires an understanding of cardiac energetics, as described next.

Some Cardiac Defects Increase the Heart's Workload, Which Leads to Cardiac Hypertrophy

Cardiac defects often compromise the heart's ability to supply the systemic organs with the blood flow they need to support their metabolism. Compensating for such a *heart failure* or *pump failure* frequently requires one or both ventricles to pump more blood than normal or to pump blood at a higher pressure than normal. These adaptations increase the workload of the heart. A persistent increase in cardiac workload leads, over several weeks, to cardiac hypertrophy. A ventricle that must pump more blood volume than normal will develop some hypertrophy, whereas a ventricle that must pump blood at a higher pressure than normal develops a huge hypertrophy. This observation is the basis for the clinical aphorism, "Pressure work is harder for the heart [i.e., causes more hypertrophy] than volume work." To understand the physiological reason of this difference, we must delve into cardiac muscle energetics. To get started, it is useful to consider the analogous case of skeletal muscle hypertrophy in response to increased workload (physical conditioning).

A skeletal muscle does work by exerting a force while shortening. The useful mechanical work (*external work*) done by a skeletal muscle is equal to the force developed by the contracting muscle, multiplied by the distance moved during one contraction, multiplied by the number of contractions (that is, work equals force multiplied by distance). Therefore the external work done by a skeletal muscle can be increased by increasing the forcefulness of contraction, the distance moved, or the number of contractions. In weightlifting conditioning, the emphasis is on performing a few very forceful contractions of skeletal muscle. In contrast, conditioning that involves repetitive, low-force contractions of skeletal muscle (e.g., running, swimming) emphasizes primarily the distance and duration components of skeletal muscle work. Both "weight work" and "distance work" lead to skeletal muscle hypertrophy. However, a common observation is that weight work causes substantially more hypertrophy than does distance work. The basis for this difference is that weight work involves the generation of huge amounts of *internal work* (*wasted work*), which appears as heat. This large expenditure of energy on internal work greatly increases the *total work* (external work plus internal work) being done during weight lifting compared with distance running. It is the increase in total work of muscle, not just the external work, that is the primary stimulus for hypertrophy.

The heart does work by pumping blood. The useful mechanical work (external work) done by any pump is equal to the pressure generated by the pump, multiplied by the volume of fluid that is pumped in one pump stroke, multiplied by the number of pump strokes. Therefore the external work done by the left ventricle in 1 minute is equal to the pressure generated, multiplied by the stroke volume, multiplied by the heart rate. As shown in the following equation, the pressure generated by the left ventricle can be approximated by the average (mean) pressure in the aorta:

$$\textit{Minute work of left ventricle} = \text{Mean aortic pressure} \times \text{Stroke volume} \times \text{Heart rate}$$

The external work done by the ventricle in one cardiac cycle is called the *stroke work,* as follows:

$$\textit{Stroke work of left ventricle} = \text{Mean aortic pressure} \times \text{Stroke volume}$$

(The work of the right ventricle can be calculated in a similar way, but using mean pulmonary artery pressure.)

In accordance with the analogy to skeletal muscle conditioning, the mean aortic pressure is analogous to the force developed by the contracting skeletal muscle; the stroke volume is analogous to the distance moved during one contraction; and the heart rate is analogous to the number of contractions. Obviously, the external work done by the left ventricle could be increased by increasing the pressure that the left ventricle generates during systolic ejection, by increasing the stroke volume, or by increasing the heart rate. For example, a 50% increase in ventricular external work could result from a 50% increase in the left ventricular systolic pressure, a 50% increase in the left ventricular stroke volume, or a 50% increase in the heart rate. Any of these changes would lead, over a period of weeks, to left ventricular hypertrophy. However, an increase in the ventricular systolic pressure causes a much more pronounced hypertrophy than does an increase in the stroke volume or heart rate. The basis for this difference is that, to initiate the ejection of blood during each ventricular systole, the ventricular muscle must first contract strongly enough to raise ventricular pressure from its low, end-diastolic level up to a level that exceeds aortic pressure (see transition from *point A* to *point B* in Fig. 21.1). This phase of *isovolumetric contraction* requires the expenditure of considerable muscle energy. However, because there is no ejection of blood during this phase, the energy takes the form of internal work (wasted work, which appears as heat). The external (pumping) work of the ventricle begins only when ventricular pressure exceeds aortic pressure. Therefore any circumstance that elevates the ventricular pressure required to initiate systolic ejection will substantially increase the ventricle's *total work* (external work plus internal work). It is the total work of the cardiac muscle, not just the external work, that is the primary stimulus for hypertrophy.

Under normal resting conditions, about 85% of the metabolic energy consumed by the heart appears as heat (internal work), and only 15% appears as external work. A physicist would say that the heart has a "thermodynamic efficiency" of about 15%. However, the heart's "thermodynamic efficiency" depends on the type of work being done by the ventricles. The heart becomes less efficient when the external work is increased by increasing the systolic pressure. Conversely, the heart becomes more efficient when the external work is increased by an increase in the volume of blood pumped.

The dominant role of pressure in determining total ventricular energy consumption is evident from a comparison of the work done by the left and right ventricles. The stroke volume and heart rate are equivalent for the left and right ventricles, but the systolic pressure generated is about five times higher in the left ventricle than in the right (mean aortic pressure is about five times higher than mean pulmonary artery pressure). Therefore the external work

Vet-ebooks.com

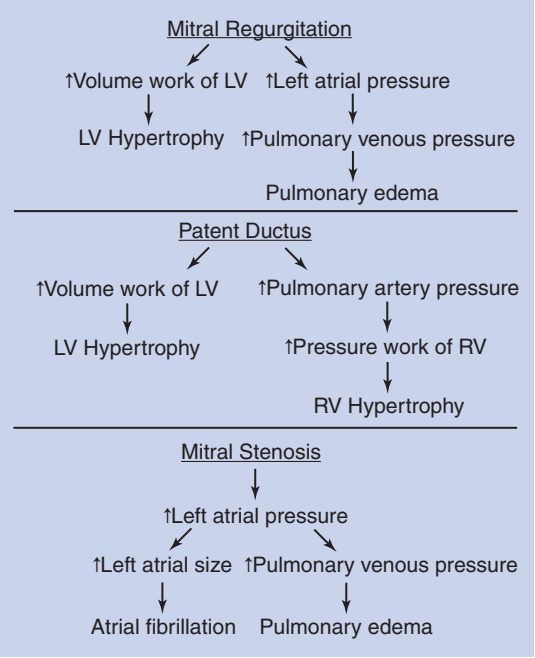
done by the left ventricle is approximately five times greater than the external work done by the right ventricle. However, the total metabolic energy consumption (total work) of the left ventricle is much more than five times greater than the energy consumption of the right ventricle, because the extra external work performed by the left ventricle is in the form of greater pressure. That is, the internal (wasted) work of the left ventricle is hugely greater than the internal (wasted) work of the right ventricle. Therefore almost all the metabolic energy required by the heart is utilized by the left ventricle; almost all the oxygen consumed by the heart is consumed by the left ventricle; and almost all the coronary blood flow is delivered to the left ventricular muscle. Because of the high amount of pressure work done by the left ventricle compared with the right ventricle, the left ventricle normally develops much heavier and thicker muscle walls than the right ventricle.

A clinical observation from human medicine provides a further illustration of how an increase in the ventricular pressure work leads to ventricular hypertrophy. Many adult humans have hypertension. In most of these patients, cardiac output is normal. Their arterial blood pressure is elevated because of an increased resistance to blood flow in the systemic arterioles. An elevated left ventricular pressure is required to force the cardiac output through these constricted systemic arterioles. The increased pressure work done by the left ventricle in hypertensive patients results in a striking left ventricular hypertrophy.

Up to a point, ventricular hypertrophy is an appropriate and beneficial adaptation to an increased workload imposed on the ventricular muscle. However, excessive hypertrophy is deleterious for three reasons. First, enlargement of the ventricular muscle restricts the opening of the aortic valve (or pulmonic valve, in the case of right ventricular hypertrophy). A vicious cycle develops. Ventricular hypertrophy leads to aortic or pulmonic stenosis, which necessitates that the ventricle generates an even greater systolic pressure to eject blood, which leads to more ventricular hypertrophy, and so on. A second complication of excessive hypertrophy is that the coronary circulation may be unable to provide enough blood flow to meet the increased metabolic demand of the massive ventricular muscle, particularly during exercise. Inadequate coronary blood flow is especially likely if the coronary vessels have become constricted because of coronary artery disease (atherosclerosis). As a result, patients with ventricular hypertrophy and coronary artery disease are at high risk for cardiac ischemia, myocardial infarction, ventricular arrhythmias, and sudden death, especially during periods of exercise. This explains why the all-too-common combination of hypertension and coronary artery disease is such a serious problem in human medicine. Fortunately, coronary artery disease is rare in most animals. The third complication of cardiac hypertrophy is that the cellular growth factors that mediate the hypertrophy also predispose the cardiac muscle to apoptosis (programmed cell death).

The Pathophysiological Consequences of Cardiac Defects Are Direct Results of the Abnormal Pressures, Volumes, and Workloads Created in the Cardiac Chambers

Fig. 21.10 summarizes the consequences associated with some common cardiac defects. First, consider *mitral regurgitation.* With each contraction of the left ventricle, a normal volume of blood is ejected into the aorta, and an additional volume of blood is ejected backward (through the regurgitant valve) into the left atrium. As a result, there is an increase in the volume work performed by



• **Fig. 21.10** Pathophysiological consequences of several common cardiac defects. *LV,* Left ventricle; *RV,* right ventricle.

the left ventricle. Therefore mild-to-moderate left ventricular hypertrophy develops. Also, in a heart with mitral regurgitation, the left atrium becomes distended, and left atrial pressure increases, as does pulmonary venous pressure. Elevated pressure in the pulmonary blood vessels forces water and electrolytes out of the bloodstream and into the pulmonary interstitial spaces, causing *pulmonary edema,* which can lead to respiratory dysfunction (see Chapter 50).

The consequences of mitral regurgitation are usually more noticeable during exercise than during rest. One reason is that, despite the regurgitation, the left ventricle can usually adapt enough, through hypertrophy and an increase in heart rate, to maintain a normal cardiac output into the aorta, and therefore into the systemic circulation, at rest. Also, despite some pulmonary edema, the oxygenation of the blood is sufficient to meet the animal's needs during rest. During exercise, however, the output of the left ventricle into the systemic circulation must increase severalfold to supply adequate blood to exercising skeletal muscle. Also, the delivery of oxygen into the blood must increase severalfold. Despite the hypertrophy, the left ventricle may not be able to deliver adequate blood flow to the systemic circulation during exercise if mitral regurgitation is serious. Also, pulmonary edema may prevent delivery of enough oxygen into the blood to support the metabolism of the exercising animal.

Consider next the abnormalities associated with *aortic stenosis* (not shown in Fig. 21.10). To eject a normal volume of blood with each beat through a stenotic aortic valve, the left ventricle must develop an abnormally high systolic pressure. This increases the pressure work of the left ventricle, which leads to a marked left ventricular hypertrophy. The hypertrophy has the desirable effect of increasing the contractility of the left ventricular muscle so that it can generate the increased pressure required to maintain normal cardiac output. As hypertrophy progresses, however, the ventricular muscle begins to impinge on the aortic outflow pathway, which further hampers the ability of the ventricle to eject

Vet-ebooks.com

blood. In a sense, the hypertrophic ventricular muscle "gets in its own way" or becomes muscle bound. The resulting limitation in aortic outflow is much more likely to be a problem during exercise than at rest. A patient with aortic stenosis may be able to function normally at rest but characteristically exhibits exercise intolerance.

PDA is a defect that typically results in both left and right ventricular hypertrophy (Fig. 21.10). In a typical patient with a PDA, the left ventricle pumps a near-normal volume of blood per minute to the systemic circulation and pumps two to three times that volume of blood per minute through the PDA. As a result, the volume work done by the left ventricle greatly exceeds normal amounts, which leads to left ventricular hypertrophy. The blood flowing through the PDA enters the pulmonary artery, and thus pulmonary arterial pressure exceeds normal levels. This in turn increases the pressure work that must be done by the right ventricle. The right ventricle receives a near-normal volume of blood back from the systemic circulation each minute, and the right ventricle must generate an elevated systolic pressure to eject this blood into the high-pressure pulmonary artery. The increased pressure work for the right ventricle is a powerful stimulus for hypertrophy, and pronounced right ventricular hypertrophy develops.

As a patient with PDA grows, exercise intolerance becomes evident. Despite hypertrophy, the left ventricle cannot supply both the increased blood flow needed by growing, exercising skeletal muscles, and also the blood that flows through the PDA. In patients with PDA, the pulmonary artery and the pulmonary blood vessels must carry not only the blood that is pumped by the right ventricle (as in a normal animal) but also the blood that is pumped through the PDA. In a severe case, pulmonary blood flow can be more than four times greater than normal. The resulting increases in pulmonary vascular pressure can lead to pulmonary edema. Surgical repair of a PDA in a young animal leads to a rapid reversal of all these cardiovascular and pulmonary abnormalities.

An understanding of the preceding examples should make it easy to predict the pathological consequences of a VSD. These consequences include increased volume work of the left ventricle, moderate left ventricular hypertrophy, increased volume and pressure work of the right ventricle, pronounced right ventricular hypertrophy, increased blood flow through the lungs, possible pulmonary edema, and probable exercise intolerance. It should also be clear why pulmonic stenosis leads to increased pressure work for the right ventricle and pronounced right ventricular hypertrophy (see Clinical Correlations).

Fig. 21.10 also summarizes the pathological consequences associated with the diastolic murmur of *mitral stenosis.* The left atrial pressure must exceed normal levels in order to force a normal volume of blood through the stenotic mitral valve and into the left ventricle during each ventricular diastole. The elevated left atrial pressure distends the left atrium. There may be some hypertrophy of the atrial muscle. One problem is that atrial action potentials tend to become discoordinated in a distended atrium, and atrial fibrillation is a common consequence. Also, the increase in the left atrial pressure causes blood to back up and accumulate in the pulmonary blood vessels, so pulmonary edema is likely. It might seem that the backup of blood in the pulmonary vessels would eventually also increase the pressure in the pulmonary artery and thereby increase the pressure work of the right ventricle. In other words, one might predict that mitral stenosis would lead to right ventricular hypertrophy. This prediction is logical, but in practice, animals with greatly elevated left atrial pressures usually die of the effects of pulmonary edema before right ventricular pressures have had a chance to become high enough to induce right ventricular hypertrophy. Therefore mitral stenosis does not generally lead to hypertrophy of either ventricle.

The defect of *aortic regurgitation* leads to left ventricular hypertrophy. With each systole, the left ventricle must eject an abnormally large volume of blood into the aorta. Of this, a normal volume of blood goes on into the systemic circuit; the rest is simply regurgitated back from the aorta into the left ventricle during diastole. Thus the volume work of the left ventricle is increased to above-normal levels, and left ventricular pressures may rise as well. Both these factors stimulate left ventricular hypertrophy. In severe cases of aortic regurgitation, diastolic ventricular pressure becomes elevated (because during diastole the left ventricle receives blood from both the left atrium and the aorta). This leads to increases in left atrial pressure, and pulmonary edema may develop.

Consideration of the abnormalities associated with cardiac defects is important for two reasons. First, these defects and their consequences are often encountered in veterinary medicine. Second, this discussion illustrates how the clinical signs and consequences of disease states can be understood and predicted in a rational way, on the basis of an understanding of basic principles of cardiac physiology.

CLINICAL CORRELATIONS

Pulmonic Stenosis

History

A 6-month-old female schnauzer is referred to your clinic because of a heart murmur that was detected during a routine health care visit. The puppy is fairly active but is slightly smaller than her female littermates. She also tires more quickly than her littermates when they play together.

Clinical Examination

All physical parameters are normal except for a systolic heart murmur that can be heard best over the left third to fourth intercostal space. Femoral pulses are normal, and the jugular veins are not distended. Electrocardiography reveals that the dog is in normal sinus rhythm with a heart rate of 118 beats/min. The PR interval is normal. However, the major QRS deflection is negative in leads I and aV_F. Also, deep S waves are noted in leads II and III, and the QRS complexes are slightly prolonged as a result of the wide S waves. Thoracic radiographs show right ventricular enlargement; the right border of the cardiac silhouette is more rounded, and closer to the right thoracic wall, than normal.

A catheter is inserted into the jugular vein, and the following pressures are measured as the catheter is advanced through the right side of the heart and into the pulmonary artery: central venous pressure (mean right atrial pressure), 8 mm Hg (normal, 3 mm Hg); right ventricular systolic pressure, 122 mm Hg (normal, 20 mm Hg); and pulmonary artery systolic pressure, 16 mm Hg (normal, 20 mm Hg).

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

The jugular catheter is withdrawn until the catheter tip is in the right ventricle. Additional radiographs are then taken while a radiopaque dye is injected through the catheter. These radiographs reveal that the right ventricular outflow tract is narrowed just below the pulmonic valve and that the pulmonic valve does not open widely during ventricular systole.

Comment

The young age of this dog and the absence of other signs of illness suggest that the murmur results from a congenital cardiac abnormality. Murmurs are graded on a scale of I through VI, with VI being the most severe. This dog's murmur is graded IV. A systolic murmur can result from aortic or pulmonic stenosis, mitral or tricuspid regurgitation, or a ventricular septal defect (see Fig. 21.8). On the basis of the location from which this murmur can be heard best, aortic or pulmonic stenosis is the most likely cause. All the additional clinical evidence supports a diagnosis of pulmonic stenosis.

The ECG indicates that the sinoatrial node is acting as the pacemaker and that the AV node is conducting each atrial action potential into the ventricles. However, the abnormalities observed in the polarities and shapes of the QRS complex are indicative of right ventricular hypertrophy, and the radiographs corroborate this finding. Pulmonic stenosis leads to right ventricular hypertrophy, because the right ventricle must generate much higher pressures than normal during systole in order to eject blood through the narrow outflow tract.

Normally, the pulmonic valve opens widely during systole, and the ventricular systolic pressure closely matches the pulmonary artery systolic pressure. In this dog, there is a difference of 106 mm Hg between right ventricular systolic pressure and the systolic pressure in the pulmonary artery just beyond the pulmonic valve. This difference indicates a severe pulmonic obstruction. The degree of obstruction can be visualized on the radiographs taken during dye injection.

Right ventricular hypertrophy is one of two adaptive responses that help this dog maintain a near-normal right ventricular stroke volume, despite the pulmonic stenosis. The other adaptive response is that the mean right atrial pressure is higher than normal (8 vs. 3 mm Hg). The right atrial pressure is elevated because blood backs up or accumulates in areas upstream from the stenosis (i.e., in the right ventricle, right atrium, and systemic veins). The elevated atrial pressure is adaptive because it increases the right ventricular preload, which increases the end-diastolic volume, which (according to Starling's law of the heart) helps keep the right ventricular stroke volume at a normal level, despite the stenosis. The right atrial pressure is not quite high enough in this dog to cause systemic edema or abdominal ascites (see Chapter 23). However, both these signs are sometimes seen in dogs with severe pulmonic stenosis because excessively elevated right atrial pressure leads to marked backup of blood and increased pressure within the systemic capillaries.

The combined effects of right ventricular hypertrophy and elevated right ventricular preload allow this dog's heart to pump a near-normal stroke volume during rest. However, the pulmonic obstruction limits the increase in stroke volume that can occur during exercise. The resulting limitation in cardiac output accounts for this dog's lack of stamina during exercise. Over a prolonged period, such a limitation in cardiac output will likely stunt growth.

Treatment

Theoretically, the best treatment for pulmonic stenosis is to remove the obstruction surgically. A valve dilator can be used or an artificial conduit can be installed across the stenotic valve. Although seriously affected dogs require such interventional treatment, dogs with mild-to-moderate pulmonic stenosis can lead sedentary lives without any treatment.

Some evidence indicates that the adverse effects of pulmonic stenosis can be minimized by the administration of β-adrenergic antagonists (e.g., propranolol) or calcium channel blockers (e.g., verapamil). Although the mechanism and efficacy of these drugs remain unclear, there is speculation that these drugs are beneficial because they limit ventricular contractility, which limits the work of the heart. Because an increase in cardiac work is the stimulus for hypertrophy, a drug that limits the increase in work also limits the hypertrophy. Although moderate hypertrophy can be adaptive (as explained earlier), excessive hypertrophy is detrimental for two reasons. First, the enlarged ventricular muscle can crowd the pulmonic outflow tract, worsening the stenosis. Second, the coronary circulation may be unable to deliver the increased amounts of blood flow required by the massive ventricular muscle.

Older Horse With Exercise Intolerance

History

A 22-year-old thoroughbred mare is presented for exercise intolerance. The owner uses her for trail riding and some low-level eventing. The mare has had some mild arthritis during her career, but in the last 2 to 3 months she has seemed reluctant to work, takes longer to recover after rides, and acts lethargic. Vaccinations and deworming are current.

Clinical Examination

The mare appears to be slightly underweight. She is responsive but quiet (more quiet than normal, according to the owner). Her temperature is normal; pulse and respiration are slightly increased. Her mucous membranes are darker pink than normal (suggesting reduced blood flow), but capillary refill time is not abnormally long. She has a grade IV systolic murmur on the left side, most consistent with mitral regurgitation. Her lungs are normal on auscultation. No other abnormal findings are noted on physical examination. The mare is lunged for several minutes and reauscultated. No additional abnormalities are detected, except the mare's heart rate and respiratory rate seem to take longer than normal to return to their resting levels. A blood sample is taken for analysis.

Comment

Results of the complete blood count (CBC) and serum chemistry are within normal limits. Echocardiography reveals mitral regurgitation associated with fibrotic thickening of the mitral valve. The *chordae tendineae* are intact. There is some dilation (eccentric hypertrophy) of the left ventricle but not of the left atrium.

Mitral valve thickening and insufficiency often develop with age, and mitral regurgitation is likely limiting this mare's left ventricular performance. The resulting tendency for inefficient pumping of blood into the systemic circulation can account for the decreased perfusion of the mucous membranes at rest and for the exercise intolerance and listlessness noticed by the owner. With each systolic contraction, the left ventricle is pumping blood both forward, into the aorta, and backward, through the leaky mitral valve and into the left atrium. The mild left ventricular hypertrophy and dilation are likely adaptive responses to this increased volume work. Animals with more severe mitral regurgitation also have left atrial dilation, associated with a much poorer prognosis than if there is no dilation or only left ventricular dilation.

Treatment

No medical treatment is indicated at this time. However, the owner needs to decrease the work by the mare. The mare should have only light, nonstressful activity. A follow-up examination is recommended in 3 to 6 months to assess the rate of progression of the mitral valve disease. If marked progression is noted at that time, the mare should be retired.

Vet-ebooks.com

Practice Questions

1. During a normal cardiac cycle, which of the following events happens soonest after the first heart sound is heard?
- **a.** Atrial contraction begins.
- **b.** The mitral valve opens.
- **c.** Atrial depolarization begins.
- **d.** Ventricular ejection begins.
- **e.** The aortic valve closes.

2. In the normal cardiac cycle:
- **a.** The P wave in the electrocardiogram coincides with the beginning of ventricular ejection.
- **b.** The second heart sound coincides with the beginning of isovolumetric relaxation.
- **c.** Left ventricular pressure reaches its highest level just as the aortic valve closes.
- **d.** Aortic pressure reaches its highest level at the beginning of ventricular systole.
- **e.** The mitral valve is open throughout ventricular diastole.

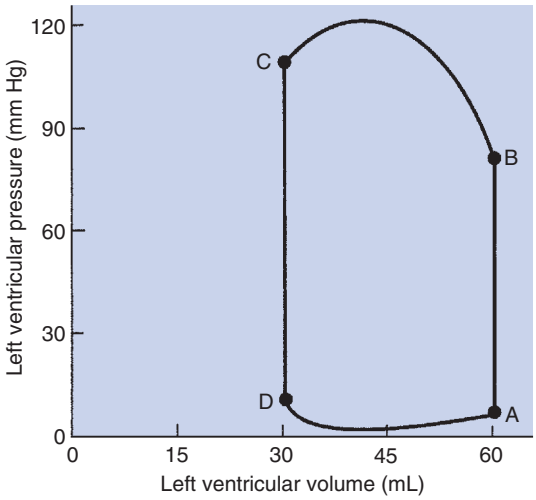
3. An abnormality that causes a sustained decrease in left ventricular compliance is most likely to result in an increased:
- **a.** Left ventricular stroke volume
- **b.** Left atrial volume and pressure
- **c.** Left ventricular afterload
- **d.** Pulmonary blood flow
- **e.** Left ventricular end-diastolic volume

4. Which of the following cause-and-effect statements is true for a normal heart?
- **a.** Sympathetic activation causes end-systolic ventricular volume to increase.
- **b.** An increase in ventricular preload causes end-diastolic ventricular volume to decrease.
- **c.** Pacing the heart at a high rate causes stroke volume to decrease.
- **d.** An increase in ventricular contractility causes systolic duration to increase.
- **e.** An increase in ventricular contractility causes the external work of the heart to decrease.

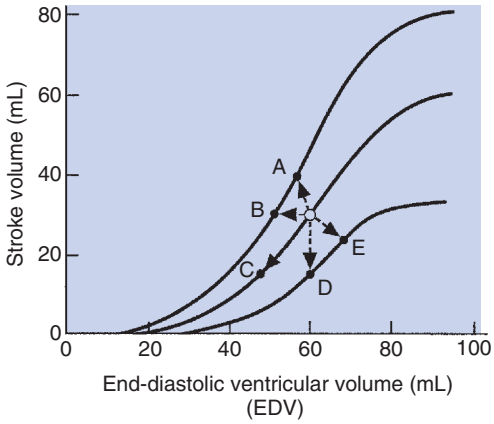
5. Fig. 21.11 shows a plot of the changes in pressure and volume that occur in the left ventricle during one normal cardiac cycle. Which of the following is true regarding this plot?
- **a.** Point D marks the beginning of isovolumetric relaxation.
- **b.** Point B marks the closure of the aortic valve.
- **c.** Point C marks the opening of the mitral valve.
- **d.** Point A marks the beginning of isovolumetric contraction.
- **e.** Point D marks the beginning of ventricular systole.

6. Starting at the open circle in Fig. 21.12, which point would be reached after the contractility decreased and the preload increased?
- **a.** Point A
- **b.** Point B
- **c.** Point C
- **d.** Point D
- **e.** Point E

7. You examine a 7-year-old poodle and find evidence of a systolic murmur (no diastolic murmur), pulmonary edema (indicated by rapid, noisy respiration and cough), left ventricular hypertrophy (no right ventricular hypertrophy), and exercise intolerance. The most likely explanation for the symptoms is:
- **a.** Mitral regurgitation
- **b.** Mitral stenosis
- **c.** Aortic regurgitation
- **d.** Pulmonic stenosis
- **e.** Patent ductus arteriosus



• **Fig. 21.11** Closed loop depicts the changes in left ventricular pressure and volume that occur during one cardiac cycle. Practice Question 5 is based on this graph. The first step in understanding the figure is to determine whether the normal sequence of events proceeds clockwise or counterclockwise around the loop. To make this distinction, recall that the ventricles fill when ventricular pressure is low and they empty when ventricular pressure is high. Next, identify the phases of the cardiac cycle that correspond with each limb of the loop. Finally, determine what happens to the mitral and aortic valves at each corner of the loop. *Hint: A, B, C,* and *D* in this figure match the similarly labeled points in Fig. 21.1 (on the graph of *ventricular pressure*).



• **Fig. 21.12** Practice Question 6 is based on this graph of three ventricular function curves.

Vet-ebooks.com

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Elsevier; 2017.

Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. St. Louis: Elsevier; 2017.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2015.

Hinchcliff KW, Geor RJ. *Equine Exercise Physiology: The Science of Exercise in the Athletic Horse*. Edinburgh; New York: Saunders; 2008.

Katz AM. *Physiology of the Heart*. 5th ed. Baltimore: Lippincott, Williams & Wilkins; 2010.

Keene BW, Smith FWK, Tilley LP, Hansen B. *Rapid Interpretation of Heart and Lung Sounds: A Guide to Cardiac and Respiratory Auscultation in Dogs and Cats*. 3rd ed. St. Louis: Saunders; 2015.

Koeppen BM, Stanton BA. *Berne & Levy Physiology*. 7th ed. Philadelphia: Elsevier; 2018.

Marr C, Bowen M, eds. *Cardiology of the Horse*. 2nd ed. Philadelphia: Saunders; 2011.

Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Philadelphia: Mosby; 2013.

Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. St. Louis: Elsevier; 2016.

Vet-ebooks.com

22

The Systemic and Pulmonary Circulations

ROBERT B. STEPHENSON

KEY POINTS

1. Blood pressure represents a potential energy that propels blood through the circulation.
2. Vascular resistance is defined as perfusion pressure divided by flow.
3. The net resistance of the systemic circulation is called the total peripheral resistance.
4. Mean arterial pressure is determined by the cardiac output and the total peripheral resistance.
5. Blood flow to each organ is determined by perfusion pressure and by the organ's vascular resistance.
6. The pulmonary circulation offers much less resistance to blood flow than does the systemic circulation.
7. Arterial pressures are measured in terms of systolic, diastolic, and mean levels.
8. Pulse pressure increases when the stroke volume increases, heart rate decreases, or aortic compliance decreases.

Blood Pressure Represents a Potential Energy That Propels Blood Through the Circulation

The *systemic circulation* has the aorta as its inlet point and the *venae cavae* as its outlet. The remainder of the circulation (i.e., right heart, pulmonary circuit, and left heart) is, by definition, the *central circulation.*

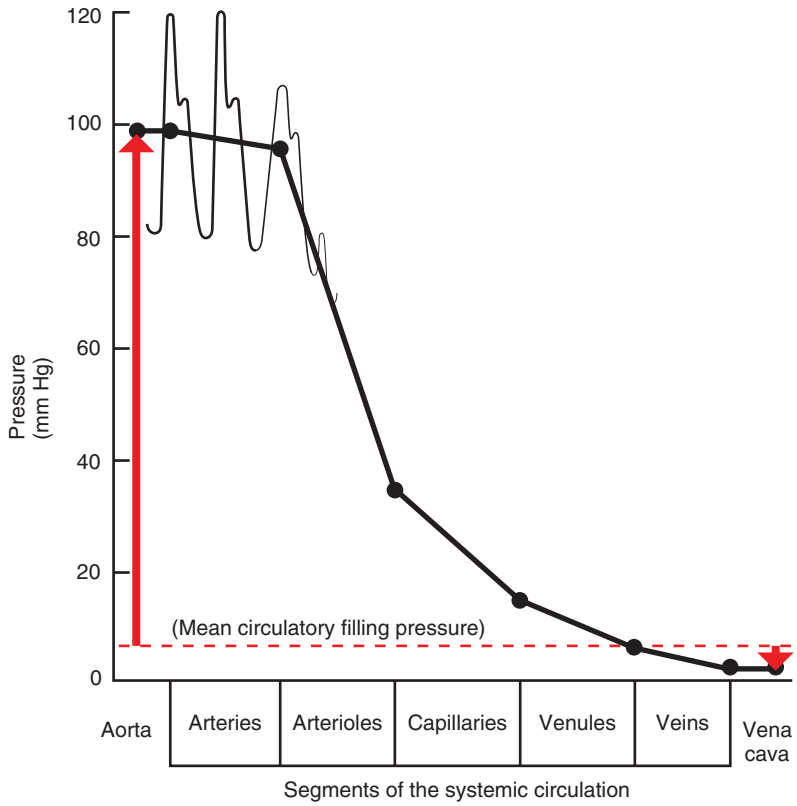
Fig. 22.1 shows the normal pressure profile in the systemic circulation. This figure portrays the pressures that would be measured if a miniature pressure gauge were inserted into the various vessels that blood passes through in its journey through the systemic circulation. Although the data are for a dog, blood pressures are remarkably consistent across a wide range of mammalian species and sizes. Blood pressure is highest in the aorta (typically, mean aortic pressure is 98 mm Hg) and lowest in the venae cavae (typically, 3 mm Hg). The difference (95 mm Hg) between these pressures constitutes the driving force for the movement of blood, by bulk flow, through the systemic circulation. As discussed in Chapter 18, such a pressure difference between the inlet and outlet of a tube (or system of tubes) is called *perfusion pressure difference* (or more commonly, just *perfusion pressure*).

Aortic blood pressure can be thought of as the potential energy available to move blood. The decrease in pressure in the sequential segments of the systemic circuit represents the amount of this potential energy that is "used up" in moving blood through each segment. Pressure energy is used up through *friction,* which is generated as the molecules and cells of blood rub against each other and against the walls of the blood vessels. The energy used up through friction is converted to heat, although the actual increase in the temperature of the blood and blood vessels as a result of friction is very small.

The amount of blood pressure energy used up in each of the sequential segments of the systemic circulation depends on the amount of friction or resistance that the blood encounters. The aorta and large arteries offer very little resistance to blood flow (very little friction), so the blood pressure decreases only a little in these vessels (from 98 to about 95 mm Hg). The greatest pressure decrease (greatest loss of pressure energy through friction) occurs as blood flows through arterioles; that is, the resistance to blood flow is greater in the arterioles than in any other segment of the systemic circulation. The capillaries and the venules offer a substantial resistance to blood flow, but the resistance (and therefore the pressure decrease) is not as great in these vessels as it is in the arterioles. The large veins and the venae cavae are low-resistance vessels, so little pressure energy is expended in driving the blood flow through these vessels.

The pumping of blood by the heart maintains the pressure difference between the aorta and the venae cavae. If the heart stops, blood continues to flow for a few moments from the aorta toward the venae cavae. As this blood leaves the aorta, the aortic walls become less distended, and the blood pressure inside the aorta decreases. As extra blood accumulates in the venae cavae, they become more distended than before, and the blood pressure inside the venae cavae increases. Soon, there is no pressure difference between the aorta and the venae cavae. Blood flow in the systemic circuit ceases, and the pressure everywhere in the systemic circulation is the same. This pressure, in such a static circulation, is called the *mean circulatory filling pressure;* a typical value is about 7 mm Hg. The mean circulatory filling pressure is greater than zero (i.e., above atmospheric pressure), because there is a "fullness" to the circulation; that is, even if the heart stops, blood still distends the vessels that contain it. The vessel walls, being elastic, recoil ("push back") against this distention, and this accounts for the persistence of pressure in the circulation even if the heart stops. If a transfusion of blood is administered to an animal with the heart stopped, the

Vet-ebooks.com



• **Fig. 22.1** Graph of the blood pressures (hydrostatic pressures) that typically exist in the systemic circulation of a dog at rest (*black dots and solid black line*). The blood pressure in the aorta and arteries is pulsatile, increasing with each cardiac ejection and falling between ejections; pulsations disappear by the time the blood reaches arterioles (*fading black line*). The plotted points indicate the average (mean) values of those pulsatile pressures. Mean circulatory filling pressure (*dashed red line*) is the pressure that would persist throughout the systemic circulation if the heart were stopped. *Red arrows* show the contrasting directions and magnitudes of the pressure changes that would occur in the aorta and venae cavae if a stopped heart were restarted and cardiac output returned to normal (see text for details).

vessels become more distended, and the mean circulating filling pressure rises above 7 mm Hg. Conversely, if blood is removed from an animal with the heart stopped, the pressure everywhere falls to a level below 7 mm Hg.

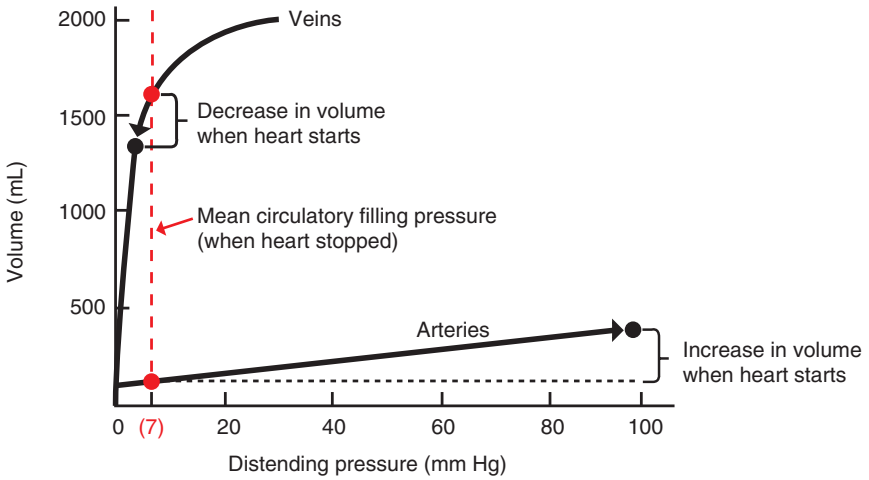
Consider what happens if the heart is restarted in an animal after the pressure has equalized everywhere at 7 mm Hg. With each heartbeat, the heart takes a stroke volume of blood out of the venae cavae and transfers this blood volume (via the pulmonary circulation) into the aorta. The volume of blood in the venae cavae decreases, so the venae cavae become less distended and vena caval pressure drops below 7 mm Hg. The volume of blood in the aorta increases, so the aorta becomes more distended and aortic pressure rises above 7 mm Hg. As illustrated in Fig. 22.1, the vena caval pressure drops about 4 mm Hg (from 7 to 3 mm Hg), and the aortic pressure rises about 91 mm Hg (from 7 to 98 mm Hg). It is important to understand why the pressure decreases only a little in the venae cavae but increases so much in the aorta, even though the volume of blood added to the aorta with each heartbeat is the same as the volume removed from the venae cavae. The reason is that the veins are much more compliant (distensible) than the arteries; one can add or remove blood from veins without changing the venous pressure very much, whereas the addition or removal of blood from arteries changes the arterial pressure a great deal.

A compliant vessel readily distends when pressure or volume is added. It yields to pressure. As already presented in Chapter 21, *compliance* is defined as the change in the volume within a vessel or a chamber divided by the associated change in distending (transmural) pressure, as follows:

$$\text{Compliance} = \frac{\Delta\text{Volume}}{\Delta\text{Transmural pressure}}$$

Compliance corresponds to the slope of a volume-versus-pressure graph. As illustrated in Fig. 22.2, veins are about 20 times more compliant than arteries (over the range of pressures typically encountered in the circulation). That is, veins accept or give up a large volume of blood without incurring much of a change in pressure. For example, veins readily expand or contract to accommodate the changes in blood volume that occur with fluid intake (e.g., drinking) or fluid loss (e.g., sweating). Veins thus function as the major blood *volume reservoirs* of the body. In contrast, arteries function as *pressure reservoirs,* providing the temporary storage site for the surge of pressure energy created with each cardiac ejection. Arteries are tough vessels with low compliance. Arteries accommodate a large increase in pressure during a cardiac ejection and then sustain the pressure high enough between cardiac ejections to provide a continuous flow of blood through the systemic circulation.

Vet-ebooks.com



• **Fig. 22.2** Contrasting relationships between volume (of blood) and distending pressure for veins and arteries. Veins are more compliant than arteries, so they hold a greater volume of blood for a given distending pressure. This concept is illustrated for a distending pressure of 7 mm Hg (*vertical dashed red line*), which is a normal value for the mean circulatory filling pressure (the pressure that would exist everywhere in the circulation if the heart were stopped, as shown in Fig. 22.1). With the heart stopped, the veins contain about 1600 mL of blood and the arteries only 125 mL (*red circles*). If the heart is restarted, its pumping action transfers some blood volume from the veins to the arteries, so venous volume and distending pressure decrease, and arterial volume and distending pressure increase (*black circles*). Because the veins are much more compliant than the arteries, the decrease in venous pressure is very small (decreases from 7 to 3 mm Hg), whereas the increase in arterial pressure is much larger (increases from 7 to 98 mm Hg).

Vascular Resistance Is Defined as Perfusion Pressure Divided by Flow

Everyday experience tells us that it is easier to force fluid through a large tube than through a small tube. For example, it is easier to drink a milkshake through a large-diameter straw than through a small-diameter straw. For a given driving force (perfusion pressure difference), the flow is higher in the large tube because it offers less resistance to flow (less friction) than the small tube. The precise definition of resistance is:

$$\text{Resistance} = \frac{\Delta\text{Pressure}}{\text{Flow}}$$

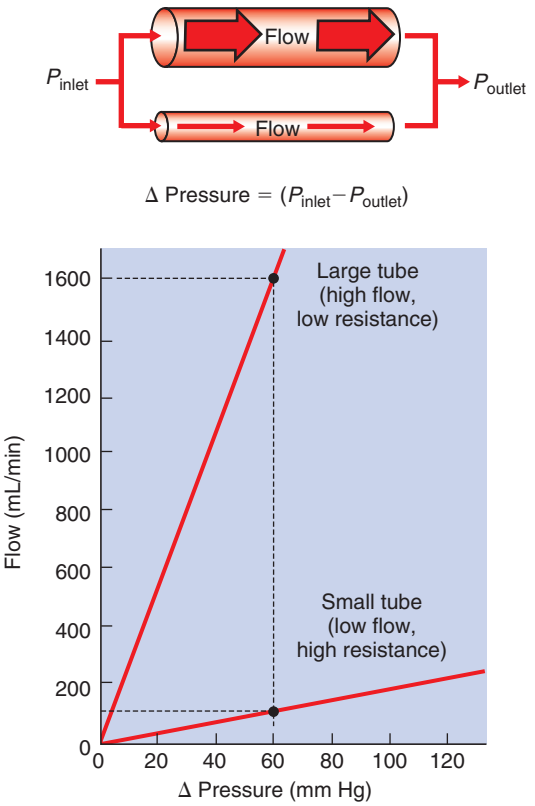
where Δ Pressure is *perfusion pressure difference,* or simply *perfusion pressure* (i.e., the pressure at the tube inlet minus the pressure at its outlet).

Fig. 22.3 presents these concepts in pictorial and graphic form. The dashed lines in this figure indicate that a perfusion pressure of 60 mm Hg causes a flow of 1600 milliliters per minute (mL/min) through the large tube. Thus the resistance of the large tube is 37.5 mm Hg/L/min. The same perfusion pressure (60 mm Hg) causes a flow of only 100 mL/min through the small tube. The resistance of the small tube is therefore 600 mm Hg/L/min. The resistance of the small tube is 16 times greater than the resistance of the large tube.

In the late 1800s, the French physicist and physician J.L.M. Poiseuille demonstrated the dominant effect of radius on the resistance of a tube. He showed the following:

$$\text{Resistance of a tube} \cong \frac{8\eta l}{\pi r^4}$$

where l is the length of the tube, r is the radius, η is the viscosity of the fluid flowing through the tube, and π has its usual meaning.



• **Fig. 22.3** Relationship between fluid flow and perfusion pressure (Δ Pressure) for two tubes. The perfusion pressure is the pressure at the inlet (P_{inlet}) of the tube minus the pressure at the outlet (P_{outlet}). In this example, the larger tube has twice the radius of the smaller tube. For a given perfusion pressure, the flow through the larger tube is 16 times greater than the flow through the smaller tube. That is, the resistance of the larger tube is one-sixteenth the resistance of the smaller tube.

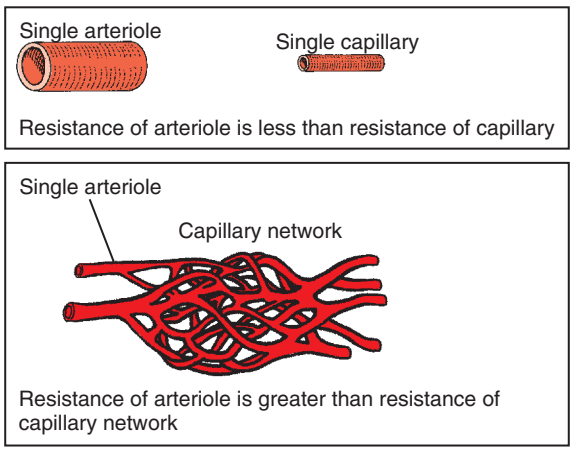
Vet-ebooks.com

This equation (*Poiseuille's law*) emphasizes that radius (r) is the dominant factor influencing the resistance of a tube; resistance varies inversely with the fourth power of radius. Doubling the radius of a tube decreases its resistance by a factor of 16 (2^4). This explains why using a larger diameter straw makes it so much easier to drink a milkshake. Resistance is also influenced by the length (l) of the tube; it is harder to force fluid through a long tube than through a short tube of the same radius. The final determinant of resistance is the viscosity (η) of the fluid. The higher the viscosity of the fluid, the higher is the resistance to its flow through a tube. For example, honey is more viscous than water, so a tube offers a higher resistance to the flow of honey than to the flow of water.

As already described in connection with Fig. 22.1, the arterioles are the segment of the circulation in which there is the greatest decrease in blood pressure (greatest loss of pressure energy through friction). That is, the arterioles are the segment of the systemic circulation with the highest resistance to blood flow. It may seem paradoxical that the arterioles are the site of highest resistance when the capillaries are smaller vessels. After all, Poiseuille's law and Fig. 22.3 emphasize that a smaller tube has a much higher resistance than a larger tube. The resolution of this paradox is presented in Fig. 22.4. It is true that each capillary has a smaller radius and therefore a greater resistance than each arteriole. However, each arteriole in the body distributes blood into a network of many capillaries, and the *net resistance* of all those capillaries is less than the resistance of the single arteriole that delivers blood to them. It is only because each arteriole delivers blood to so many capillaries that the net resistance of the capillaries is less than the resistance of the arteriole.

Arterioles are the site not only of the highest resistance in the circulation but also of adjustable resistance. Variation in arteriolar resistance is the main factor that determines how much blood flows through each organ in the body; an increase in arteriolar resistance in an organ decreases the blood flow through that organ and vice versa. Arterioles change their resistance, moment to moment, by changing their radius. (The length of an arteriole does not change, at least not over the short term.) The walls of arterioles are relatively thick and muscular. Contraction of the arteriolar smooth muscle decreases the radius of arterioles, and this *vasoconstriction* substantially increases resistance to blood flow. Relaxation of the smooth muscle allows the radius of the vessels to increase, and this *vasodilation* substantially reduces the resistance to blood flow.

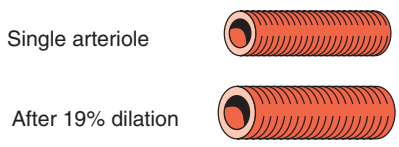
Fig. 22.5 illustrates that a small change in the radius of arterioles in an organ (a brain in this case) brings about a large change in resistance and therefore in blood flow. In this example, the arterial pressure is 93 mm Hg and the venous pressure is 3 mm Hg, so the perfusion pressure is 90 mm Hg. The brain blood flow is initially observed to be 90 mL/min. Based on the mathematical definition of resistance, the resistance of the brain blood vessels is 1 mm Hg/mL/min. Most of this resistance is provided by the brain arterioles. Next, consider the consequence of a slight vasodilation, such that the radius of the arterioles increases by 19% (e.g., from a radius of 1.00 to a radius of 1.19). Recall from Poiseuille's law that the resistance varies inversely as the fourth power of the radius. Because 1.19^4 equals 2.00, a 19% increase in radius cuts the resistance in half. Decreasing the brain's resistance by half (to 0.5 mm Hg/mL/min) would double the brain blood flow (to 180 mL/min).



• **Fig. 22.4** The resistance of a single arteriole is less than the resistance of a single capillary, because arterioles are larger in diameter. However, each arteriole supplies blood to a whole network of capillaries, and the resistance of an arteriole is greater than the resistance of the capillary network that it supplies with blood.

Initially, blood flow through the brain = 90 mL/min
(as shown in figure on right)

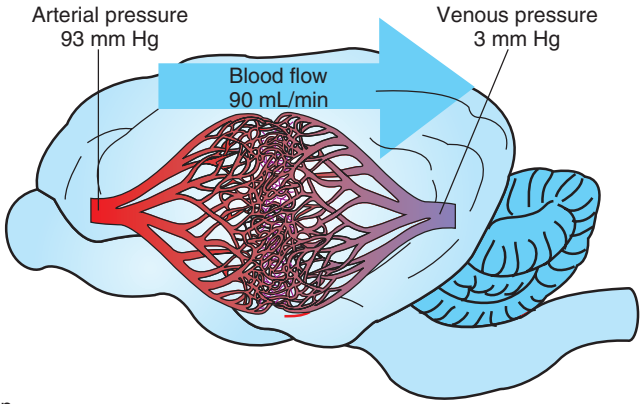
$$\text{Resistance} = \frac{\Delta\text{Pressure}}{\text{Blood flow}} = \frac{(93–3)\text{ mm Hg}}{90\text{ mL/min}} = 1\text{ mm Hg/mL/min}$$



19% increase in arteriolar radius cuts resistance in half,
and doubles blood flow (see discussion in text).

After vasodilation, resistance = 0.5 mm Hg/mL/min

$$\text{Blood flow} = \frac{\Delta\text{Pressure}}{\text{Resistance}} = \frac{(93–3)\text{ mm Hg}}{0.5\text{ mm Hg/mL/min}} = 180\text{ mL/min}$$



• **Fig. 22.5** Example illustrating that a small arteriolar dilation (vasodilation) would substantially increase blood flow to an organ (brain, in this case).

Vet-ebooks.com

The Net Resistance of the Systemic Circulation Is Called the Total Peripheral Resistance

Resistance is always defined as a pressure difference (perfusion pressure) divided by flow. In a calculation of the resistance of the systemic circulation, the perfusion pressure is the pressure in the aorta minus the pressure in the venae cavae. The flow is the total amount of blood that flows through the systemic circuit, which is equal to the cardiac output:

$$\text{Systemic vascular resistance (SVR)} = \frac{(\text{Mean aortic pressure} - \text{Mean vena caval pressure})}{\text{Cardiac output}}$$

A synonym for systemic vascular resistance is *total peripheral resistance* (*TPR*).

For a typical dog at rest, the mean aortic pressure is 98 mm Hg, the mean vena caval pressure is 3 mm Hg, and the cardiac output is 2.5 L/min. Under these conditions, TPR is 38 mm Hg/L/min, which means that it takes a driving pressure of 38 mm Hg to force 1 L/min of blood through the systemic circuit.

Because the pressure in the venae cavae is usually close to zero, it is sometimes ignored in the calculation of TPR. The resultant simplified equation states that TPR is approximately equal to mean aortic pressure divided by the cardiac output. Usually, this equation is rearranged to form the statement that the mean aortic blood pressure (Pa) is approximately equal to the cardiac output (CO) multiplied by TPR:

$$Pa \cong CO \times TPR$$

This equation expresses one of the central concepts in cardiovascular physiology, namely that mean aortic blood pressure is determined by two, and only two, factors. Thus if the aortic pressure is increased, it must be because the cardiac output increased, the TPR increased, or both. There are no other possibilities.

Mean Arterial Pressure Is Determined by the Cardiac Output and the Total Peripheral Resistance

Three examples illustrate the application of the concept that the mean aortic blood pressure is determined by cardiac output and TPR. First, in the most common form of human essential hypertension, the cardiac output is normal. The blood pressure is elevated because of excessively constricted systemic arterioles, which increases TPR above normal. What remains unclear about human essential hypertension is why the arterioles are constricted. High blood pressure is a serious health problem in human medicine because patients with uncontrolled hypertension often develop cardiac hypertrophy, and they are at high risk for cardiac arrhythmias, myocardial infarction, renal failure, and stroke. Naturally occurring hypertension is rare in veterinary species, although several techniques have been developed to induce hypertension in laboratory animals for research purposes.

Severe hemorrhage or dehydration is another condition in which the arterial pressure becomes abnormal, and it provides several distinct contrasts to chronic hypertension. Hemorrhage and dehydration are often encountered in veterinary medicine. Also, the arterial pressure is reduced in these conditions, not elevated. The cause of the decreased pressure is a decreased cardiac output. Hemorrhage or dehydration characteristically reduces the cardiac preload, which reduces the stroke volume and cardiac output. TPR is actually increased above normal because the body constricts the arterioles in the kidneys, splanchnic circulation, and resting skeletal muscle. Vasoconstriction in these organs minimizes the fall in arterial pressure and diverts the available cardiac output to the organs most critical for moment-to-moment survival, which include the brain, exercising skeletal muscle, and the heart (i.e., coronary circulation).

The response to vigorous exercise provides a third application of the concept that the mean aortic blood pressure is determined by the cardiac output and TPR. As in hemorrhage, exercise causes the cardiac output and TPR to change in opposite directions. In exercise, however, the cardiac output is elevated, and TPR is decreased. TPR decreases because the arterioles in the exercising skeletal muscle dilate, which increases skeletal muscle blood flow. During vigorous exercise, TPR decreases to about one-fourth of its resting value. The cardiac output increases about fourfold. The result is that the aortic pressure is negligibly changed. Fig. 22.6 depicts the cardiovascular adjustments to vigorous exercise.

Blood Flow to Each Organ Is Determined by Perfusion Pressure and by the Organ's Vascular Resistance

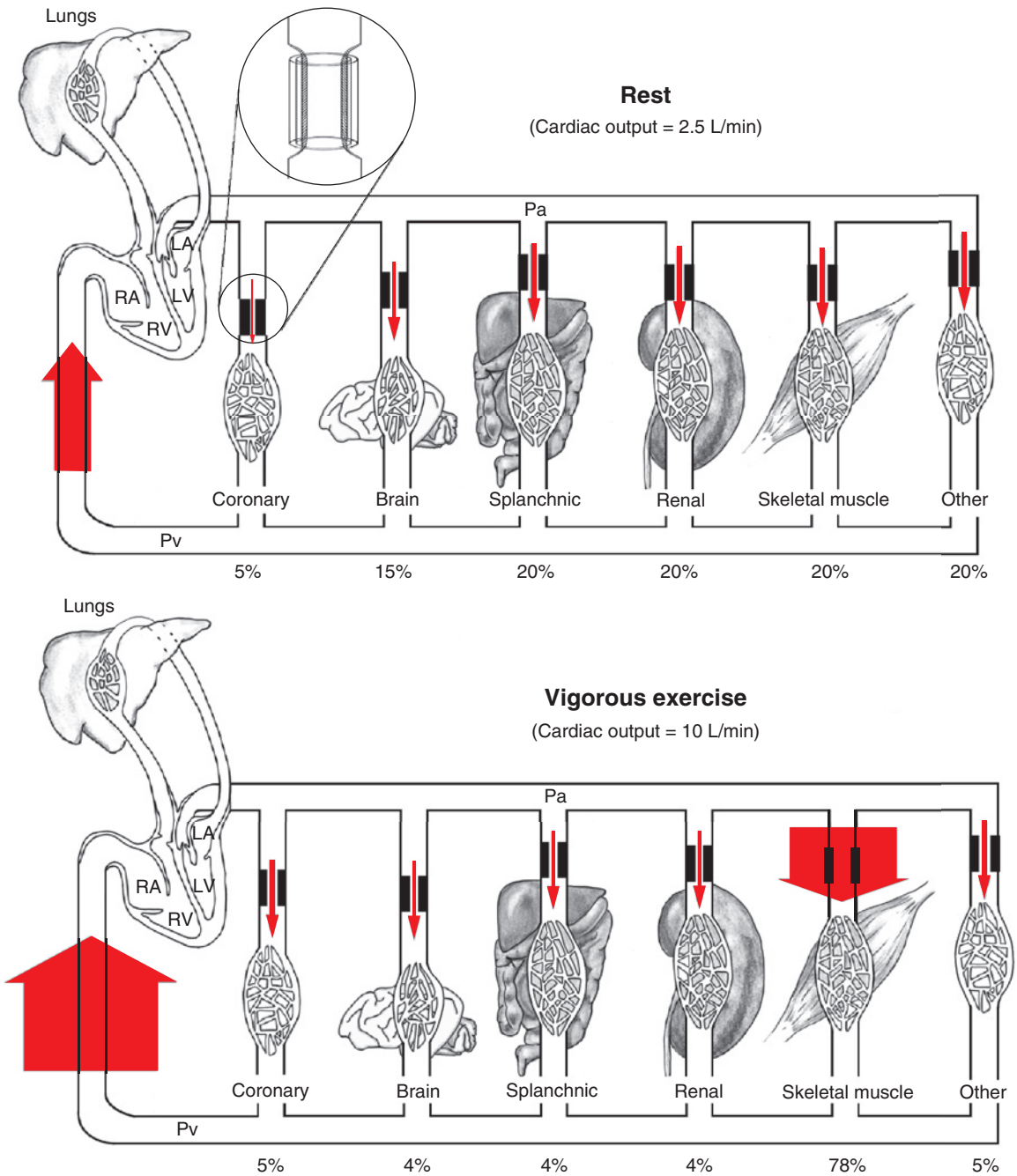
If the equation that defines resistance is solved for flow, the result is:

$$\text{Flow} = \frac{\Delta\text{Pressure}}{\text{Resistance}}$$

As applied to the blood flow through any organ, this equation points out that the blood flow is determined by perfusion pressure (mean arterial pressure minus mean venous pressure) and by the resistance of the organ's blood vessels. There are no other factors. All the organs of the systemic circulation receive arterial blood flow via branches of the aorta, so all are exposed to essentially the same arterial pressure. Similarly, the venous blood from all the organs of the systemic circulation is collected into the venae cavae, so under usual circumstances, mean venous pressure is the same for all organs. Given that all the systemic organs are exposed to nearly the same perfusion pressure, it follows that the differences in blood flow to the various organs result solely from their different vascular resistances. As already explained, the vascular resistance of an organ is determined primarily by the diameter of its arterioles. Thus arteriolar vasodilation and vasoconstriction are the primary mechanisms that increase or decrease the blood flow in one organ relative to another organ.

Fig. 22.6 illustrates how changes in the vascular resistance of various organs alter the distribution of cardiac output among the organs. In a typical dog at rest, the arteriolar resistances are similar in the splanchnic, renal, and skeletal vascular beds. Therefore each of these beds receives about the same blood flow (indicated in Fig. 22.6 by arrows of equal width). During exercise, skeletal muscle arterioles dilate greatly, almost doubling in diameter, which decreases their resistance to blood flow by a factor of almost 16. Therefore the skeletal muscle blood flow increases almost 16-fold (from 0.5 to 7.8 L/min). Also during exercise, coronary arterioles dilate, and the coronary blood flow increases dramatically. Brain arterioles remain the same, so the brain blood flow is unchanged. By contrast,

Vet-ebooks.com



• **Fig. 22.6** Cardiac output and its distribution compared during rest (*top*) and vigorous exercise (*bottom*) in a typical large dog. The width of the *red arrows* denotes the amount of blood flow. The flow of blood into the right side of the heart (which is equal to the cardiac output) is represented by the *very wide arrows* on the left. The cardiac output increases fourfold during exercise (from 2.5 L/min to 10.0 L/min). The entire cardiac output passes through the lungs and then is pumped by the left ventricle (*LV*) into the systemic arterial system (*horizontal tube across top*). The systemic arteries deliver blood to each of the systemic vascular beds, which are grouped here into *Coronary, Brain, Splanchnic, Renal, Skeletal muscle,* and *Other.* In each systemic organ, blood must pass through high-resistance arterioles (*heavy black bars*) before reaching the capillaries. The arterioles act as adjustable cuffs or constrictors (see magnified view, *top*). The proportion of the total cardiac output that passes through each organ is indicated by a percentage at the bottom. Because each organ is exposed to the same arterial pressure (*Pa*) and venous pressure (*Pv*), the proportion of cardiac output that each organ receives is determined by its resistance. Resistance is determined primarily by the arteriolar diameter, which is indicated by the size of the opening between the heavy black bars. During vigorous exercise, skeletal muscle arterioles dilate enough that the blood flow to the exercising muscles may increase almost 16-fold (from 0.5 L/min at rest to 7.8 L/min in this example). Coronary arterioles also dilate, and the coronary blood flow increases about fourfold, which meets the increased demand by the heart muscle for oxygen. Vasoconstriction causes a small decrease in blood flow to the splanchnic and renal circulations. Blood flow to the brain is basically unchanged, although the percentage of total cardiac output received by the brain decreases. *RV,* Right ventricle; *LA,* left atrium; *RA,* right atrium.

Vet-ebooks.com

the arterioles in the splanchnic and renal circulations constrict slightly during exercise, which causes splanchnic and renal resistance to increase by about 20%. Therefore the splanchnic and renal blood flows decrease by about 20% (from 0.5 to 0.4 L/min each).

This discussion of blood flow during exercise describes the responses of a normal dog with a healthy heart. Such a dog can readily increase its cardiac output enough to meet the increased blood flow needs of the skeletal and cardiac muscle. As a consequence, the arterial pressure (and hence the perfusion pressure) is very similar during rest and exercise. By contrast, a dog with heart failure cannot increase its cardiac output much above its resting level. Therefore the arterial pressure (and perfusion pressure) declines during exercise, and none of the organs receives the blood flow that it requires. This is why animals with heart failure exhibit weakness, fatigue, and exercise intolerance. (Additional complications of heart failure are discussed in Chapter 26.) The point for now is that the equation that relates blood flow, perfusion pressure, and vascular resistance is fundamental and inescapable; this relationship is profoundly important to an understanding of cardiovascular function and dysfunction.

The Pulmonary Circulation Offers Much Less Resistance to Blood Flow Than Does the Systemic Circulation

As with any other resistance, pulmonary resistance is calculated as a pressure difference (perfusion pressure) divided by a flow. The perfusion pressure that forces blood through the pulmonary circuit is the pressure in the pulmonary artery minus the pressure in the pulmonary veins. The flow that traverses the pulmonary circuit is equal to the cardiac output. Therefore:

$$\text{Pulmonary vascular resistance} = \frac{\text{(Mean pulmonary artery pressure} - \text{Mean pulmonary venous pressure)}}{\text{Cardiac output}}$$

For a typical dog at rest, the mean pulmonary arterial pressure is 13 mm Hg, the mean pulmonary venous pressure is 5 mm Hg, and the cardiac output is 2.5 L/min. Thus pulmonary resistance is 3.2 mm Hg/L/min. Note that this is only about one-twelfth the resistance of the systemic circulation. Because the majority of the systemic resistance is in the arterioles, it is valid to conclude that the pulmonary arterioles offer substantially less resistance than do the systemic arterioles.

The entire cardiac output passes through the lungs, so a fourfold increase in cardiac output during exercise also necessitates a fourfold increase in pulmonary blood flow. Because pulmonary blood vessels are quite compliant, they readily distend to accept the increase in blood flow. Because even a small increase in vessel radius greatly decreases resistance (in accordance with Poiseuille's law, as mentioned earlier), the resistance of the pulmonary blood vessels drops greatly during exercise. The decreased pulmonary resistance during exercise is advantageous because it allows the pulmonary flow to increase greatly without necessitating a large increase in the pulmonary arterial pressure.

Chapters 46 and 47 present additional details about the characteristics of pulmonary blood flow, including an explanation of the mechanisms that adjust the vascular resistance in various regions of the lungs so that the amount of blood that flows through each region of the lungs is appropriately matched to the amount of fresh air delivered to the alveoli in that region (*ventilation-perfusion matching*).

Arterial Pressures Are Measured in Terms of Systolic, Diastolic, and Mean Levels

Blood pressure within the aorta is not constant but rather is *pulsatile,* as shown in Figs. 21.1, 22.1 and 22.7. Pulmonary arterial pressure is also pulsatile. With each cardiac ejection, the aorta and pulmonary artery become distended with blood, which causes the pressures within these vessels to increase to peak values, called *systolic arterial pressures.* Between cardiac ejections (i.e., during ventricular diastole), blood continues to flow out of the aorta and pulmonary artery into the systemic and pulmonary circulations, respectively. As the volume of blood remaining in these large arteries decreases, the arteries become less distended, so arterial pressure decreases. Pressure continues to decrease until the next cardiac ejection begins. The minimal pressure reached before each new ejection is called the *diastolic arterial pressure.* Fig. 22.7 illustrates typical values for systolic and diastolic pressures.

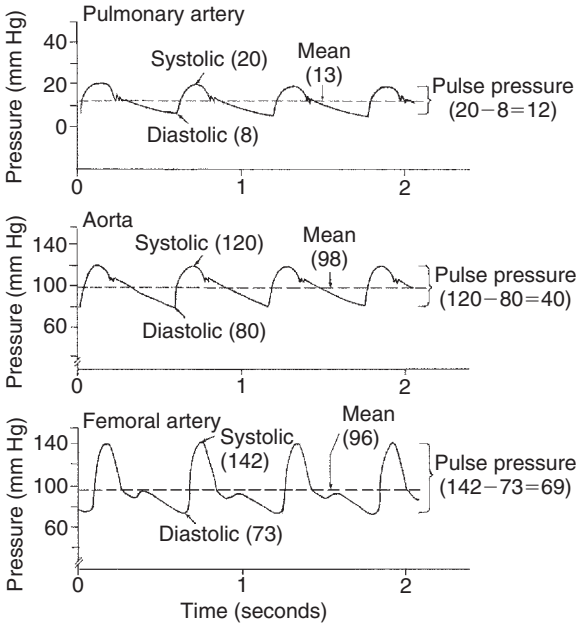
The amplitude of the pressure pulsations in an artery is called the *pulse pressure,* specifically:

$$\text{Aortic pulse pressure} = (\text{Aortic systolic pressure} - \text{Aortic diastolic pressure})$$

and

$$\text{Pulmonary artery pulse pressure} = (\text{Pulmonary artery systolic pressure} - \text{Pulmonary artery diastolic pressure})$$

Typical values for pulse pressure are given in Fig. 22.7. Note how much lower the systolic, diastolic, and pulse pressures are in the pulmonary artery than in the aorta. These differences illustrate



• **Fig. 22.7** Blood pressure in the large arteries is pulsatile. The pressure patterns typical of the pulmonary artery, aorta, and femoral artery of the dog are shown.

Vet-ebooks.com

why the pulmonary circulation is called the *low-pressure, low-resistance side of the circulation,* whereas the systemic circulation is called the *high-pressure, high-resistance side of the circulation.*

It is important to distinguish among systolic pressure, diastolic pressure, and pulse pressure, and to distinguish all of them from *mean pressure.* Mean aortic pressure is the average pressure in the aorta over the course of one or more complete cardiac cycles. Likewise, mean pulmonary artery pressure is the average pressure in that vessel. Obviously, the mean pressure in an artery is somewhere between the systolic (maximal) and diastolic (minimal) pressure levels. However, because the pressure waveforms in arteries are not symmetrical, the mean pressure is generally not exactly midway between the systolic and diastolic pressures.

A popular approximation is that mean pressure is about one-third of the way up from diastolic toward systolic pressure; that is:

$$\text{Mean arterial pressure} \cong \text{Diastolic pressure} + \tfrac{1}{3}\,\text{Pulse pressure}$$

Fig. 22.7 reveals that this is *not* a valid approximation for the determination of mean pressure in the aorta. However, the approximation is a good one for pressures measured in the femoral artery or in most other major arteries distal to the aorta. The reason that the rule applies in the distal arteries but not in the aorta is that the waveform of the arterial pressure pulsations changes as the pulses move away from the heart. The physics behind the changing pressure waveform is complex; however, one consequence is that the mean level of pressure in the distal arteries is well below the midpoint between diastolic and systolic arterial pressures (see pressure waveform for femoral artery in Fig. 22.7).

Another consequence of the aforementioned changing arterial pressure waveform is that pulse pressure typically *increases* as blood flows from the aorta into the distal arteries. However, the mean pressure necessarily *decreases* in accordance with the principle of the conservation of energy. As stated earlier, mean arterial pressure is a measure of the potential energy in the bloodstream, and this potential energy is used up (converted into heat by friction) as blood flows from the aorta through the systemic circulation. The aorta and large arteries offer only a small resistance to blood flow, which explains why mean arterial pressure decreases only 1 to 3 mm Hg between the aorta and the femoral artery (see Fig. 22.7). Most of the resistance to blood flow is found in the arterioles and capillaries. Therefore the largest decrements in *mean pressure* occur in the arterioles and capillaries (see Fig. 22.1).

An important point to remember is that mean aortic pressure (not systolic, diastolic, or pulse pressure) must be used when calculating TPR as:

$$\text{Total peripheral resistance} = \frac{\text{(Mean aortic pressure} - \text{Mean vena caval pressure)}}{\text{Cardiac output}}$$

Likewise, mean pulmonary artery pressure (not systolic, diastolic, or pulse pressure) must be used when calculating pulmonary vascular resistance as:

$$\text{Pulmonary vascular resistance} = \frac{\text{(Mean pulmonary arterial pressure} - \text{Mean pulmonary vein pressure)}}{\text{Cardiac output}}$$

Unfortunately, making an accurate measurement of a mean vascular pressure requires inserting a needle or catheter into the vessel of interest. The first direct measurement of mean arterial blood pressure was carried out by Stephen Hales, an English clergyman. In about 1730, Hales inserted a tube (catheter) into the femoral artery of a conscious horse and found that blood rose in the tube to a height of more than 8 feet. An 8-foot column of blood represents a pressure of more than 180 mm Hg, almost twice the mean arterial pressure expected in a normal resting animal. The high pressure undoubtedly reflected the physical and emotional distress of the horse, which was restrained upside down during the episode. In the present day, arterial catheterization (with anesthetic agents to reduce pain) is routine in human medicine (e.g., in cardiac catheterization laboratories) and is becoming more common in veterinary medicine. However, the lesson that physical or emotional distress can dramatically increase blood pressure is as relevant today as it was in Hales' time.

In human medicine, systolic and diastolic pressures in the brachial artery can be measured quite accurately with a blood pressure cuff and stethoscope. Mean arterial pressure can then be approximated using the equation given earlier. Blood pressure cuffs are less frequently used on veterinary species, but the pulse is often palpated by placing the fingertips over a major artery, such as the femoral artery. Palpation of an artery allows the clinician to sense the pulse pressure on the basis of the magnitude of the pulsations felt in the artery. A low pulse pressure is referred to as a *thready,* or weak, pulse. A high pulse pressure may be called a *bounding,* or strong, pulse.

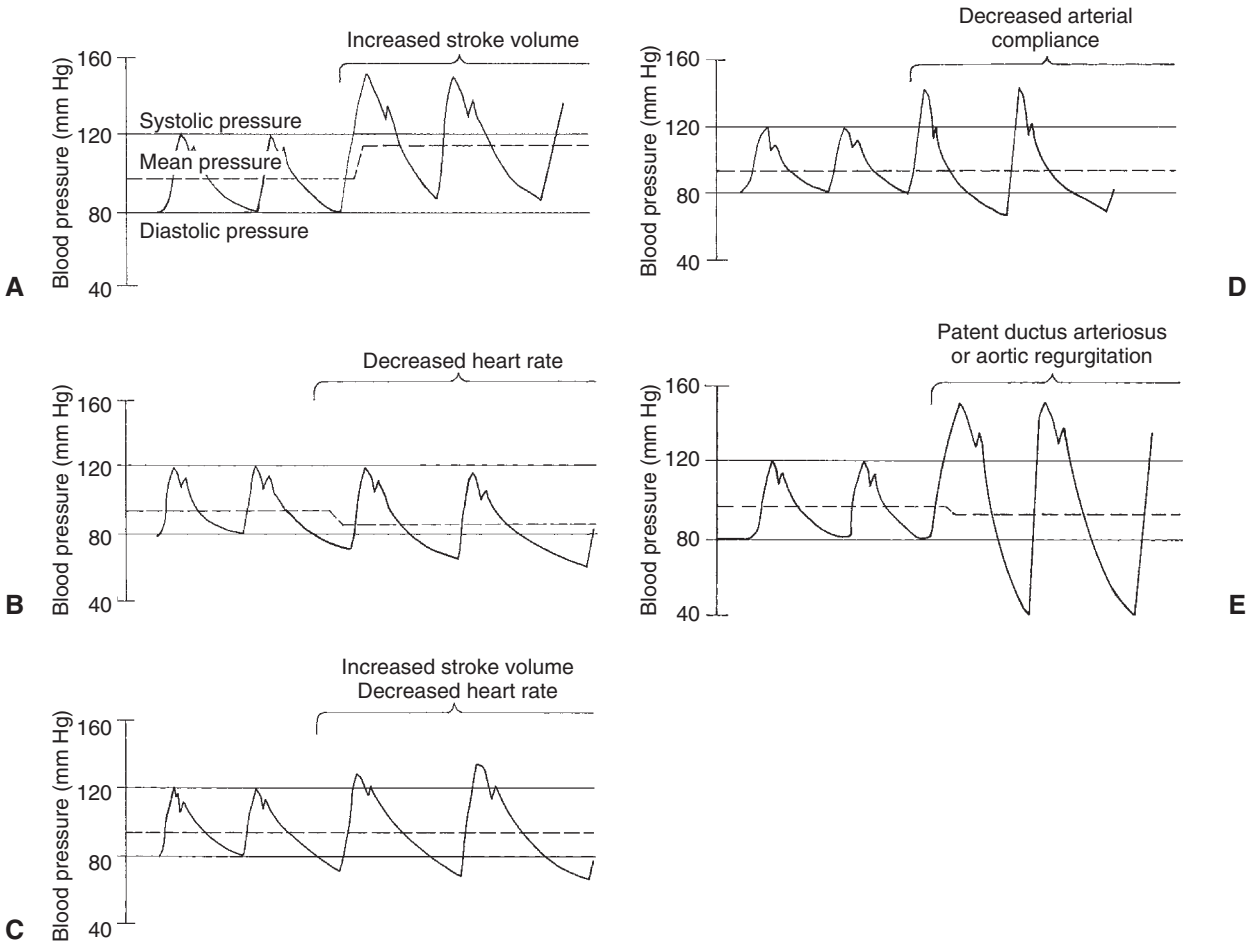
Pulse Pressure Increases When the Stroke Volume Increases, Heart Rate Decreases, or Aortic Compliance Decreases

Because the arterial pulse is so frequently palpated in patients, it is important for the veterinary clinician to understand the factors that typically influence pulse pressure. First, an increase in stroke volume tends to increase pulse pressure. Because cardiac ejections create the arterial pulsations in the first place, it is not surprising that larger ejections create larger pulsations. Fig. 22.8A depicts this effect and shows that an increase in stroke volume also increases mean arterial pressure. Mean pressure increases as well, because an increased stroke volume increases cardiac output.

A second factor that tends to increase pulse pressure is a decrease in heart rate. In the time between cardiac ejections, pressure in the aorta decreases because blood continues to leave the aorta and to flow out through the arteries, arterioles, and capillaries of the systemic circulation. Aortic pressure falls to a minimal (diastolic) level before being boosted again by the next cardiac ejection. When heart rate decreases, there is a longer time between ejections and therefore a longer time for blood to run out of the aorta and into the systemic circulation. As a consequence, blood pressure in the aorta falls to a lower level before the next cardiac ejection, and pulse pressure is increased (Fig. 22.8B). A decrease in heart rate results in a decrease in cardiac output, so a decrease in heart rate decreases the mean arterial pressure (Fig. 22.8B).

Fig. 22.8C shows the effect of a simultaneous increase in stroke volume and decrease in heart rate. In this example, the increase in stroke volume and decrease in heart rate are balanced, such that cardiac output, which is stroke volume multiplied by heart rate, remains unchanged. Therefore mean arterial pressure remains unchanged. However, pulse pressure is greatly increased as a result of the combined effects of an increase in stroke volume and a decrease in heart rate. In humans and some animals, aerobic conditioning leads to increased stroke volume and decreased heart

Vet-ebooks.com



• **Fig. 22.8** Various conditions that increase arterial pulse pressure are compared with regard to their effects on systolic pressure, diastolic pressure, and mean pressure (see text).

rate at rest. In consequence, a well-trained athlete at rest typically has a normal mean arterial pressure but a markedly elevated pulse pressure. Arterial palpation of an athlete at rest typically reveals a strong, slow pulse.

A decrease in arterial compliance (stiffening of the arteries or "hardening of the arteries") is a third factor that tends to increase pulse pressure (see Fig. 22.8D). With each ventricular systole, the heart ejects a stroke volume of blood into the aorta and large arteries, which distends these vessels. If these vessels become stiff, a greater increase in pressure is required to bring about the same distension. Arterial stiffening also decreases diastolic arterial pressure. This effect is more difficult to grasp intuitively but should not be surprising. Just as aortic pressure rises to higher-than-normal systolic levels when the heart ejects blood into a stiff aorta, so does aortic pressure fall to lower-than-normal diastolic levels when blood runs out of the stiff aorta between cardiac ejections. The higher systolic pressure and lower diastolic pressure are simply two direct consequences of the same phenomenon: decreased arterial compliance. The major arteries tend to become stiffer as a result of the normal aging process, which accounts for the increase in pulse pressure typical in older humans and some animals.

In general, cardiac output is not much affected by arterial stiffening. A healthy ventricle is able to generate the higher systolic pressures needed to eject blood into a stiff arterial system, although ventricular hypertrophy is sometimes triggered. Stiff arteries typically retain their large diameters, and therefore the resistance of these big vessels remains low. The arterioles (the site of most of the peripheral resistance) remain normal. Consequently, arterial stiffening generally has very little effect on TPR. *Mean* arterial pressure (the product of cardiac output and TPR) is therefore generally not much changed by arterial stiffening.

In summary, pulse pressure tends to be increased by increased stroke volume, decreased heart rate, or decreased arterial compliance.

Some of the cardiac defects that produce murmurs also cause characteristic changes in pulse pressure. For example, a patient with patent ductus arteriosus has a large left ventricular stroke volume, which elevates aortic systolic pressure. Aortic diastolic pressure is much lower than normal because, between cardiac ejections, blood runs out of the aorta by two pathways: into the systemic circuit and through the open ductus. Consequently, aortic pulse pressure is dramatically increased (Fig. 22.8E). Aortic regurgitation causes a similar, characteristic increase in pulse pressure. During diastole, blood leaves the aorta through two pathways: forward into the systemic circuit and backward (through the incompetent valve) into the left ventricle. This additional run-off of blood from the aorta during diastole causes aortic diastolic pressure to be lower than normal. Stroke volume is elevated because, with each systole, the left ventricle ejects both the regurgitant blood and also the blood that has returned to it through the normal venous pathway. The increased stroke volume causes aortic systolic pressure to be higher than normal.

Vet-ebooks.com

CLINICAL CORRELATIONS

Canine Heartworm Disease With Pulmonary Embolism

History

You examine a 6-year-old male beagle that has been a hunting companion of his owner for several years. The owner reports that the dog tires more easily than usual and has developed a cough that is worse during exercise. You treated this dog for a laceration when he was 3 years old, and your records indicate that the dog was otherwise in excellent health at that time. The owner acknowledges that the dog has not been given any immunizations or heartworm prophylactic medication for the past 2 years.

Clinical Examination

On physical examination of the dog, you notice the cough reported by the owner and an apparent, modest accumulation of fluid in the abdominal cavity (*ascites*). You also auscultate a systolic murmur that is loudest over the left third and fourth intercostal spaces. The chest radiograph and electrocardiogram show evidence of right ventricular enlargement. In addition, the pulmonary vessels are more prominent than normal on the radiograph and are tortuous (twisted). You suspect canine heartworm disease. You obtain a blood sample, some of which you submit for an enzyme-linked immunosorbent assay (ELISA) to test for heartworm antigen. Additionally, you use a pipette to apply a sample of the buffy coat (from the centrifuge tube) onto a glass slide for microscopical examination. You see microfilaria of the type shed by adult canine heartworms (*Dirofilaria immitis*), and the ELISA is positive for the presence of *D. immitis* antigen. You diagnose canine heartworm parasitism.

Comment

Mosquitoes transfer the microfilaria from the bloodstream of an infected dog to the bloodstream of a noninfected dog. The microfilaria develop into adult worms, which grow to a length of 10 to 20 cm while clinging to the walls of the pulmonary artery and its major branches. Heartworm infestation typically causes pulmonary arterial vessels to become enlarged and tortuous. In heavily infested dogs, adult worms also reside in the right ventricle and right ventricular outflow tract, where they cause pulmonic stenosis. The resulting turbulence during right ventricular ejection accounts for the murmur heard in this dog. The pulmonic stenosis and the increased pulmonary resistance created by the worms also lead to right ventricular hypertrophy, exercise intolerance, and ascites (review the Clinical Correlation on pulmonic stenosis in Chapter 21 for an explanation of why these complications develop). An additional problem is that the adult worms release vasoactive substances into the circulation, which disrupt some of the normal mechanisms that adjust arteriolar diameter, control blood flow, and regulate arterial pressure. Heavily infested dogs become very ill.

Treatment

You advise the owner that the dog should be treated with an arsenic-containing medication that kills adult worms over several days. You also warn the owner that the treatment of severely infested dogs is risky. Dead adult worms break away from the right ventricle and pulmonary artery and lodge in smaller pulmonary vessels. These vascular occlusions (*pulmonary emboli*) restrict pulmonary blood flow and reduce cardiac output. Therefore it is necessary to keep the dog in a quiet, unstressed state for 8 to 10 days after beginning treatment. In addition to restricting pulmonary blood flow, the emboli are likely to cause inflammation and blood clots in the lungs. Pulmonary edema is expected. Pulmonary blood vessels may break down, allowing blood to enter the airways of the lungs. Respiratory failure is possible. Antiinflammatory drugs are sometimes administered to reduce these complications.

With the owner's consent, you keep the dog at your clinic for 2 days (to allow him to become accustomed to the surroundings) and then begin treatment. During the next week, the dog becomes even more lethargic than before and begins to cough up blood. The dog has a low-grade fever (102°–103° F), and his ascites becomes worse. However, his systolic murmur begins to fade. After 1 week, all the clinical signs have greatly improved. The dog is sent home for a prolonged period of recuperation. The long-term prognosis is good.

Dummy Foal: Hypoxemic Ischemic Encephalopathy

History

A 14-year-old thoroughbred mare is presented for *dystocia* (difficult birth). The foal (a filly) was pulled with some difficulty. The filly was slow to stand and did not nurse voluntarily for several hours. The mare was *stripped* (milked) of colostrum, which was fed to the foal by nasogastric tube.

Clinical Examination

The foal has a slightly low temperature and increased pulse and respiratory rates. The mucous membranes are tacky to the touch (dehydrated) and dark pink in color (indicating poor perfusion and/or poor oxygenation). Capillary refill time is prolonged (consistent with poor perfusion). The foal has a marked murmur similar to that heard with a patent *ductus arteriosus.* Peripheral pulses are decreased (weak), and distal extremities are cool. Gastrointestinal motility is decreased. The foal appears mature physically, but she is acting dysmature when she attempts to stand, nurse, or lie down. Blood studies reveal that the foal is not septic, but she is hypoxemic, has evidence of poor kidney function, and is acidotic.

Comment

Hypoxemic ischemic encephalopathy (HIE) occurs when a foal receives decreased oxygen for some time. This can occur before, during, or after foaling. With a dystocia, after the water breaks and while the foal is being pulled, the oxygen supply to the foal is decreased. The foal must rely on anaerobic metabolism during the period of low oxygen, which results in acidosis. Decreased oxygen also causes pulmonary blood vessels to constrict (*hypoxic pulmonary vasoconstriction,* discussed further in Chapter 46). The resulting increase in pulmonary vascular resistance causes blood to back up or accumulate in the pulmonary artery, right ventricle, and right atrium, and this increases the pressure in these chambers. If pressures in the right side of the heart exceed those in the left side, blood flow persists (from right to left) through the foramen ovale. (By contrast, when a normal foal begins to breathe, pressures in the right side fall below those in the left side of the heart, so the foramen ovale closes.) The blood that flows through the foramen ovale in this foal reaches the aorta without passing through the lungs, and therefore without being oxygenated at all (*right-to-left shunt*).

Treatment

The foal needs oxygen to reverse the hypoxic pulmonary vasoconstriction and the consequent high pressure in the right heart, persistent flow through the foramen ovale, and hypoxemia. Supplemental oxygen can be provided by *nasal insufflation* (tube placed in nasal cavity for delivery of oxygen). Additionally, the foal will be given drugs, such as dopamine, to increase cardiac contractility, cardiac output, and blood pressure. This treatment, in addition to intravenous fluids, will likely improve blood flow to the vital organs, including the brain and kidneys. Improved respiratory and renal function will reverse the acidosis. Foals with HIE often develop other complications, which need to be addressed as they arise.

Vet-ebooks.com

Practice Questions

1. The magnitude of pulsations in blood pressure (caused by the pulsatile ejection of blood from the heart) is greatest in the:
- **a.** Arteries
- **b.** Arterioles
- **c.** Capillaries
- **d.** Veins
- **e.** Venae cavae

2. Which of the following is a correct comparison between segments of the systemic circulation?
- **a.** The aorta and large arteries have a higher resistance to blood flow than the capillaries.
- **b.** The arterioles have a higher resistance to blood flow than the capillaries.
- **c.** The veins have a higher resistance to blood flow than the capillaries.
- **d.** The aorta and large arteries have a higher compliance than the veins.
- **e.** The aorta and large arteries contain a greater volume of blood than the veins and venae cavae.

3. If the heart suddenly stops beating in an otherwise normal dog, which of the following would be most likely to increase?
- **a.** Mean circulatory filling pressure
- **b.** Mean aortic pressure
- **c.** Vena caval pressure
- **d.** Perfusion pressure for the systemic circulation
- **e.** Perfusion pressure for the pulmonary circulation

4. Which of the following would cause mean aortic pressure to increase?
- **a.** Stroke volume increases from 30 to 40 mL, and heart rate decreases from 100 to 60 beats/min.
- **b.** Arterial compliance decreases.
- **c.** Cardiac output decreases.
- **d.** Arterioles throughout the body dilate.
- **e.** TPR increases.

5. If aortic compliance decreases while heart rate, cardiac output, and total peripheral resistance (TPR) remain unchanged:
- **a.** Pulse pressure will be unchanged.
- **b.** Pulse pressure will increase.
- **c.** Pulse pressure will decrease.
- **d.** One cannot know the effect on pulse pressure because stroke volume may have changed.
- **e.** One cannot know the effect on pulse pressure because mean aortic pressure may have changed.

6. The following measurements are made on a dog exhibiting distress following surgery: heart rate, 80 beats/min; stroke volume, 30 mL; mean aortic pressure, 96 mm Hg; mean pulmonary artery pressure, 26 mm Hg; left atrial pressure, 5 mm Hg; and right atrial pressure, 12 mm Hg. Which of the following is the best estimate for the TPR of this dog?
- **a.** 8.75 mm Hg/L/min
- **b.** 10.83 mm Hg/L/min
- **c.** 29.17 mm Hg/L/min
- **d.** 35.00 mm Hg/L/min
- **e.** 40.00 mm Hg/L/min

7. For the dog described in Question 6, the perfusion pressure for the systemic circulation is approximately _____ the perfusion pressure for the pulmonary circulation.
- **a.** Half as much as
- **b.** Equal to
- **c.** Twice as high as
- **d.** Four times higher than
- **e.** Nine times higher than

8. The blood flow through the brain of a resting dog would be decreased most by a 20% decrease in the:
- **a.** Mean arterial pressure
- **b.** Heart rate
- **c.** Systemic perfusion pressure
- **d.** Radius of arteries in the brain
- **e.** Radius of arterioles in the brain

Bibliography

Bonagura JD, Twedt DC. *Kirk's Current Veterinary Therapy XV*. Elsevier; 2014.

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.

Bowman DD. *Georgis' Parasitology for Veterinarians*. 10th ed. Elsevier; 2013.

Kumar V, Addas AK, Aster JC. *Robbins and Cotran Pathologic Basis of Disease*. 9th ed. Elsevier; 2014.

Knottenbelt DC, Holdstock N, Madigan JE. *Equine Neonatology: Medicine and Surgery*. New York: Saunders Elsevier; 2004.

Milnor WR. *Cardiovascular Physiology*. New York: Oxford University Press; 1990.

Mohrman DE, Heller LJ. *Cardiovascular Physiology*. 8th ed. McGraw-Hill; 2013.

Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.

Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2015.

Wilkins PA. *Equine Neonatology, An Issue of Veterinary Clinics of North America: Equine Practice (The Clinics: Veterinary Medicine)*. Elsevier; 2015.

Zachary JF, ed. *Pathologic Basis of Veterinary Disease Expert Consult*. 6th ed. Elsevier; 2016.

Vet-ebooks.com

23

Capillaries and Fluid Exchange

ROBERT B. STEPHENSON

KEY POINTS

1. Capillaries, the smallest blood vessels, are the sites for the exchange of water and solutes between the bloodstream and the interstitial fluid.
2. Lipid-soluble substances diffuse readily through capillary walls, whereas lipid-insoluble substances must pass through capillary pores or clefts.
3. Fick's law of diffusion is a simple mathematical accounting of the physical factors that affect the rate of diffusion.
4. Water moves across capillary walls both by diffusion (osmosis) and by bulk flow.
5. The Starling equation quantifies the interaction of oncotic and hydrostatic forces acting on water.
6. Several common physiological changes alter the normal balance of Starling forces and increase the filtration of water out of capillaries.
7. Edema is a clinically noticeable excess of interstitial fluid.

Capillaries, the Smallest Blood Vessels, Are the Sites for the Exchange of Water and Solutes Between the Bloodstream and the Interstitial Fluid

Because of their small size, the capillaries are sometimes called the *microcirculation.* They are also called the *exchange vessels,* because the exchange of water and solutes between the bloodstream and interstitial fluid takes place primarily across the walls of the capillaries. Each type of blood vessel in the body is structurally suited for its particular function, and the walls of the capillaries are especially well adapted for their exchange function.

Fig. 23.1 shows the contrasting features of the walls of the various types of blood vessels in the systemic circulation. The distinguishing feature of the walls of the aorta and large arteries is the presence of a large amount of elastic material along with smooth muscle. These vessels are called the *elastic vessels;* elasticity is necessary because the aorta and large arteries must distend with each pulsatile ejection of blood from the heart. The arterial walls are also strong and quite stiff (low compliance). There is no contradiction in saying that the arteries are both elastic and stiff. *Elasticity* denotes dispensability and an ability to return to the original shape after the distending force or pressure is removed. *Compliance* is a measure of how much force or pressure is required to achieve distention. The arteries are elastic, but a high pressure (systolic pressure) is required to distend them.

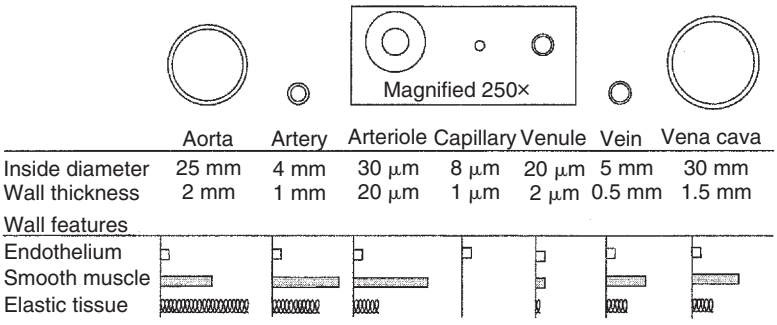
Small arteries, and particularly arterioles, have relatively thick walls with less elastic tissue and a predominance of smooth muscle, so they are called the *muscular vessels.* Contraction and relaxation of the smooth muscle enables these vessels to constrict or dilate, which varies their resistance to blood flow. The muscular vessels vary the total peripheral resistance and direct blood flow toward or away from particular organs or regions within an organ.

Capillaries are the smallest vessels, being about 8 μm in diameter and about 0.5 mm long. Capillaries are so small that red blood cells (7.5 μm in diameter) must squeeze through in single file. Capillary walls consist of a single layer of endothelial cells. The small diameter of the capillaries and the thinness of their walls facilitate the exchange of water and solutes between the blood within capillaries and the interstitial fluid immediately outside the capillaries.

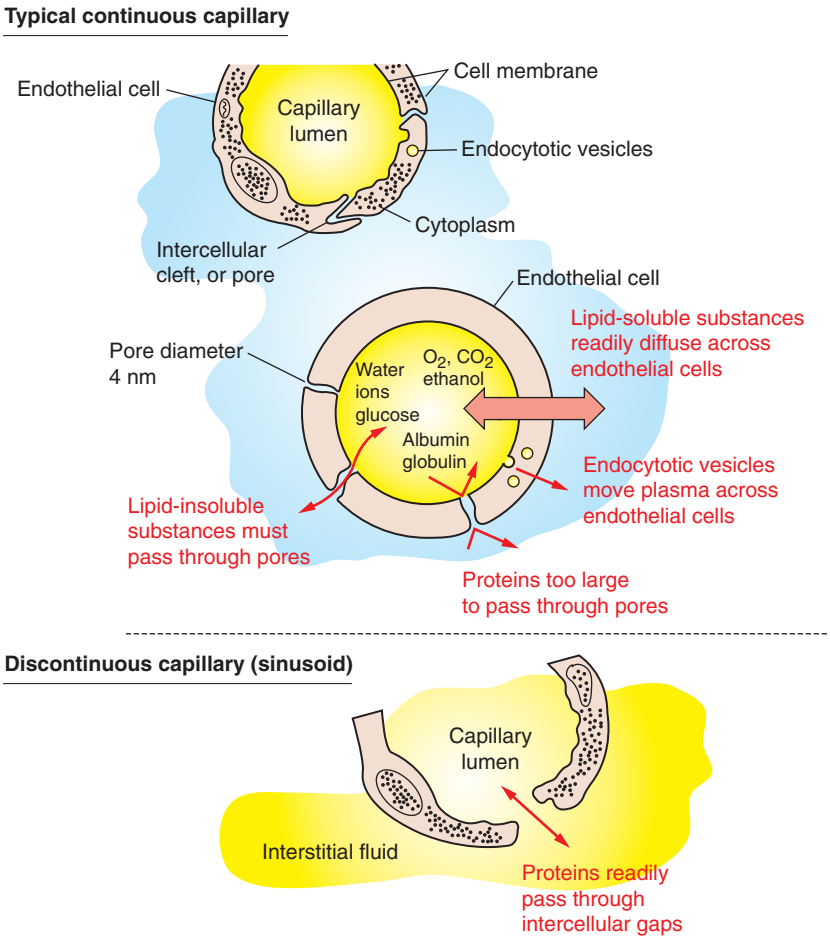
Venules and veins are larger than capillaries, and they have thicker walls. Venules and veins have both elastic tissue and smooth muscle in their walls. However, the walls of veins are not as thick or as muscular as the walls of arteries or arterioles. The primary role of veins is to serve as *reservoir vessels.* Veins are very compliant, and many veins in the body are normally in a state of partial collapse. Therefore veins can accommodate substantial changes in blood volume without much change in venous pressure.

Capillaries form a network (see Fig. 18.4). In most tissues, the capillary network is so dense that each cell of the tissue is within 100 μm (0.1 mm) of a capillary. However, not all the capillaries of a tissue carry blood flow at all times. In most tissues the arterioles alternate between constriction and dilation, so blood flow is periodically reduced or even stopped in capillaries. Also, in some tissues (e.g., intestinal circulation), tiny cuffs of smooth muscle encircle capillaries at the points where they branch off from arterioles. Contraction of these *precapillary sphincters* can reduce or stop the flow of blood in individual capillaries. When the metabolic rate of a tissue increases (and therefore its need for blood flow increases), the arterioles and precapillary sphincters still constrict periodically, but they spend more time in the dilated (relaxed) state. This increases the fraction of capillaries in which blood is flowing at any one time. At maximal metabolic rate (e.g., maximal exercise in a skeletal muscle), arterioles become maximally dilated, and blood flows through all the capillaries all the time. Sending blood flow to all the capillaries not only increases the total blood flow through a tissue but also minimizes the distance between each cell of the tissue and the nearest capillary carrying a bulk flow of blood. Both these effects speed up diffusional exchange between the capillary blood and the tissue cells.

Vet-ebooks.com



• **Fig. 23.1** Each type of blood vessel in the systemic circulation is specifically suited to its particular function by its size, wall thickness, and wall composition. In this drawing, each type of vessel is shown in cross-section. The drawings are to scale (note that the arteriole, capillary, and venule are magnified 250 times to make them visible). Also shown are the relative proportions of the three most important types of tissue found in blood vessel walls.



• **Fig. 23.2** Capillaries in cross-section. Typical continuous capillaries have small clefts, or pores, between endothelial cells (*top*). Water and small, lipid-insoluble solutes move between the capillary plasma (*yellow*) and the interstitial fluid (*blue*) through these pores (*center*). Plasma protein molecules are too large to pass through the pores. However, plasma proteins, along with other plasma constituents, are taken into endocytotic vesicles, which can deliver their contents into the interstitial fluid via exocytosis, although this is a very slow process. In contrast, lipid-soluble substances can diffuse directly, and extremely quickly, through the capillary endothelial cells. The size of the clefts between endothelial cells varies greatly from tissue to tissue, with the smallest being in brain capillaries and the largest being in the discontinuous capillaries, or sinusoids, such as those in the liver (*bottom*).

Lipid-Soluble Substances Diffuse Readily Through Capillary Walls, Whereas Lipid-Insoluble Substances Must Pass Through Capillary Pores or Clefts

The rate of diffusional exchange between capillary blood and the surrounding interstitial fluid depends both on the properties of the substances being exchanged and the features of the capillary wall. Small, lipid-soluble substances (e.g., dissolved oxygen and carbon dioxide, fatty acids, ethanol, and some hormones) readily dissolve in the cell membranes of the endothelial cells that form the capillary walls. Such lipid-soluble substances can diffuse very rapidly through the endothelial cells from blood to interstitial fluid, or vice-versa. In contrast, lipid-insoluble substances (e.g., ions, glucose, and amino acids) do not dissolve in cell membranes and so cannot diffuse through the endothelial cells. Instead, such substances must pass through the *pores* or *clefts* that exist between the endothelial cells (Fig. 23.2). These pores (clefts) create extremely tiny, water-filled channels between the capillary blood and the interstitial fluid. In a typical capillary, the pores constitute only about 1% of the total wall surface area. Therefore diffusional movement of lipid-insoluble substances across capillary walls is much slower than the movement of lipid-soluble substances, because the lipid-insoluble substances are restricted to passage through the capillary pores.

The most widespread type of capillary in the body is called continuous because the endothelial cells form a continuous tube, except for the tiny, water-filled pores between the endothelial cells. In typical continuous capillaries, the diameter of the pores is about

Vet-ebooks.com

4 nm, which is large enough to permit the passage of water and of the small lipid-insoluble solutes in plasma and interstitial fluid (e.g., ions, glucose, and amino acids). The plasma protein molecules, however, are a little too large to pass through pores of this size. Blood cells, of course, are far too large to pass through such small openings (see Fig. 18.7).

The main route for the delivery of plasma proteins into the interstitial fluid is through the three-step process of *transcytosis.* The first step is *pinocytosis* (a form of *endocytosis*), which involves the invagination of the capillary endothelial cell membrane to form an intracellular vesicle that contains plasma, including plasma proteins (see Fig. 23.2). Second, via random motion, some of these vesicles cross the capillary endothelial cell from the side facing the bloodstream to the side facing the interstitial fluid. In the third step, these vesicles fuse with the endothelial cell membrane on the interstitial fluid side; the vesicles discharge their contents into the interstitial space. This third step is called *exocytosis.* The delivery of plasma constituents into the interstitial fluid by transcytosis is very slow compared with the passage of small, lipid-insoluble substances through capillary pores, and extremely slow compared with the diffusion of lipid-soluble substances through endothelial cells.

The size of the capillary pores or clefts varies from tissue to tissue. Two extremes are found in the brain and the liver. In brain capillaries, the junctions between adjacent endothelial cells are so tight that ions, glucose, and amino acids cannot pass through. Yet brain neurons require glucose to carry out their normal metabolism. Glucose is moved across the brain capillary endothelial cells by means of specialized transport proteins (*glucose carriers*) embedded in the cell membranes of the endothelial cells. The energy to drive this *facilitated diffusion* comes from the glucose concentration difference between the blood plasma and the brain interstitial fluid. The endothelial cell membranes of brain capillaries also contain specialized carrier proteins that can mediate *active transport* of particular ions and amino acids from the blood to the brain interstitial fluid or vice versa. The brain capillaries are unique in being able to carry out such selective, unidirectional active transport. The tight junctions between endothelial cells in brain capillaries compose part of a barrier between the bloodstream and the brain tissue called the *blood-brain barrier* (also discussed in Chapter 15). One function of the blood-brain barrier is to allow regulation of the concentration of particular ions and amino acids in the brain interstitial fluid; another function is to protect brain neurons from exposure to toxic substances that may be in the blood.

In the liver, the clefts between capillary endothelial cells are exceptionally large, typically exceeding 100 nm in width (see Fig. 23.2, bottom). Even plasma proteins such as albumin and globulin can readily pass through such large clefts. These vessels are called *discontinuous capillaries* (or *sinusoids*). Large gaps between endothelial cells are an appropriate feature for capillaries in the liver. Plasma proteins are produced by liver cells (*hepatocytes*), and the large gaps permit the newly synthesized protein molecules to enter the bloodstream. The large gaps are also appropriate for the role of the liver in detoxification. Some toxins become bound to plasma proteins in the bloodstream, then are removed from the blood by the liver and chemically changed into less toxic substances. Discontinuous (sinusoidal) capillaries are also found in the spleen and bone marrow.

Fenestrated capillaries ("capillaries with windows") present an additional variation on capillary pores. Fenestrae are holes or perforations through (not between) endothelial cells. Fenestrae are typically 50 to 80 nm in diameter, which is larger than the intercellular clefts of typical continuous capillaries but smaller than the clefts of discontinuous capillaries. Very fine diaphragms span most fenestrae. These diaphragms block the passage of plasma proteins, but permit the passage of water and both lipid-soluble and lipid-insoluble substances. The diaphragms are analogous to window screens, which let in air and dust but keep mosquitos out. It is thought that fenestrae are formed when endocytotic and exocytotic vesicles line up and merge, thus creating a temporary water-filled channel through an endothelial cell. Fenestrated capillaries are characteristically found in places where large amounts of fluid and solutes must pass into or out of capillaries (e.g., gastrointestinal tract, endocrine glands, and kidneys).

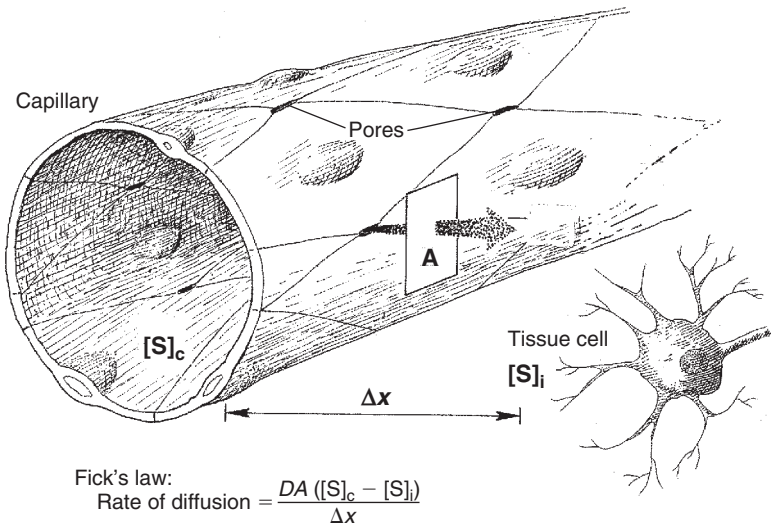
Fick's Law of Diffusion Is a Simple Mathematical Accounting of the Physical Factors That Affect the Rate of Diffusion

In general, movement of solutes between the bloodstream and interstitial fluid occurs via simple, passive diffusion. Salient exceptions (already discussed) occur in the brain, where glucose moves through capillary endothelial cells via facilitated diffusion, and particular ions and amino acids are actively transported. The rate of passive diffusional exchange between capillary blood and interstitial fluid is affected by the distance involved, the size of the capillary pores (or fenestrae, when present), and the properties of the diffusing substance (i.e., lipid soluble vs. lipid insoluble). The German physiologist Adolph Fick incorporated all these factors into an equation: *Fick's law of diffusion.* Fig. 23.3 shows how Fick's law applies to the diffusional exchange between capillary fluid and interstitial fluid. The rate of passive diffusion of any substance (S) depends, first, on the *concentration difference,* that is, the difference between the concentration of the substance in capillary fluid and its concentration in interstitial fluid. Diffusion is driven by this concentration difference, and diffusion always proceeds from the area of higher concentration toward the area of lower concentration. Next, the rate of diffusion is determined by the *area available for diffusion,* the term A in the equation. For lipid-soluble substances, this area is equivalent to the total surface area of the capillaries. For lipid-insoluble substances, this area is much smaller, equal to the area of the pores (or clefts) between capillary endothelial cells (plus the area of fenestrae, when present).

The term Δx in the equation represents the *distance* over which diffusion must occur. Functionally, Δx equals the distance from a tissue cell to the nearest capillary carrying blood by bulk flow (see Fig. 23.3). The greater the distance from the tissue cells to the capillaries, the slower is the rate of diffusional exchange of substances between that cell and the capillary blood; therefore Δx appears in the denominator in the equation.

The term D in the equation is a *diffusion coefficient.* The value of D increases with temperature because diffusion depends on the random (Brownian) motion of particles in solution, and the velocity of Brownian motion increases with temperature. D also depends on the molecular weight of the diffusing substance and on its solubility. For example, D for carbon dioxide is about 20 times greater than D for oxygen, because carbon dioxide is much more soluble in body fluid than oxygen. As a result, carbon dioxide diffuses much more rapidly than does oxygen for a given concentration difference, area, and diffusion distance. This difference in diffusion speed is inconsequential under normal physiological conditions. In certain disease states, however, the area available

Vet-ebooks.com



• **Fig. 23.3** According to Fick's law, four factors affect the rate of diffusion of a particular substance S from the capillary plasma to the interstitial fluid that surrounds a tissue cell: first, $[S]_c - [S]_i$, the difference between the concentration of S in the capillary plasma and in interstitial fluid; second, *A,* the area available for diffusion; third, Δx, the distance involved; and fourth, *D,* the diffusion coefficient for substance S.

for diffusion decreases, and the diffusion distance increases. Under these conditions, the delivery of oxygen to the metabolizing cells of a tissue generally becomes critically impaired before the removal of carbon dioxide from the cells becomes inadequate.

Several of the factors that affect the rate of diffusion are physiologically adjustable. For example, in skeletal muscle at rest, the arterioles cycle between closed and partially opened. Consequently, at any one moment, blood flows through only about one-fourth of the skeletal muscle capillaries. Blood simply sits still in the remainder of them. Nevertheless, this low and "part-time" blood flow through capillaries is adequate to deliver oxygen and nutrients to the resting skeletal muscle cells and to remove the small amounts of carbon dioxide and other waste products being produced by those cells. In contrast, during exercise, the metabolic rate of the skeletal muscle cells increases severalfold, as does their need for blood flow. During exercise, skeletal muscle arterioles dilate, and more and more of them remain open on a "full-time" basis as the level of exercise increases. Consequently, blood flow through the capillaries increases dramatically and becomes more continuous.

Arteriolar vasodilation acts in three ways to speed the delivery of oxygen and metabolic substrates to the exercising muscle cells and to facilitate the removal of carbon dioxide and other metabolic waste products. First, when more capillaries carry blood, the area available for diffusion (*A* in Fick's diffusion equation) is increased. Second, because more capillaries carry blood, the distance between each exercising skeletal muscle cell and the nearest open capillary (Δx in the diffusion equation) is decreased. Third, the driving force for diffusion of oxygen (the oxygen concentration difference between the capillary blood and the interstitial fluid) is increased. The concentration difference is increased because (1) the greater blood flow brings more freshly oxygenated blood into the capillaries, and (2) the rapid utilization of oxygen by the exercising skeletal muscle cells decreases the concentration of oxygen within these cells and therefore within the surrounding interstitial fluid.

The same factors that increase the rate of oxygen diffusion during exercise also increase the rate of delivery of glucose and other nutrients. Furthermore, the same factors act to increase the rate at which carbon dioxide and other metabolic products are removed from the tissue cells and into the bloodstream. In the case of carbon dioxide and other metabolic products, the concentration is higher in the cells than in the capillary plasma, so diffusional movement is from the metabolizing cells toward the bloodstream.

Water Moves Across Capillary Walls Both by Diffusion (Osmosis) and by Bulk Flow

The exchange of water between the capillary plasma and the interstitial fluid merits special consideration for two reasons. First, the forces that govern water movement are more complicated than the simple diffusive forces that affect solute movement. Second, a particular imbalance in these forces causes an excessive amount of water to accumulate in the interstitial space, which leads to the important clinical sign, *edema.*

As the preceding discussion emphasized, solutes such as oxygen, carbon dioxide, glucose, electrolytes, and fatty acids move between the capillary plasma and the interstitial fluid by diffusion. Water also moves by diffusion; the diffusional movement of water is called *osmosis.* The physical prerequisites for osmosis are (1) the presence of a *semipermeable membrane* (a membrane permeable to water but not to specific solutes) and (2) a difference in the total concentration of the *impermeable solutes* on the two sides of the membrane.

The capillary wall constitutes a semipermeable membrane. Water can readily pass through capillary pores; however, the pores in continuous capillaries are too small to permit the passage of plasma proteins. As a consequence, the concentration of plasma proteins is normally much higher in the capillary plasma than in the interstitial fluid. Protein concentration is typically 7 grams per deciliter (g/dL) within the capillary plasma but only 0.2 g/dL in the interstitial fluid. These dissimilar protein concentrations create an osmotic imbalance. As a consequence, water molecules tend to move by osmosis from the interstitial fluid into the capillary blood plasma. (Remember that water moves by osmosis toward the side of the semipermeable membrane with the higher concentration of impermeable solute.)

The tendency for water to move by diffusion is quantified as *osmotic pressure* (see Chapter 1). The normal osmotic pressure created by the proteins in the plasma is 25 mm Hg; that is, the osmotic effect of the plasma proteins is equivalent to a hydrostatic pressure of 25 mm Hg driving water into the capillaries. The osmotic pressure created by the plasma proteins is also called *protein oncotic pressure* or *oncotic pressure* or *colloid osmotic pressure.* (The term *colloid* is used because the plasma proteins are not in a true solution but rather in a colloidal suspension.)

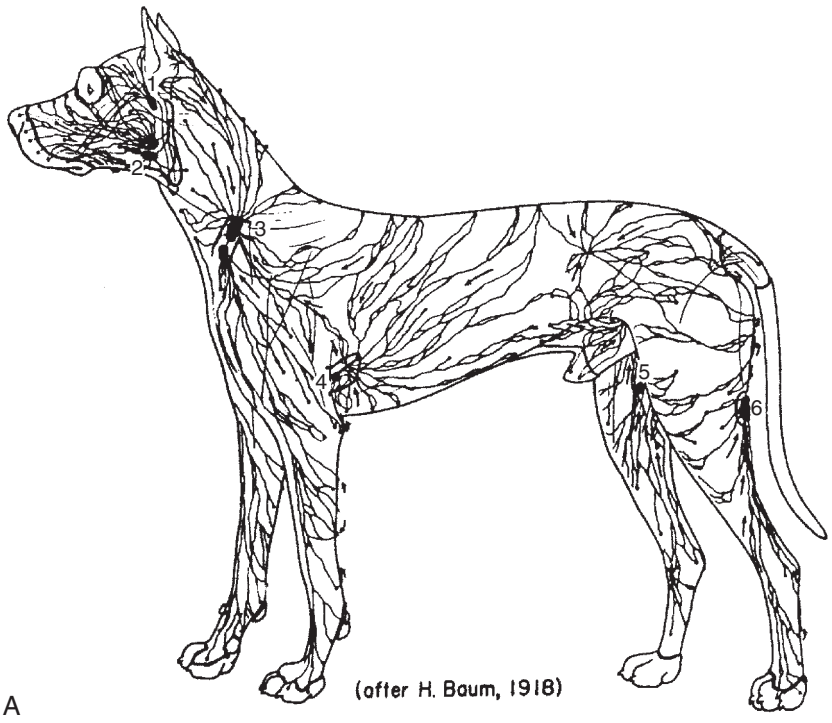
Vet-ebooks.com

The plasma proteins in the interstitial fluid also exert an osmotic effect. However, because the concentration of plasma proteins in interstitial fluid is normally quite low, the oncotic pressure created in the interstitial fluid by these proteins is normally only about 1 mm Hg. The imbalance of oncotic pressures (higher in the capillary fluid than in the interstitial fluid) creates the net driving force for the diffusion (osmotic movement) of water from the interstitial fluid into the capillaries.

The movement of water into a capillary is called *reabsorption.* The movement of water in the opposite direction, from the capillary plasma into the interstitial fluid, is called *filtration.* The *oncotic pressure difference* normally favors reabsorption. The net driving force for osmosis is calculated by subtracting the oncotic pressure of interstitial fluid from the oncotic pressure of capillary plasma (e.g., 25 mm Hg − 1 mm Hg = 24 mm Hg).

In addition to being affected by diffusional (osmotic) forces, water responds to hydrostatic pressure differences across the capillary wall. Hydrostatic pressure differences cause water to move by bulk flow; in this case the bulk flow occurs through the capillary pores. The hydrostatic pressure within the capillaries (capillary blood pressure) is higher at the arteriolar end of capillaries than at the venous end (see Fig. 22.1). However, a representative average capillary hydrostatic pressure would be about 18 mm Hg (i.e., about 18 mm Hg higher than atmospheric pressure). In contrast, interstitial fluid hydrostatic pressure is normally about −7 mm Hg. The negative sign simply means that interstitial fluid pressure is *less* (although only slightly less) than atmospheric pressure. The negative interstitial fluid pressure (−7 mm Hg) together with the positive capillary hydrostatic pressure (18 mm Hg) creates a net hydrostatic pressure difference of 25 mm Hg across the wall of a typical capillary (i.e., 18 mm Hg − (−7 mm Hg) = 25 mm Hg). Because hydrostatic pressure is higher in the capillary plasma than in the interstitial fluid, the net hydrostatic pressure difference tends to force water out of the capillaries and into the interstitial spaces; that is, the *hydrostatic pressure difference* favors filtration.

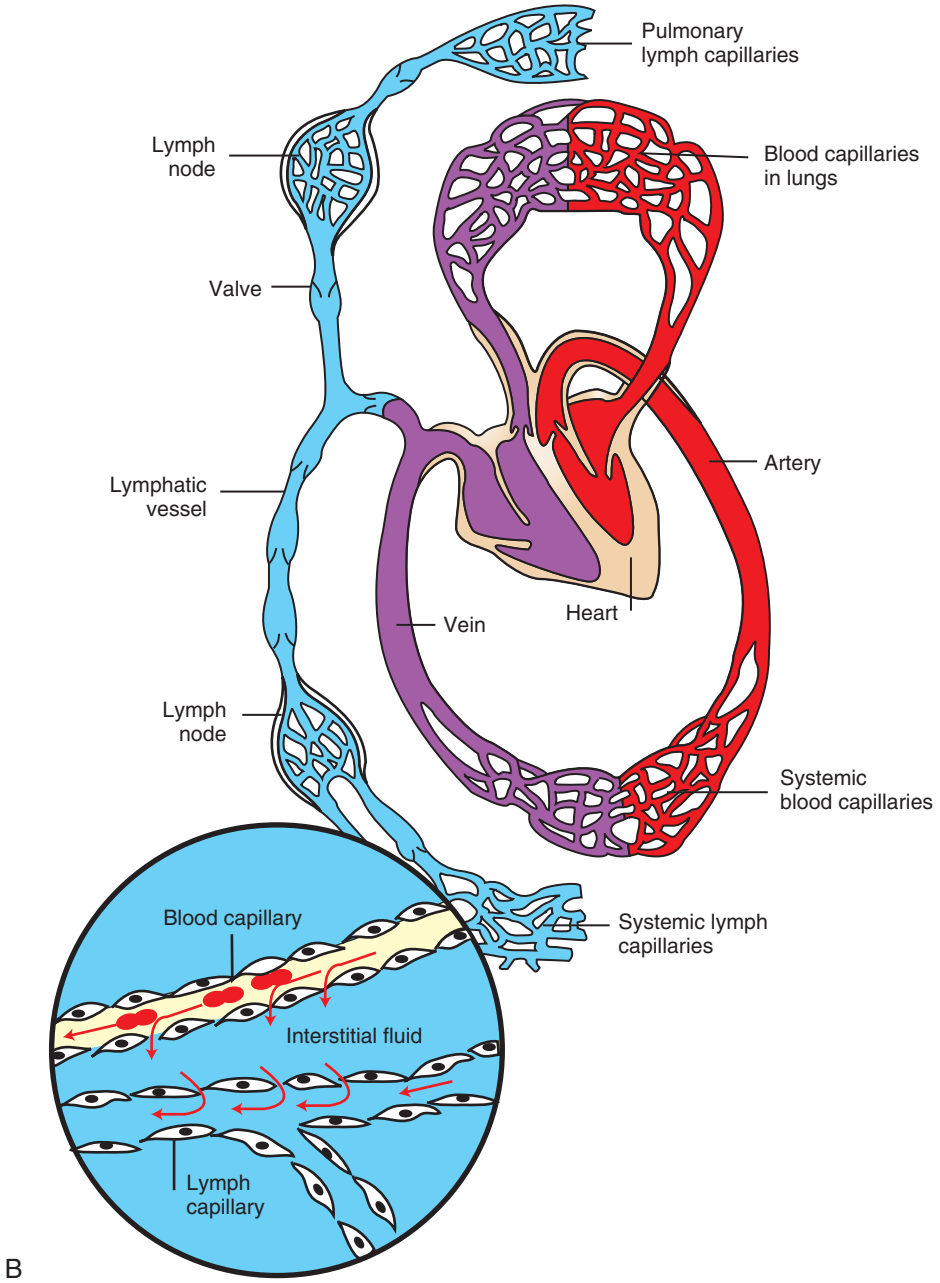
In most capillaries of the systemic circulation, the hydrostatic pressure difference (which favors filtration) almost balances the oncotic pressure difference (which favors reabsorption). However, the balance is rarely perfect. Usually, the hydrostatic pressure difference slightly exceeds the oncotic pressure difference, so there is a small net filtration of water out of the capillaries. This water would simply accumulate in the interstitial spaces and cause swelling there if not for the lymph vessels, which collect excess interstitial fluid and return it to the bloodstream through the subclavian veins (Fig. 23.4).



A

• **Fig. 23.4** (A) Anatomical and (B) schematic overviews of the lymphatic system. The lymphatic vessels collect excess interstitial fluid from tissues throughout the body (including the lungs) and carry it to the subclavian veins, where the lymph reenters the bloodstream. Lymph moves through lymph vessels via bulk flow. Lymph flow is promoted by the massaging action exerted on lymph vessels by contraction and relaxation of skeletal muscles and (in the lungs) by the pressure variations accompanying inspiration and expiration. The lymph vessels contain one-way valves, which prevent the backflow of lymph. Thus massaging actions propel lymph in one direction only: toward the subclavian vein. In addition, some lymph vessels have smooth muscle in their walls; the alternating contraction and relaxation of this smooth muscle also propels lymph flow toward the subclavian veins. The *numbers* in A identify the major lymph nodes. The *magnified inset* in B depicts the typical, net filtration of fluid out of a blood capillary and into the interstitial space. Under normal conditions, this excess interstitial fluid is collected and carried away by the lymph capillaries. Plasma is indicated in *yellow,* interstitial fluid and lymph in *blue.* Three red blood cells are depicted in the blood capillary. (*A* from Getty R. *Sisson and Grossman's the Anatomy of the Domestic Animal,* vol 2. Philadelphia: Saunders; 1975.) *Continued*

Vet-ebooks.com



• **Fig. 23.4, cont'd**

As already mentioned, capillary hydrostatic pressure and interstitial fluid hydrostatic pressure are, by convention, measured relative to atmospheric pressure. Thus to say that interstitial pressure is normally "negative" does not imply that a vacuum exists in the interstitium but only that the interstitial pressure is slightly below atmospheric pressure. If all the interstitial spaces of the body had a hydrostatic pressure higher than atmospheric pressure, all parts of the body would bulge outward. The subatmospheric interstitial fluid pressure probably accounts for the fact that the skin normally stays snug against the underlying tissue and that some body surfaces normally have a concave shape (e.g., axillary space, orbits of the eyes).

The Starling Equation Quantifies the Interaction of Oncotic and Hydrostatic Forces Acting on Water

The following equation expresses mathematically the interaction between osmotic pressures and hydrostatic pressures in determination of the net force (net pressure) acting on water:

$$\text{Net pressure} = (P_c - P_i) - (\pi_c - \pi_i)$$

where P_c is capillary hydrostatic pressure, P_i is interstitial fluid hydrostatic pressure, π_c is capillary plasma oncotic pressure, and

Vet-ebooks.com

π_i is interstitial fluid oncotic pressure. For the systemic circulation, the nominal values for these pressures (as already mentioned) are as follows:

P_c = 18 mm Hg
P_i = −7 mm Hg
π_c = 25 mm Hg
π_i = 1 mm Hg

The solution of this equation, with nominal values inserted for each term, is:

$$\begin{aligned}\text{Net pressure} &= [18\text{ mm Hg} - (-7\text{ mm Hg})] \\ &\quad - [(25\text{ mm Hg} - 1\text{ mm Hg})] \\ &= +1\text{ mm Hg}\end{aligned}$$

A positive net pressure favors filtration (a negative net pressure would indicate that reabsorption is favored). The small magnitude of the typical net pressure (1 mm Hg) indicates that the hydrostatic and osmotic forces that affect water are almost in balance (i.e., there is only a slight tendency for filtration). The quantitative analysis of how oncotic and hydrostatic pressures affect water movement across capillary walls was first derived by Ernest Henry Starling (the same scientist for whom Starling's law of the heart is named). Therefore the oncotic and hydrostatic pressures that act on water are often called *Starling forces.* Furthermore, the tendency for the net oncotic effect to be closely balanced by the net hydrostatic effect is often referred to as the *balance of Starling forces.* Starling realized that the actual rate of water movement across capillary walls (*transcapillary water flux*) is affected both by the magnitude of the imbalance between hydrostatic and oncotic forces and by the permeability of the capillary wall to water. These ideas are expressed in the following equation, which indicates that the transcapillary water flux is equal to the permeability of the capillary wall (given as the filtration coefficient K_f) multiplied by the net pressure:

$$\text{Transcapillary water flux} = K_f[(P_c - P_i) - (\pi_c - \pi_i)]$$

Examination of this equation reveals that the tendency for the filtration of water out of capillaries can be enhanced by (1) increasing the hydrostatic pressure difference between capillary blood and interstitial fluid, (2) decreasing the osmotic tendency for water to be reabsorbed, or (3) increasing the permeability of the capillary to water (i.e., increasing the filtration coefficient).

Several Common Physiological Changes Alter the Normal Balance of Starling Forces and Increase the Filtration of Water Out of Capillaries

An increase in capillary hydrostatic pressure (P_c) favors a greater filtration of water. Capillary hydrostatic pressure can be increased by an increase in arterial blood pressure or by a decrease in arteriolar resistance. An increase in arterial pressure causes more blood pressure to be transmitted down through the arterioles and into the capillaries. Likewise, a decrease in arteriolar resistance (e.g., a dilation of the arterioles) allows a greater portion of the arterial pressure to be transmitted into the capillaries. Capillary hydrostatic pressure can also be increased by a "backing up" (or "damming up") of venous blood. For example, an increase in central venous pressure causes blood to accumulate in the systemic veins and capillaries and raises capillary pressure. An obstruction to venous outflow (e.g., too tight a dressing on a limb) also causes blood to back up in the veins and capillaries of the limb, which increases capillary hydrostatic pressure.

The primary determinant of interstitial fluid hydrostatic pressure is the volume of fluid present in the interstitial space. An accumulation of interstitial fluid increases interstitial hydrostatic pressure. Removal of interstitial fluid decreases the pressure. As stated earlier, interstitial fluid hydrostatic pressure is usually slightly subatmospheric (e.g., −7 mm Hg). When interstitial fluid hydrostatic pressure rises above atmospheric pressure, the accumulation of interstitial fluid becomes clinically noticeable as a swelling or *edema.*

The net oncotic pressure depends on the concentrations of proteins in the capillary plasma and in the interstitial fluid. As already stated, the normal protein concentration in plasma is 7 g/dL, which results in a plasma oncotic pressure of 25 mm Hg. Any increase above normal in the concentration of proteins in the capillary plasma (*hyperproteinemia*) raises the plasma oncotic pressure above normal, and vice versa.

Similarly, changes in the interstitial protein concentration alter interstitial fluid oncotic pressure. In most organs of the systemic circulation protein molecules do not readily pass through the capillary pores or clefts. As already described, the main route for the delivery of plasma proteins into the interstitial fluid is through the process of *transcytosis.* An increase in the rate of transcytosis increases the delivery of plasma proteins into the interstitial space and therefore increases interstitial fluid oncotic pressure. In addition, abnormal circumstances (e.g., tissue inflammation) can cause the capillary pores to open widely enough to allow plasma proteins to pass through and into the interstitial fluid.

Excess interstitial fluid and the plasma proteins it contains are removed from the interstitial space through lymph flow. The smallest lymphatic vessels (*lymphatic capillaries*) are structured much like blood capillaries. One salient difference is that the clefts between the endothelial cells of lymphatic capillaries are large enough to readily accommodate the passage of plasma protein molecules. Therefore when excess interstitial fluid flows through these clefts and enters lymph capillaries, any plasma proteins present in the interstitial fluid are also carried into the lymph capillaries. The lymphatic capillary fluid, containing these plasma proteins, then flows into progressively larger lymphatic vessels, ending up in the thorax, where the lymphatic fluid reenters the bloodstream via the subclavian veins (see Fig. 23.4).

The role of lymphatic flow in counteracting the accumulation of excessive interstitial fluid is especially important in the lungs. Lung capillaries are more permeable to plasma proteins than are most capillaries in the systemic circulation. As a result, the oncotic pressure of interstitial fluid in the lungs is normally rather high (nominally 18 mm Hg). Capillary hydrostatic pressure in the lungs is generally about 12 mm Hg. (This value is lower than the capillary hydrostatic pressure in systemic capillaries because pulmonary arterial pressure is so much lower than systemic arterial pressure.) Interstitial hydrostatic pressure in the lungs is generally about −4 mm Hg (the same as intrapleural pressure). Summation of these Starling forces for lung capillaries yields the following:

$$\begin{aligned}\text{Net pressure} &= [12\text{ mm Hg} - (-4\text{ mm Hg})] \\ &\quad - [(25\text{ mm Hg} - 18\text{ mm Hg})] \\ &= +9\text{ mm Hg}\end{aligned}$$

A net pressure of + 9 mm Hg indicates that there is a substantial driving force for filtration of fluid out of the capillaries and into the lung interstitial spaces. The lung interstitial spaces would fill

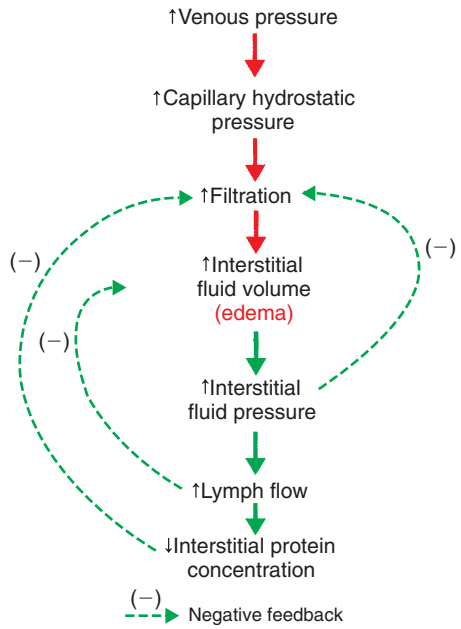
Vet-ebooks.com

rapidly with water, and pulmonary edema would develop, were it not for the well-developed system of lymph vessels in the lungs. These vessels continuously remove interstitial fluid and prevent its excessive accumulation.

Edema Is a Clinically Noticeable Excess of Interstitial Fluid

Edema is a common clinical problem. Edema results either from excessive filtration of fluid out of capillaries or from depressed removal of excess interstitial fluid by the lymph system. An increase in venous blood pressure is one common cause of excessive filtration. Increased venous pressure can result from the application of a too-tight dressing on the extremity of an animal. The resulting constriction of the veins impedes the outflow of venous blood from the limb. Blood backs up in the limb veins, which increases venous pressure. Blood then backs up in the capillaries and increases capillary hydrostatic pressure. As shown in Fig. 23.5, the increase in capillary hydrostatic pressure leads to excessive filtration of capillary fluid into the interstitial space. When this accumulation of fluids becomes clinically noticeable, the patient is said to exhibit edema.

Other causes of increased venous pressure are severe pulmonic stenosis (see Clinical Correlation for Chapter 21) and severe heartworm disease (see Clinical Correlation for Chapter 22). In these conditions, an excessive volume of blood accumulates in the right atrium and systemic veins. The resulting increase in venous pressure causes blood to back up in the systemic capillaries and this increases capillary hydrostatic pressure and leads to edema, as shown in Fig. 23.5.



• **Fig. 23.5** Increase in venous pressure leads to increase in interstitial fluid volume (edema). The *dashed lines (negative feedback)* indicate the counteracting effects of the three safety factors against edema. First, an increase in interstitial fluid hydrostatic pressure reduces the rate of filtration back toward normal. Second, an increase in lymph flow reduces interstitial fluid volume back toward normal. Third, a decrease in interstitial fluid protein concentration reduces the rate of filtration back toward normal.

Whatever the cause of an increase in venous pressure, three factors (*safety factors*) limit the degree of the resulting edema. All three safety factors depend on an increased interstitial fluid volume leading to an increase in interstitial fluid hydrostatic pressure. The first safety factor is that the increased interstitial fluid pressure acts directly to oppose or limit filtration. Interstitial fluid pressure does not need to rise above capillary hydrostatic pressure to limit edema. Any increase in interstitial fluid pressure (e.g., from a normal value of −7 to + 2 mm Hg) helps change the net balance of the Starling forces in the direction of reducing filtration.

The second safety factor against edema is that increased interstitial fluid pressure promotes the entry of interstitial fluid into lymphatic capillaries. Increased lymph flow removes edema fluid from the tissue and therefore helps limit the degree of edema.

The third safety factor is an indirect consequence of increased lymph flow. Recall that interstitial fluid normally contains a small amount of plasma protein, usually the result of transcytosis. This protein exerts a small but significant oncotic pressure that favors filtration. Under the circumstance of increased capillary hydrostatic pressure, the increased capillary filtration delivers fluid into the interstitial space that is relatively free of proteins. Meanwhile, the elevated lymph flow carries away not only interstitial fluid but also the proteins that it contains. The combination of increased filtration adding fluid to the interstitial space, together with increased lymph flow removing interstitial fluid and the protein it contains, leads to "flushing" of proteins out of the interstitial fluid. The resulting decrease in interstitial fluid oncotic pressure helps reduce the excess filtration back toward normal.

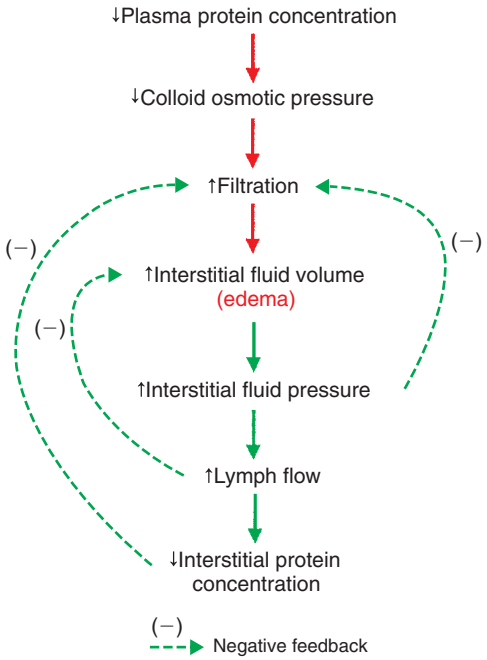
To summarize, an increase in venous pressure leads to an increase in capillary hydrostatic pressure, which increases filtration and edema develops. Three safety factors then come into play to reduce filtration back toward normal and to limit the degree of edema. A steady-state degree of edema is eventually reached, in which interstitial fluid is removed by lymph vessels as fast as it is filtered.

The *systemic edema* that results from an increase in systemic venous pressure is often most noticeable in *dependent regions* (regions of the body below heart level), where the downward pull of gravity augments the elevation in venous pressure. In large quadrupeds, dependent regions typically include the abdomen and extremities. When edema develops in the abdominal organs, excess interstitial fluid tends to ooze out of the edematous tissues and accumulate in the peritoneal space. Excessive fluid in the peritoneum is called *ascites.*

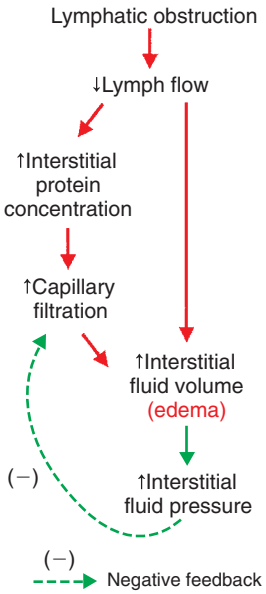
Marked systemic edema and ascites is common in patients with right ventricular heart failure. By contrast, failure of the left ventricle leads to *pulmonary edema.* Ineffective pumping by the left ventricle results in increased blood volume and increased pressure in the left atrium and pulmonary veins. This elevated pressure extends back into the pulmonary capillaries, which increases capillary filtration in the lung tissue. In severe cases of pulmonary edema, some of the excess interstitial fluid oozes into the alveoli and bronchial airways. Such a patient typically coughs up a frothy fluid. Excess edema fluid may also ooze into the intrapleural space, which is called *pleural effusion.* The consequences of heart failure are discussed more fully in Chapter 26.

A decreased plasma protein concentration (*hypoproteinemia*) is another common cause of edema (Fig. 23.6). One cause of hypoproteinemia is a decrease in the rate of plasma protein production by the liver. This occurs in malnutrition and leads to the clinical syndrome of *kwashiorkor.* Victims of kwashiorkor typically look emaciated, except that the abdomen is grossly distended by

Vet-ebooks.com



• **Fig. 23.6** Decrease in plasma protein concentration leads to edema, but the degree of edema is limited by the same three safety factors as shown in Fig. 23.5.



• **Fig. 23.7** Lymphatic obstruction leads to edema. Lymphedema is clinically troublesome because only one of the normal three safety factors is operative to limit the degree of edema.

edema and ascites. Another cause of abnormally low plasma protein concentration is an increase in the rate of loss of plasma proteins from the body. Protein loss occurs in kidney disease. For example, in *nephrotic syndrome,* the kidney glomerular capillaries become permeable to plasma proteins. Plasma proteins leave the bloodstream and enter the urinary tubules (nephrons) of the kidney. A chronic loss of proteins in the urine reduces the plasma protein concentration. Therefore the presence of substantial amounts of plasma protein in the urine is an alarming clinical sign.

Severe burns also cause the loss of plasma proteins from the body. The capillaries of burned skin become very permeable to both fluid and proteins. Substantial amounts of plasma can leave the body through these damaged capillaries. The presence of plasma proteins in the fluid weeping from a burn site accounts for the typical yellow color of that fluid. If the water and electrolytes lost through burns are replaced through drinking or an intravenous administration of fluids, and if the plasma proteins are not also replaced, the plasma protein concentration in the blood decreases.

Whether it results from decreased production or increased loss, hypoproteinemia leads to a decrease in plasma colloid osmotic pressure. This alters the balance of the Starling forces in a direction that favors excessive filtration of fluid from the capillaries (see Fig. 23.6). Interstitial fluid accumulates and edema is noticed. However, the same three safety factors that limit edema in the case of increased venous pressure (see Fig. 23.5) also operate in the case of decreased plasma protein concentration. The degree of edema is limited by: (1) an increased interstitial fluid pressure, (2) an increased lymph flow, and (3) a decreased interstitial protein concentration.

Another cause of edema is lymphatic obstruction. Clinically, this situation is called *lymphedema.* The passage of lymph through lymph nodes can be impaired by inflammation of the nodal tissue or cancerous tumors growing within the nodes. Also, in certain parasitical diseases, microfilariae lodge in the lymph nodes and obstruct lymph flow. Filarial parasites cause the pronounced edema seen in cases of *elephantiasis.* Lymphedema also occurs as a secondary consequence of surgical procedures that damage lymph nodes. A common example of this in human medicine is the edema of the arm that follows radical mastectomy. The removal of axillary lymph nodes during radical mastectomy creates scar tissue that impairs lymphatic drainage from the arm.

Fig. 23.7 traces the causes of edema after lymphatic obstruction and shows why lymphedema is clinically so troublesome. Lymphatic obstruction decreases lymph flow. Interstitial fluid accumulates in the tissues, instead of being removed by the lymph, and edema results. Decreased lymph flow also leads, indirectly, to an increase in interstitial protein concentration. As previously mentioned, plasma proteins continually enter the interstitial fluid, via the slow process of transcytosis. Lymph flow carries away not only interstitial fluid but also the proteins that it contains. The combination of filtration adding fluid to the interstitial space, together with lymph flow removing both fluid and protein, leads to "flushing" of proteins out of the interstitial fluid, which keeps interstitial protein concentration very low. Lymphatic obstruction curtails this flushing of proteins from the interstitial fluid.

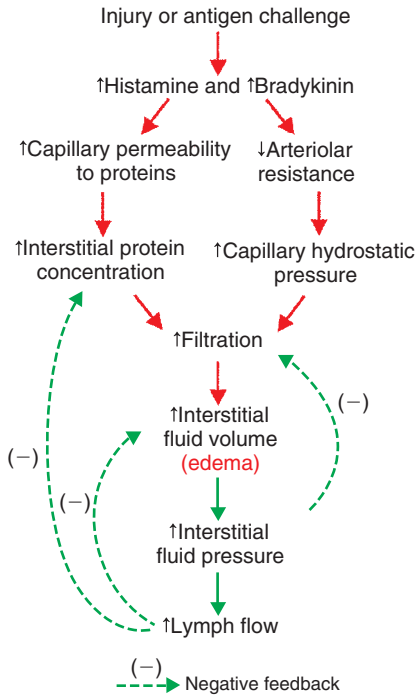
With lymphatic obstruction, the accumulation of edema fluid raises interstitial fluid pressure, which acts as a safety factor by reducing capillary filtration. However, the second and third safety factors discussed earlier are absent in the case of lymphedema because these factors depend on an increase in lymph flow. In lymphedema, a decreased lymph flow is the causative problem, so there cannot be an increased lymph flow (second safety factor) to help carry away the excess interstitial fluid. Moreover, when lymph flow is impaired, plasma proteins accumulate in the interstitial fluid instead of being carried away in the lymph. Therefore the third safety factor that protects against edema (decreased interstitial fluid oncotic pressure) is also compromised in lymphedema.

Another cause of edema is physical injury or an allergic reaction to antigen challenges. Physical trauma, such as a scratch or a cut on the skin, results in a localized bump or swelling. A similar swelling is observed when the skin reacts to an irritating agent or

Vet-ebooks.com

antigen challenge (e.g., response to an insect bite). Inflammation and swelling can also occur in bronchial tissue during an asthmatic reaction. Edema in the airway can be life-threatening because it limits airflow to the lungs. An injury or antigen challenge stimulates the release of the chemical *histamine* from mast cells in the affected tissue and also stimulates the production of the peptide *bradykinin* in the blood and interstitial fluid. Both histamine and bradykinin have a pair of effects that cause edema in the reacting tissue (see Fig. 23.8). First, these chemicals increase the permeability of capillaries to plasma proteins. As proteins leave the bloodstream and accumulate in the interstitial space, they increase the interstitial fluid oncotic pressure, which promotes filtration of fluid. Second, histamine and bradykinin relax arteriolar smooth muscle. The arterioles dilate, and the resulting decrease in arteriolar resistance allows more of the arterial blood pressure to impinge on the capillaries. This leads to an increase in the capillary hydrostatic pressure, which promotes filtration. Although histamine and bradykinin promote excess filtration and edema through two mechanisms, all three safety factors that protect against edema are intact and act to limit the degree of edema.

Other situations also cause edema, but the examples discussed here cover some of the most common causes of clinical edema. These examples also reinforce an understanding of the interplay of the osmotic (oncotic) and hydrostatic forces that act on water to govern its filtration out of capillaries or its reabsorption into capillaries.



• **Fig. 23.8** Histamine mediates the changes that lead to edema in response to a physical injury or an antigen challenge. The normal three safety factors against edema are intact and help to limit the degree of edema. Treatment with an antihistamine (a drug that blocks histamine receptors on arterioles and capillaries) would also help to reduce the edema.

CLINICAL CORRELATIONS

Acute Protein-Losing Enteropathy in a Horse

History

You are called to a home a few miles from your clinic by parents concerned about their daughter's 4-year-old Quarter Horse mare. They report that the horse is listless and has had diarrhea for 2 days.

Clinical Examination

You arrive at the client's home to find that the horse is stabled in a small barn with poor ventilation and no access to pasture. Low-quality grass hay is stacked in the barn. On physical examination, you find the horse to be somewhat emaciated with dry mucous membranes, a foul-smelling diarrhea, and an elevated heart rate (tachycardia). When you pinch a section of the horse's skin, it falls back to the normal position slowly, which indicates dehydration. The horse's temperature is within a normal range.

You take a blood sample and then begin an intravenous administration of polyionic fluid (lactated Ringer's solution). You tell the clients that you will return later. Analysis of the blood sample indicates a hematocrit of 55% (normal range for the horse, 35%–45%) and a plasma protein concentration of 4.5 g/dL (normal range, 5.9–7.8 g/dL). You become concerned that the administration of fluids, without replacement of plasma proteins, will exacerbate the horse's hypoproteinemia, so you arrange to obtain plasma from a donor horse. You return with the plasma and find that the sick horse is still listless. Edema is now evident along the horse's ventral abdomen and in the limbs.

Comment

Acute enteropathy (intestinal disorder) often causes diarrhea. The loss of water and solutes leads to dehydration; blood volume and interstitial fluid volume are both reduced. The hematocrit (fraction of cells in blood) is typically elevated because fluid is being removed from the bloodstream, but blood cells are not. In some forms of enteropathy (called *protein-losing enteropathy*), the capillaries in the intestine become leaky to plasma proteins. Albumin, in particular, moves from the bloodstream into the intestinal lumen and is eliminated in the feces.

This horse has a severe shortage of plasma proteins. The shortage of plasma proteins probably resulted from a combination of poor nutrition (which depresses the production of plasma proteins by the liver) and protein-losing enteropathy. The deficit of plasma proteins in this horse is even more severe than indicated by the plasma protein concentration of 4.5 g/dL, because this value is the net result of two opposing processes. The loss of protein in the diarrhea lowered the plasma protein concentration, but the loss of water (dehydration) caused a decrease in plasma volume and therefore increased the concentration of the remaining proteins in the plasma.

The development of edema in this horse was predictable. The administration of intravenous fluids added water and electrolytes (but no proteins) to the circulating blood volume, and this caused a further reduction in the concentration of the plasma proteins remaining in the bloodstream. The resulting further decrease in plasma oncotic pressure led to excessive filtration of fluid out of capillaries and into the interstitial space. The consequence was edema, especially in the dependent regions of the body (ventral abdomen and legs). Restoration of a normal plasma protein concentration would reverse the edema.

Treatment

Bacterial or parasitical infections are a common cause of protein-losing enteropathy. If this horse had a fever, an infectious cause would be more likely. Acute enteropathy without fever (as in this case) is often self-limiting. Therefore the aim of treatment should be to remedy the dehydration, the electrolyte loss, and the plasma protein deficit. Intravenous administration of plasma, in addition to polyionic fluids, is usually effective. In some cases, antibiotics are also indicated, because enteropathy involves inflammation of

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

the intestinal wall, which can allow transmural migration of bacteria (and toxic bacterial products) from the gastrointestinal tract into the peritoneum. Important steps for long-term health in this horse would include better nutrition, regular deworming, and improved stable management.

Vasculitis in a Mare

History

A client calls you about a 20-year-old Quarter horse mare that he recently adopted from a rescue organization. The mare has poor appetite, is depressed, and has swollen legs. The rescue organization did not give the mare vaccines or perform a Coggins test for equine infectious anemia (EIA).

Clinical Examination

The mare has an increased temperature (105° F) and an abnormally high heart rate and respiratory rate. You notice petechia (small hemorrhages) on her mucous membranes and ears. Her capillary refill time (CRT) is prolonged. Her gastrointestinal borborgymi (rumbling sounds indicating gastrointestinal motility) is decreased. Her legs are warm and swollen, and applying light pressure to them seems to cause pain. You suspect the mare may have EIA, so you recommend keeping her in quarantine, and you draw blood samples for laboratory analysis, including a Coggins test.

Comment

Equine infectious anemia is caused by a retrovirus. Horse-to-horse transmission can occur via direct contact with body fluids, but more frequently the retrovirus is carried by horse flies or deer flies. Infected horses can develop acute or chronic disease. Acute disease is characterized by fever, inappentence, lethargy, dependent edema, petechia, and icteric (yellow) mucous membranes. Blood studies from affected horses typically reveal anemia, hemolysis (rupture of red blood cells), thrombocytopenia (low platelet count), hypoalbuminemia (low plasma albumin), and hyperglobulinemia (elevated plasma globulins). The infection leads to hemolysis (and hence the anemia). The breakdown of hemoglobin from the ruptured red cells yields several substances, including bilirubin, which is bright yellow (hence the icteric mucous membranes). The infection also causes vasculitis (inflammation of blood vessels), and this accounts for several of the other signs, including the warm, swollen, and painful legs. Specifically, vascular inflammation causes capillary permeability to increase, and this allows plasma proteins to escape from the capillaries into the interstitial fluid. The result is increased interstitial fluid oncotic pressure and decreased capillary plasma oncotic pressure. This leads to increased capillary filtration (and hence the edema). Also, vasculitis leads to the breakdown of some blood vessel walls, resulting in localized hemorrhage (hence the petechia).

Treatment

EIA is a reportable disease. There is no cure. Most infected horses are euthanized because they must otherwise be quarantined for life to minimize spread of the disease. Affected horses must also be branded and tested frequently. If an owner chooses to keep the horse, treatment is supportive. Even with treatment, horses may succumb to the disease either during the acute or chronic phase of infection.

Practice Questions

1. The rate of diffusion of glucose molecules from capillary blood to interstitial fluid is most directly affected by the:
- **a.** Capillary plasma oncotic pressure.
- **b.** Interstitial fluid hydrostatic pressure.
- **c.** Size and number of capillary pores.
- **d.** Amount of oxygen in the blood.
- **e.** Hematocrit.

2. For continuous capillaries, such as those found in skeletal muscle, the capillary surface area available for diffusion is greatest for which of the following?
- **a.** Glucose
- **b.** Dissolved oxygen
- **c.** Ions, such as Na^+, K^+, and Cl^-
- **d.** Amino acids
- **e.** Plasma proteins

3. An increase in the venous resistance (e.g., too tight a bandage on an extremity) causes edema because:
- **a.** Lymph flow increases.
- **b.** Pinocytosis cannot remove excess interstitial fluid fast enough.
- **c.** Capillary hydrostatic pressure increases.
- **d.** The interstitial proteins block up lymph channels.
- **e.** The interstitial fluid pressure decreases.

4. Which of the following will NOT cause pulmonary edema?
- **a.** An increase in pulmonary capillary permeability to protein
- **b.** A blockage of pulmonary lymph vessels
- **c.** An increase in left atrial pressure
- **d.** A constriction of pulmonary arterioles
- **e.** Left-sided heart failure

5. A patient with a form of protein-losing kidney disease has a plasma colloid osmotic pressure of 10 mm Hg. The patient has edema but is not getting any worse. Blood pressure and heart rate are normal. Which of the following is probably preventing further edema?
- **a.** Increased capillary hydrostatic pressure
- **b.** Decreased lymph flow
- **c.** Decreased concentration of plasma proteins in the plasma
- **d.** Increased interstitial fluid oncotic pressure
- **e.** Increased interstitial fluid hydrostatic pressure

6. The following parameters exist in the microcirculation of a skeletal muscle during a period of vigorous exercise:
P_c (capillary hydrostatic pressure) = 34 mm Hg
P_i (interstitial fluid hydrostatic pressure) = 10 mm Hg
π_c (capillary plasma oncotic pressure) = 24 mm Hg
π_i (interstitial fluid oncotic pressure) = 3 mm Hg
Which of the following is *true*?
- **a.** These conditions favor filtration.
- **b.** These conditions favor reabsorption.
- **c.** These conditions favor neither filtration nor reabsorption.
- **d.** It is not clear what these conditions favor because the concentration of plasma protein is not specified.
- **e.** It is not clear what these conditions favor because the rate of lymph flow is not specified.

Vet-ebooks.com

7. Histamine and bradykinin cause edema by increasing BOTH:
   a. Capillary permeability to plasma protein AND interstitial hydrostatic pressure.
   b. Interstitial colloid osmotic pressure AND lymph flow.
   c. Capillary hydrostatic pressure AND capillary permeability to plasma protein.
   d. Capillary hydrostatic pressure AND plasma colloid osmotic pressure.
   e. Lymph flow AND interstitial hydrostatic pressure.

8. During a 30-minute hemorrhage, a horse loses a substantial volume of blood. The horse's mean arterial pressure decreases from 90 to 75 mm Hg, and the heart rate increases from 40 to 90 beats/min. The skin becomes cool and the mucous membranes become pale, suggesting marked vasoconstriction. Because hemorrhage involves the loss of whole blood (both plasma and cells), you might expect that, soon after such a hemorrhage, the horse's remaining blood would still have a normal composition. However, you take a blood sample and discover that the hematocrit is abnormally low (only 28%). Which of the following would most likely account for the decrease in hematocrit observed after the hemorrhage?
   a. Excessive capillary filtration has caused interstitial fluid pressure to increase above normal.
   b. Excessive capillary filtration has caused capillary colloid osmotic pressure to increase above normal.
   c. Arteriolar constriction has caused capillary hydrostatic pressure to increase above normal.
   d. Low capillary hydrostatic pressure has caused interstitial fluid to be reabsorbed into the bloodstream.
   e. Many blood cells have been filtered out of capillaries and into the interstitial fluid.

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.

Koeppen BM, Stanton BA. *Berne & Levy Physiology*. 7th ed. Elsevier; 2017.

Kumar V, Addas AK, Aster JC. *Robbins and Cotran Pathologic Basis of Disease*. 9th ed. Elsevier; 2014.

Lavoie J, Hinchcliff KW, eds. *Blackwell's Five-Minute Veterinary Consult: Equine*. 2nd ed. Wiley-Blackwell; 2009.

Milnor WR. *Cardiovascular Physiology*. New York: Oxford University Press; 1990.

Mohrman DE, Heller LJ. *Cardiovascular Physiology*. 8th ed. McGraw-Hill; 2013.

Mortillaro NA, Taylor AE, eds. *The Pathophysiology of the Microcirculation*. Boca Raton, Fla: CRC Press; 1994.

Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.

Reed SM, Bayly WM, Sellon DC. *Equine Internal Medicine*. 4th ed. Elsevier; 2017.

Scallan J, Huxley VH, Korthuis RJ. *Capillary Fluid Exchange: Regulation, Function, and Pathology*. San Rafael, Calif: Morgan & Claypool Life Sciences; 2010.

Smith BP, ed. *Large Animal Internal Medicine*. 5th ed. Elsevier; 2014.

Sprayberry KA, Robinson NE. *Robinson's Current Therapy in Equine Medicine*. 7th ed. Elsevier; 2014.

Yuan SY, Rigor RR. *Regulation of Endothelial Barrier Function (Integrated Systems Physiology, From Molecule to Function to Disease)*. Morgan & Claypool Life Sciences; 2011.

Vet-ebooks.com

24

Local Control of Blood Flow

ROBERT B. STEPHENSON

KEY POINTS

1. Vascular resistance is affected by intrinsic and extrinsic control mechanisms.
2. Metabolic control of blood flow is a local mechanism that matches the blood flow of a tissue to its metabolic rate.
3. Autoregulation is a relative constancy of blood flow in an organ despite changes in perfusion pressure.
4. Many chemical signals act locally (as paracrines) to exert important control on vascular resistance.
5. Regardless of the status of arterioles, mechanical compression can reduce blood flow to a tissue.

Vascular Resistance Is Affected by Intrinsic and Extrinsic Control Mechanisms

As described in Chapter 22, the blood flow through any organ or tissue is determined by the perfusion pressure (arterial pressure minus venous pressure) and by the resistance of the blood vessels of the organ (and by no other factors), as follows:

$$\text{Blood flow} = \text{Perfusion pressure} \div \text{Vascular resistance}$$

Normally, all the organs of the systemic circulation are exposed to the same perfusion pressure. Therefore differences in blood flow to the various organs result from their different vascular resistances. The vascular resistance of an organ is determined mainly by the diameter of its arterioles. Thus arteriolar vasodilation or vasoconstriction is the predominant mechanism that increases or decreases the blood flow in one organ relative to another organ.

It is useful to divide the factors that affect arteriolar resistance into intrinsic and extrinsic mechanisms. *Extrinsic mechanisms* act from outside an organ or tissue, through nerves or hormones, to alter arteriolar resistance. By contrast, *intrinsic* (or *local*) *mechanisms* act within an organ or tissue; they do not rely on an external influence. For example, as described in Chapter 23, histamine is released from mast cells of a tissue in response to injury or during an allergic reaction. Histamine acts locally on the arteriolar smooth muscle to relax it. Dilation of the arterioles decreases arteriolar resistance and therefore increases blood flow to the tissue. Histamine is an example of a *paracrine:* a substance released from one type of cell that acts on another cell type in the vicinity. Paracrine signaling molecules move by diffusion, which is why paracrine signaling is only effective over very short distances. A second example of intrinsic control is the arteriolar dilation and increased blood flow during exercise in skeletal muscle. This example illustrates the general phenomenon of *metabolic control* of blood flow, by which tissues tend to increase their blood flow whenever their metabolic rate increases; neither nerves nor hormones are required.

Although the arterioles in all tissues are affected by both intrinsic and extrinsic mechanisms, intrinsic mechanisms predominate over extrinsic mechanisms in the control of arterioles in the brain, heart (i.e., coronary circulation), and working skeletal muscle. By contrast, extrinsic mechanisms predominate over intrinsic mechanisms in the control of blood flow to the kidneys, splanchnic organs, and resting skeletal muscle. Skin is an example of an organ in which both intrinsic and extrinsic control mechanisms have strong influences. In general, intrinsic control dominates extrinsic control in the *critical organs:* those that must have sufficient blood flow to meet their metabolic needs on a second-by-second basis for an animal to survive. Extrinsic control dominates intrinsic control in organs that can withstand temporary reductions in blood flow (and metabolism) to make extra blood available for the critical organs.

To rationalize the distinction between "critical" and "noncritical" organs, envision a lion crouching in the tall grass on an African savannah. Nearby, an antelope is grazing. A moment later, both the lion and the antelope are engaged in maximal exercise, as the lion attempts to capture its breakfast and the antelope tries to escape. For each, survival depends on adequate delivery of blood flow to the brain, heart, and working skeletal muscle. During the chase, the kidneys, splanchnic organs, and any nonexercising skeletal muscle become noncritical, in the sense that their blood flow can be reduced to make extra blood available for the critical organs (brain, heart, and working skeletal muscle).

Metabolic Control of Blood Flow Is a Local Mechanism That Matches the Blood Flow of a Tissue to Its Metabolic Rate

Metabolic control of blood flow is the most important local control mechanism. For example, metabolic control accounts for the huge increase in blood flow through a skeletal muscle as it goes from rest to maximal exercise. The functional significance of metabolic control of blood flow is that it matches the blood flow in a tissue to the metabolic rate of the tissue. An increase in tissue blood flow in response to increased metabolic rate is called *active hyperemia* (*hyper* means "elevated," *emia* refers to blood, and *active* implies an increased metabolic rate).

Metabolic control of blood flow works by means of chemical changes within the tissue. When the metabolic rate of a tissue increases, its consumption of oxygen increases, and there is an increased rate of production of metabolic products, including carbon dioxide, adenosine, and lactic acid. Also, some potassium ions

Vet-ebooks.com

TABLE 24.1 Chemical Signals Important in Local Control of Systemic Arterioles[a]

Chemical Signal	Source	Effect
Signals Related to Metabolism		
Oxygen	Delivered by arterial blood; consumed in aerobic metabolism	Vasoconstriction. (Rapid metabolism depletes O_2, which causes vasodilation.)
Carbon dioxide	Produced by aerobic metabolism	Vasodilation
Potassium ions (K^+)	Released from rapidly metabolizing cells	Vasodilation
Adenosine	Released from rapidly metabolizing cells	Vasodilation
Metabolic acids (e.g., lactic acid)	Produced by anaerobic metabolism	Vasodilation
Other Local Chemical Signals (Paracrines)		
Endothelin-1 (ET1)	Endothelial cells	Vasoconstriction
Nitric oxide (NO)	Endothelial cells and some parasympathetic nerve endings	Vasodilation
Thromboxane A_2 (TXA_2)	Platelets	Vasoconstriction (also increases platelet aggregation)
Prostacyclin (PGI_2)	Endothelial cells	Vasodilation (also decreases platelet aggregation)
Histamine	Mast cells	Vasodilation (also increases capillary permeability)
Bradykinin	Globulins in blood or tissue fluid	Vasodilation (also increases capillary permeability)

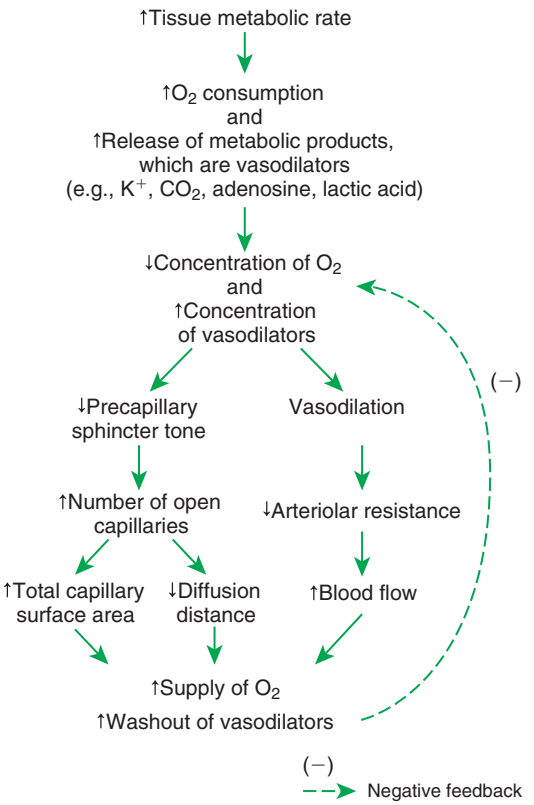
[a]Some of these chemical signals have different effects on pulmonary blood vessels than on systemic vessels. A high level of oxygen, for example, causes dilation of pulmonary vessels, whereas the effect in systemic vessels is vasoconstriction. See Chapter 46 for more details.

(K^+) escape from rapidly metabolizing cells, and these ions accumulate in the interstitial fluid. Therefore as the metabolism of a tissue increases, the interstitial concentration of oxygen decreases, and the interstitial concentrations of metabolic products and K^+ increase. All these changes have the same effect on arteriolar smooth muscle: they relax it (Table 24.1). The arterioles dilate, vascular resistance decreases, and more blood flows through the tissue.

Low levels of oxygen and high concentrations of metabolic products and K^+ also cause relaxation of the precapillary sphincters (in the tissues that have them), and this opens more of the capillaries in the tissue to blood flow. As explained in Chapter 23, when more capillaries carry blood flow, that decreases the diffusion distance between fresh, oxygenated blood and the metabolizing cells of the tissue. Opening more capillaries also increases the total capillary surface area for diffusional exchange. The net result of the increased blood flow, the decreased diffusion distance, and the increased total capillary surface area is a more rapid delivery of oxygen and other metabolic substrates to the tissue cells and a more rapid removal of metabolic waste products from the tissue.

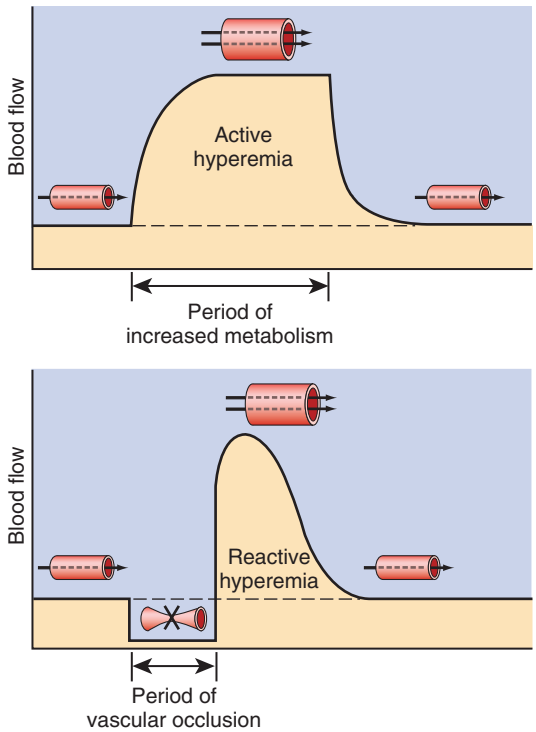
Metabolic control of blood flow involves negative feedback. The accumulation of metabolic products and the lack of oxygen initiate vasodilation, which increases blood flow. The increased blood flow hastens the removal ("wash-out") from the tissue of the accumulating metabolic products and delivers additional oxygen. A new balance is reached when the increased blood flow closely matches the increased metabolic needs of the tissue. Fig. 24.1 summarizes the major features of metabolic control of blood flow.

Reactive hyperemia is a temporary increase above normal in the flow of blood to a tissue after a period when blood flow was restricted. In this case the increased flow (*hyperemia*) is a response (*reaction*) to a period of inadequate blood flow. Mechanical compression of blood vessels is one cause of inadequate blood flow, and subsequent release of that mechanical compression elicits reactive hyperemia. This can be easily demonstrated in any accessible

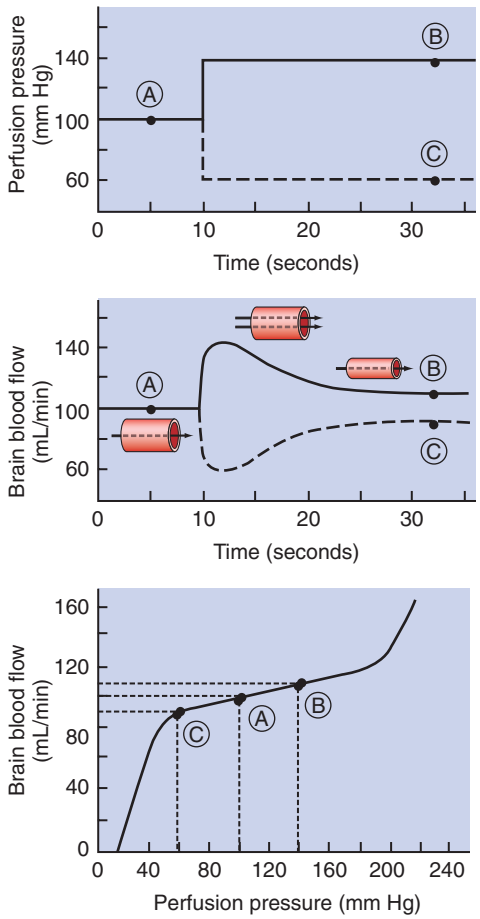


• **Fig. 24.1** Metabolic control of blood flow is a local (intrinsic) mechanism that acts within a tissue to match the blood flow to the tissue with the metabolic activity of the tissue. As a tissue becomes more active metabolically, the metabolic control mechanism increases blood flow and thereby regulates the concentration of oxygen and metabolic products in the tissue.

Vet-ebooks.com



• **Fig. 24.2** Both active hyperemia and reactive hyperemia involve increases above normal in blood flow. Both phenomena are brought about by the mechanisms for the local, metabolic control of blood flow, as discussed in the text. Sketches of tubes depict the diameter of an arteriole at various points in time; arrows within the tube indicate blood flow.



• **Fig. 24.3** Experimental demonstration of autoregulation of blood flow in the brain of a dog. Perfusion pressure was artificially increased or decreased (*top*), and the resulting changes in brain blood flow were measured (*middle*). The steady-state values of blood flow were then plotted against perfusion pressure (*bottom*). Points *A, B,* and *C* are discussed in the text. Sketches of tubes (in middle panel) depict the diameter of an arteriole at three points in time (at Point *A,* at the moment of maximum brain blood flow, and at Point *B*); arrows within the tube indicate blood flow.

nonpigmented epithelial tissue. For example, press a finger against nonpigmented skin hard enough to occlude the blood flow. Maintain the pressure for about 1 minute, then release. After release of the pressure, blood flow through the previously compressed skin will become greater than normal for several seconds, which will cause the skin to appear darker (redder) than the surrounding (noncompressed) skin.

The same metabolic control mechanisms that account for active hyperemia also explain reactive hyperemia. During the period when mechanical compression restricts blood flow, metabolism continues in the compressed tissue, metabolic products accumulate, and the local concentration of oxygen decreases. These metabolic effects cause dilation of the arterioles and a decrease in arteriolar resistance. When the mechanical obstruction to flow is removed, blood flow increases above normal until the "oxygen debt" is repaid and the excess metabolic products have been removed from the compressed tissue. Fig. 24.2 compares active and reactive hyperemia.

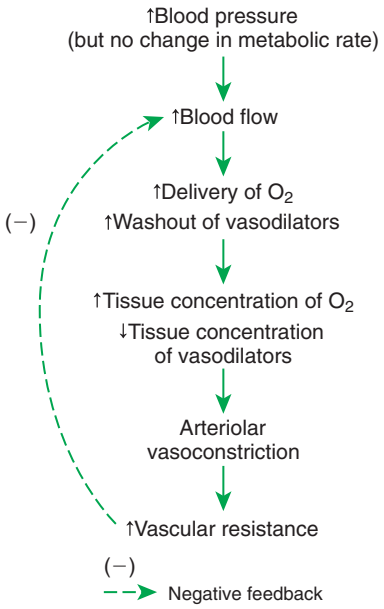
Autoregulation Is a Relative Constancy of Blood Flow in an Organ Despite Changes in Perfusion Pressure

Metabolic control mechanisms also participate in the phenomenon known as *blood flow autoregulation*. Autoregulation is evident in denervated organs and organs in which local control of blood flow is predominant over neural and humoral control (e.g., in coronary circulation, brain, and working skeletal muscle).

Fig. 24.3 summarizes an experiment that demonstrates autoregulation in the brain of a dog. Initially, the perfusion pressure (arterial pressure minus venous pressure) in this animal is 100 mm Hg, and the blood flow to the brain is 100 milliliters per minute (mL/min) (point *A*). When perfusion pressure is increased suddenly to 140 mm Hg, brain blood flow rises initially to 140 mL/min but returns toward its initial level over the next 20 to 30 seconds. Eventually, blood flow reaches a stable level of about 110 mL/min (point *B*). Conversely, if the perfusion pressure is decreased suddenly from 100 to 60 mm Hg, blood flow in the brain decreases initially to 60 mL/min but returns toward its initial level over the next 20 to 30 seconds (see *dashed lines* in the top and middle graphs of Fig. 24.3). Eventually, blood flow reaches a stable level of about 90 mL/min (point *C*). These steady-state responses are plotted in the bottom graph. The remainder of the bottom graph is obtained in a similar way; that is, perfusion pressure is set artificially to various levels, ranging from 20 to 220 mm Hg, and the resulting steady-state levels of blood flow are plotted.

Over a considerable range of perfusion pressure (about 60–190 mm Hg), relatively little change occurs in steady-state blood flow to the brain; that is, brain blood flow is autoregulated. The range of perfusion pressures over which flow remains relatively

Vet-ebooks.com



• **Fig. 24.4** The same metabolic mechanism responsible for active hyperemia and reactive hyperemia can also account for autoregulation, in which blood flow to an organ stays relatively constant despite changes in perfusion pressure.

constant is called the *autoregulatory range*. Autoregulation fails at very high and very low perfusion pressures; that is, extremely high pressures result in marked increases in blood flow, and extremely low pressures result in marked decreases in blood flow. Nevertheless, over a considerable range of perfusion pressure, autoregulation keeps blood flow in the brain relatively constant.

Fig. 24.4 shows how the metabolic control mechanisms previously described can account for the phenomenon of autoregulation. If the metabolic rate of an organ does not change but perfusion pressure is increased above normal, the increased pressure forces additional blood flow through the organ. The additional blood flow hastens the removal of metabolic products from the interstitial fluid and increases the rate of oxygen delivery to the interstitial fluid. Therefore the concentration of vasodilating metabolic products in the interstitial fluid decreases, and the concentration of oxygen in the interstitial fluid increases. These changes cause the arterioles of the tissue to constrict, which increases the resistance to blood flow above normal. The consequence is that blood flow decreases back toward its initial level, despite the continuation of the elevated perfusion pressure.

To summarize, metabolic control mechanisms bring about active hyperemia (the increase in blood flow in an organ in response to an increased metabolic rate in the absence of any blood pressure change). The same metabolic mechanisms can also account for reactive hyperemia (the temporary increase in blood flow above normal in an organ after a period of flow restriction). In addition, the same metabolic mechanisms can account for autoregulation (the relative constancy of blood flow in an organ when there has been no change in metabolic rate but blood pressure has either increased or decreased). Other mechanisms also contribute to autoregulation, and the reader may encounter discussions of these under the terms *myogenic hypothesis* and *tissue pressure hypothesis*. However, metabolic control plays a major role in autoregulation of blood flow, particularly in the critical tissues of a body (brain, coronary vessels, and exercising skeletal muscle).

Many Chemical Signals Act Locally (as Paracrines) to Exert Important Control on Vascular Resistance

As already described, metabolic control of blood flow is mediated by chemical changes that occur when tissue metabolism increases. In addition to the signaling molecules that mediate metabolic control of blood flow, there are many other chemicals that act locally, within a tissue, to affect vascular resistance and therefore blood flow. Some of these locally acting (*paracrine*) chemical signals are listed in Table 24.1.

Endothelin-1 (ET1) is released from endothelial cells in response to a variety of mechanical or chemical stimuli, especially those that traumatize the endothelium. ET1 causes vascular smooth muscle to contract, which results in vasoconstriction and a decrease in blood flow. *Nitric oxide* (NO), another signaling molecule released from endothelial cells, has the opposite effect. NO relaxes vascular smooth muscle, which results in vasodilation. One stimulus for NO release is an increase in blood flow velocity past the endothelium. NO acts locally to dilate vessels, especially small arteries, which allows them to accommodate an increased blood flow without such high flow velocities. In some tissues, most notably the erectile tissues of the external genital organs (penis and clitoris), parasympathetic nerve endings release NO and the neurotransmitter acetylcholine. The acetylcholine stimulates endothelial cells to release additional NO. The NO from the nerve endings, augmented by the NO from the endothelial cells, dilates local blood vessels, which causes engorgement of the tissues with blood, and therefore erection.

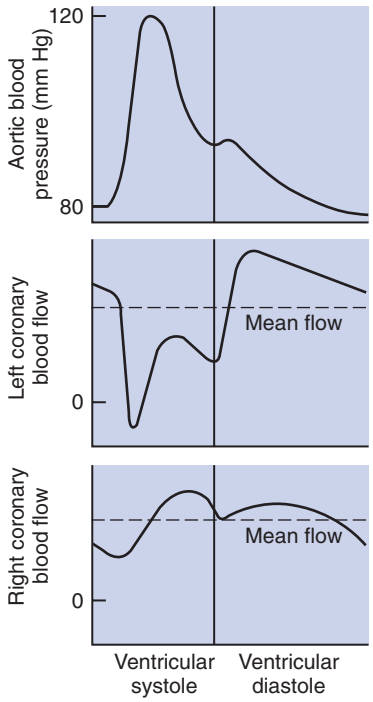
Thromboxane A2 (TXA_2) and *prostacyclin* (PGI_2) act antagonistically in the control of vascular smooth muscle and also in the control of platelet aggregation. Thus the relative balance between TXA_2 and PGI_2 is more important than the absolute level of either chemical alone. Under normal conditions, the balance ensures adequate blood flow to tissues and prevents platelet aggregation. If blood vessels become traumatized or rupture, the balance shifts in favor of TXA_2. The resulting vasoconstriction and platelet aggregation are critical in minimizing blood loss. In some pathological states, imbalances develop between TXA_2 and PGI_2. Depending on the direction of the imbalance, the result is either excessive vasoconstriction and blood coagulation or excessive vasodilation and bleeding.

Histamine, which is released from mast cells, is another locally acting vasodilator. The role of histamine in the vascular responses to tissue injury or antigen challenge is described in Chapter 23 (see Fig. 23.8). *Bradykinin* is another signaling chemical that causes vasodilation. Bradykinin is a small polypeptide that is split away by the proteolytic enzyme *kallikrein* from globulin proteins, which are present in plasma or tissue fluid. Bradykinin may also be formed in sweat glands when they are activated by acetylcholine released from sympathetic nerve endings. The resulting vasodilation of skin blood vessels, together with the evaporation of sweat, promotes heat loss from the skin. Both histamine and bradykinin exert their vasodilating effects, at least in part, by stimulating the formation of NO.

Regardless of the Status of Arterioles, Mechanical Compression Can Reduce Blood Flow to a Tissue

Mechanical compression can reduce blood flow in a tissue by literally squeezing down on all its blood vessels. The example of compressing

Vet-ebooks.com



• **Fig. 24.5** Coronary blood flow to the left ventricular muscle is greatly reduced during ventricular systole because the left ventricular muscle contracts so forcefully that it compresses the left ventricular blood vessels. Coronary blood flow to the right ventricular muscle is less affected by mechanical compression because the contractions of the right ventricle are less forceful that those of the left ventricle.

skin blood vessels for a minute and then releasing the compression has been mentioned as a way to trigger a readily visible reactive hyperemia. Long-term mechanical pressure on the skin must be avoided, however, because a prolonged period of subnormal blood flow (ischemia) leads to reversible tissue damage, then irreversible tissue damage, and eventually to cell death (necrosis). Pressure sores are an unfortunate and common example of this sequence. Three additional examples of mechanical compression are also described because of their clinical importance.

Fig. 24.5 illustrates the effect of mechanical compression on blood flow through the coronary vessels. The top tracing shows the changes in arterial (aortic) blood pressure during one complete cardiac cycle. The periods of ventricular systole and ventricular diastole are labeled at the bottom of the figure. One would expect that blood flow through the coronary circulation would be highest during ventricular systole (when the aortic pressure is highest) and that flow would be lowest during diastole (when the aortic pressure is lowest). However, the tracings of left coronary blood flow indicate that blood flow to the left ventricular muscle is actually depressed during systole and much higher during diastole. Flow even reverses (blood flows backward, *toward* the aorta) momentarily near the beginning of systole. The fact that left coronary blood flow is much lower during systole, even though the perfusion pressure is higher, implies that the resistance of the coronary vessels must be substantially higher during systole than during diastole.

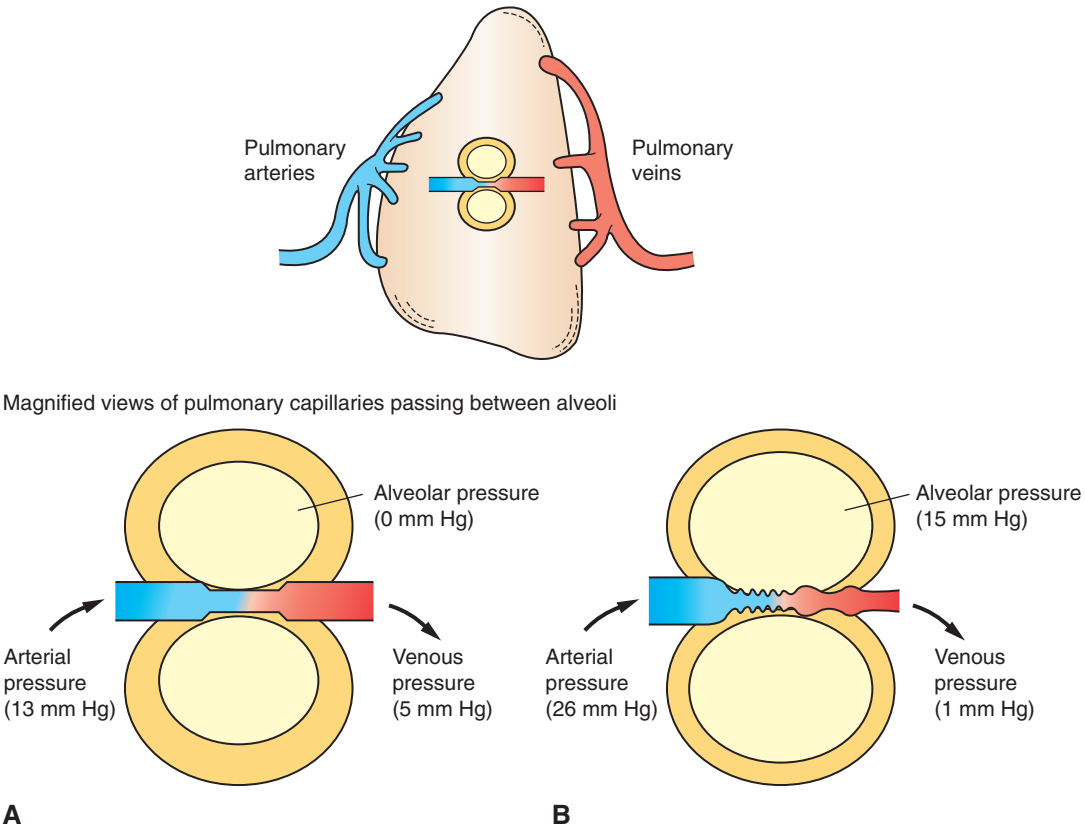
Left coronary resistance is high during systole because the contracting left ventricular muscle squeezes down on the coronary blood vessels. The coronary vessels are not constricted in this way during diastole because the ventricular muscle is relaxed. Therefore coronary vascular resistance decreases dramatically (and blood flow increases) during diastole. The bottom tracing in Fig. 24.5 indicates that mechanical compression has relatively little influence on blood flow to the right ventricular muscle; that is, the magnitude of right coronary blood flow closely follows the changes in arterial pressure (being highest during systole and lowest during diastole). Right coronary flow is not restricted by mechanical compression during systole because the right ventricle contracts with much less force than the left ventricle. The right ventricle simply does not develop sufficient compressive force to constrict its own blood vessels.

Most of the blood needed to support left ventricular metabolism must be delivered during ventricular diastole when the vessels are not compressed. This fact has great clinical significance. In a resting animal with a low heart rate, there is adequate time during diastole for the coronary vessels to supply the amount of blood needed by the ventricular tissue. During exercise, the metabolic rate of the ventricular muscle cells increases dramatically. To support the increased metabolic rate, the ventricular muscle requires much greater blood flow during exercise than at rest. However, the duration of diastole is reduced during exercise, so there is less time available for delivery of this increased flow. Nevertheless, normal, healthy coronary vessels have a sufficiently low resistance (during diastole) to supply the needed blood flow, even during maximal exercise. The situation is different, however, in animals with coronary artery disease. In animals whose coronary vessels are narrowed because of atherosclerosis, blood flow cannot increase enough to supply the needs of the vigorously working ventricular muscles. This is why ventricular ischemia develops during exercise in patients with coronary artery disease. Ischemic areas of the ventricle fail to contract normally. Ischemia can also cause arrhythmias or even ventricular fibrillation (sudden death). Coronary artery disease is more common in humans than in veterinary species, so this scenario is more likely to occur in the veterinarian than in the veterinarian's patients.

Mechanical compression caused by muscle contraction can also restrict blood flow through skeletal muscles. The blood vessels within skeletal muscles become compressed during strenuous, sustained contractions of the muscle. The compression reduces blood flow through the muscle, which can create ischemia. Ischemic muscles cannot contract with normal vigor. Ischemia also activates sensory nerve endings in the muscle, which causes pain. Activation of these muscle ischemia receptors also triggers a reflex increase in arterial pressure. The high arterial pressure is advantageous because it helps force blood flow through the skeletal muscle blood vessels, despite the compressive effects of the muscle contraction. The high arterial pressure of ischemic exercise is risky for patients with coronary artery disease, however, because high arterial pressure imposes a tremendous increase in workload on the heart. This is why patients with coronary artery disease are cautioned against types of exercise that involve strenuous, sustained muscle contractions, such as weightlifting.

Mechanical compression has important effects on the pulmonary circulation. Pulmonary vessels are more compliant than their counterparts in the systemic circulation. Greater compliance makes the pulmonary vessels more distensible but also makes them more susceptible to narrowing under the influence of mechanical compression. Moreover, because pulmonary arterial pressure is much lower than systemic arterial pressure, there is less intravascular pressure in a pulmonary vessel to oppose any external force acting to compress the vessel. Most pulmonary vessels travel within the tissues that comprise the walls of the airways, including the very thin walls of alveoli. Fig. 24.6 shows how an abnormal elevation in airway pressure can compress pulmonary blood vessels. This could happen during surgery if a patient has a tracheal tube inserted into its

Vet-ebooks.com



• **Fig. 24.6** Pulmonary blood vessels are susceptible to mechanical compression, which can be created by abnormally high pressure within the airways. (A) Under normal conditions, pulmonary arterial pressure is about 13 mm Hg and the venous pressure is about 5 mm Hg. The pressure within the pulmonary capillary depicted here would be intermediate between these two values. The pressure just outside the capillary (in the alveolar air space) is even lower; alveolar pressures typically vary between –1 mm Hg (during inspiration) and + 1 mm Hg (during expiration). Because the pressure inside pulmonary vessels is greater than the pressure outside, the vessels are not compressed. (B) If alveolar pressure increases to 15 mm Hg or higher, the pulmonary vessels become compressed. The resulting increase in pulmonary vascular resistance causes pulmonary blood flow to decrease, pulmonary arterial pressure to increase, and pulmonary venous pressure to decrease.

airway and if the tracheal tube is attached to a source of elevated pressure. The elevated pressure could be generated by a mechanical respirator that is not adjusted properly or by an anesthetist when he or she squeezes the bag attached to the tracheal tube. In either case, the pressures generated in the tracheal tube are transmitted through the airways and into the alveoli. An increase in airway pressure exerts a compressing force on the pulmonary blood vessels.

Alveolar pressures exceeding 10 to 15 mm Hg compress pulmonary blood vessels sufficiently to raise the resistance to blood flow through the lungs. As a result, blood ejected by the right ventricle dams up in the pulmonary arteries. This causes pulmonary arterial pressure to increase. An elevated pulmonary arterial pressure helps force blood through the compressed vessels. However, the increased pulmonary artery pressure also imposes an increased workload on the right ventricle. If the pressure in the airways is not excessively high, the right ventricle can generate a big enough increase in pulmonary arterial pressure to restore pulmonary blood flow almost to normal. However, with extremely high airway pressures, the right ventricle may be unable to raise pulmonary arterial pressure high enough to sustain flow. Under these conditions, pulmonary blood flow falls substantially below normal. Because the left heart can only pump as much blood as it receives via the pulmonary circulation, left ventricular output also decreases. The consequences can be fatal. The veterinary clinician must be mindful of the risks of high airway pressures whenever a patient is intubated and attached to a mechanical respiratory device.

CLINICAL CORRELATIONS

Patent Ductus Arteriosus

History

A 3-month-old female Welsh corgi is brought to your clinic by its owner who has noticed a "rumbling noise" in the dog's chest. The dog is smaller than her littermates and less playful. The dog coughs occasionally, but the cough does not produce fluid.

Clinical Examination

The dog appears to be in good health except for the occasional cough. The mucous membranes are pink, and the capillary refill time is normal (1.5 seconds). However, when you place your hand on the anterior left chest, you feel an abnormal vibration (*thrill*) with each heartbeat. With a stethoscope, you can auscultate a cardiac murmur that is loudest during systole but

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

continues throughout both systole and diastole (*continuous murmur*). The murmur is heard most loudly at the ventral third intercostal space on the left side. Expiratory sounds are slightly louder than normal. The heart rate is 152 beats/min, which you consider to be above normal for a dog of this size and age. While you are listening to the heart with the stethoscope, you palpate the femoral pulses, which are synchronized with the heart rate and very strong.

The electrocardiogram indicates that the dog has sinus tachycardia; the atrial and ventricular rates are both 152 beats/min. The R waves are abnormally large in leads II and III (2.5 and 3.5 mV, respectively). The QRS complex in lead I shows a large negative deflection followed immediately by a slightly larger positive deflection.

Thoracic radiographs show a generalized enlargement of the heart. The initial portion of the pulmonary artery is also substantially larger than normal, and the pulmonary blood vessels appear generally to be more prominent than normal.

An echocardiogram confirms the presence of a *patent ductus arteriosus* (*PDA*).

Comment

A murmur in a young, otherwise-healthy dog is most likely the result of a congenital cardiac abnormality. A continuous murmur can occur only if a defect causes turbulent flow throughout both systole and diastole. Because flow can occur only when there is a pressure gradient, the defect in this dog must be in a location where there is a substantial pressure gradient throughout the cardiac cycle. No single intracardiac defect meets this criterion; that is, a stenotic or regurgitant valve produces either a systolic murmur or a diastolic murmur but not both. A valve that is both stenotic and regurgitant produces two murmurs: one in systole and one in diastole. In such a case, however, there are brief moments during the cardiac cycle when no pressure gradient exists across the valve, so there are moments of silence between the systolic murmur and the diastolic murmur. (Admittedly, if the heart rate is high, these moments of silence are very brief, and the two murmurs can be mistaken for a continuous murmur, particularly in the case of combined aortic stenosis and regurgitation.)

The most common cardiac defect in a young animal that causes turbulent flow throughout both systole and diastole is a PDA. This vessel is normal in the fetus but should close shortly after birth. The flow through a PDA is continuous because aortic pressure is higher than pulmonary artery pressure throughout the cardiac cycle. The resulting murmur is usually heard best in the left third intercostal space. All the other clinical signs in this dog are consistent with the diagnosis of PDA. The prominence of the pulmonary vessels on the radiographs indicates that pressure and flow are abnormally high in the pulmonary artery and its branches. In a dog with a PDA, the pulmonary artery receives blood flow from both the right ventricle and the aorta, which increases both pulmonary arterial pressure and pulmonary flow.

The radiographs and electrocardiograms indicate that this dog has both right and left ventricular hypertrophy. The large R waves in leads II and III indicate left ventricular hypertrophy, and the large negative deflection during the QRS complex in lead I suggests that the right ventricle is hypertrophic as well. The left ventricle becomes hypertrophic in a dog with PDA because it is called on to pump three to five times the normal cardiac output. (It pumps a near-normal volume of blood to the organs of the systemic circulation and two to four times that much through the PDA.) The flow through the PDA is large because the PDA offers little resistance to flow. The demand on the left ventricle to pump so much blood (increased volume work) leads to left ventricular hypertrophy. The volume of blood pumped by the right ventricle is almost normal; it only needs to pump the blood that returns through the venae cavae from the systemic organs. However, the right ventricle must develop higher systolic pressures than normal to eject this blood into the pulmonary artery because pulmonary artery pressure is higher than normal, as explained earlier. This increase in pressure work leads to right ventricular hypertrophy.

Because the PDA carries so much blood away from the aorta, dogs with PDA tend to have an abnormally low aortic pressure. Diastolic pressure is particularly reduced because of the rapid outflow of blood from the aorta during ventricular diastole. Therefore PDA is typically associated with low mean aortic pressure but elevated pulse pressure (review Fig. 22.8E).

Two mechanisms work together to keep blood flow to the systemic organs almost normal despite the fact that a large fraction of cardiac output is "lost" through the PDA. First, reflex mechanisms (discussed in Chapter 25) increase sympathetic activity to the heart, which increases heart rate and contractility above normal. These sympathetic effects keep left ventricular output (and aortic pressure) sufficiently high to supply blood to the systemic organs, despite the PDA. Second, metabolic control mechanisms cause the systemic organs to vasodilate, which keeps their blood flow almost normal despite the subnormal aortic pressure.

The compensatory mechanisms just described allow most dogs with a PDA to maintain a nearly normal blood flow to the systemic organs at rest. Several months may pass before the dog's owner notices limitations in the dog's activity or growth. Eventually, however, the heart cannot increase its output sufficiently to supply the systemic blood flow needed by the muscles during exercise, so as time passes, a puppy with a PDA becomes less playful and energetic than its normal littermates. Also, if the heart is unable to supply the blood flow needed by metabolically active tissues, the owner may notice some stunting of growth. In any case, a dog with a widely open ductus has a poor long-term prognosis, unless treated.

Treatment

You show the dog's owner a diagram of the fetal circulation and explain that the *ductus arteriosus* normally closes and seals itself within 1 to 6 weeks after birth, but that the *ductus* fails to close spontaneously in about 1 of every 700 newborns (the condition is four times more common in female pups than in male pups). Treatment involves closure of the *ductus*, either by ligation during open-chest surgery or by insertion of a specially designed plug during a cardiac catheterization procedure. Most dogs treated before age 6 months go on to lead completely normal lives. However, you inform the owner that PDA is hereditary and that this puppy should probably not be used for breeding.

The owner elects to have the dog treated surgically, and the surgery is successful. The murmur and cough disappear immediately. Within 1 week, the dog is noticeably more energetic. At age 6 months, the dog has "grown into" her enlarged heart, and all physical findings are within normal limits.

Endotoxemia in a Foal

History

A 3-day-old Tennessee Walking Horse filly presents with progressive signs of lethargy, diarrhea, decreased eating, and weakness. The owners report that the filly appeared to be normal at birth, and shortly thereafter, she nursed briefly. Her condition did not cause them great concern until about a day ago.

Clinical Examination

The filly is markedly dehydrated. Although the environment is not cold, the filly has subnormal rectal temperature, suggesting that she can no longer thermoregulate. She has increased heart rate and respiratory rate. Her mucous membranes are dark red and exhibit prolonged capillary refill time, and her distal extremities feel cool. These signs indicate poor perfusion, low blood pressure, and hypoxia. She has hypermotile gut sounds and diarrhea. She only supports herself voluntarily for short periods of time. You suspect that the foal has an infection and is likely septic (bacteria and endotoxins in the blood). You submit a venous blood sample for immunoglobulin status (IgG), complete blood count (CBC) and biochemical profile, and culture. You also collect an arterial blood sample for measurement of blood gases.

Comment

This foal likely has acquired an infection from either ingestion or inhalation of contaminated liquid. Foals are frequently infected with gram-negative bacteria, and if they have not received adequate protection from antibodies in colostrum, the bacteria proliferate and release endotoxins. Circulating in the bloodstream, the bacteria and endotoxins stimulate the production of a large number of chemical mediators that cause inflammation, increased

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

capillary permeability, intravascular coagulation, cardiac depression, poor perfusion, and hypoxia. These chemical mediators include host proinflammatory intercellular signaling molecules (i.e., cytokines and chemokines), procoagulants, adhesion molecules, enzymes, and acute phase proteins (plasma concentrations change in association with inflammatory states). An additional complication is hypoproteinemia, resulting both from impaired intestinal absorption of nutrients and from loss of protein in the diarrhea.

Treatment

The bacterial infection must be treated aggressively with appropriate antibiotics. Additional treatments would include nutritional support, oxygen, and intravenous fluid therapy. The fluid therapy would include a combination of plasma (to counteract hypoproteinemia) and electrolytes (to correct dehydration). Glucose (dextrose in water) can also be given intravenously to prevent hypoglycemia. The foal must be monitored closely so that she does not become overhydrated, because she will then develop edema due to hypoproteinemia. Pulmonary edema would further jeopardize the oxygenation of blood and the delivery of adequate oxygen to the tissues. Additional drug treatments may be needed to enhance cardiac function and support blood pressure. In cases such as this one, foals are encouraged to nurse, or else are provided with milk, unless they develop ileus (a type of intestinal obstruction). Alternatively, parenteral (nonoral, often intravenous) nutrition can be provided. Antiinflammatory medications can be helpful; however, they must be used with caution because they can cause renal (kidney) failure or gastric and colonic ulcers. Prognosis is guarded in these cases because of the severity of disease and the lasting damage that it can cause in multiple organ systems (also including lungs and joints).

Practice Questions

1. Intrinsic control of blood flow is most likely to predominate over extrinsic control of blood flow in:
- **a.** The splanchnic circulation
- **b.** The kidneys
- **c.** Resting skeletal muscle
- **d.** Exercising skeletal muscle
- **e.** Skin

2. The increase in coronary blood flow during exercise is:
- **a.** Called Starling's law of the heart.
- **b.** Called reactive hyperemia.
- **c.** Caused by activation of parasympathetic nerves to the heart.
- **d.** Caused by compression of the coronary blood vessels during systole.
- **e.** Closely matched to the increased metabolic requirements of the heart muscle.

3. The metabolic control mechanism serves to match blood flow in a tissue to the tissue's metabolic activity. Which of the following variables is actually being regulated (i.e., held constant) by the metabolic control mechanism?
- **a.** Perfusion pressure applied to the tissue.
- **b.** Blood flow to the tissue.
- **c.** Arteriolar resistance in the tissue.
- **d.** Oxygen concentration in the tissue interstitial fluid.
- **e.** The number of open capillaries (i.e., the number of capillaries carrying blood flow at any one time).

4. A dog with an arterial blood pressure of 120/80 mm Hg has a cerebral blood flow of 100 mL/min. When blood pressure is increased to 130/100 mm Hg, the cerebral blood flow increases to 105 mL/min. This is an example of:
- **a.** Active hyperemia
- **b.** Autoregulation
- **c.** Reactive hyperemia
- **d.** The blood-brain barrier
- **e.** Hypoxic vasoconstriction

5. In response to an increase in perfusion pressure, the arterioles of an autoregulating organ _____, and the vascular resistance of the organ _____.
- **a.** constrict; increases
- **b.** constrict; decreases
- **c.** dilate; increases
- **d.** dilate; decreases

6. Which of the following characteristically acts as a paracrine to cause vasoconstriction in systemic arterioles?
- **a.** Carbon dioxide
- **b.** Nitric oxide
- **c.** Endothelin-1 (ET1)
- **d.** Prostacyclin (PGI2)
- **e.** Bradykinin

7. Blood flow to skeletal muscle increases dramatically during dynamic exercise (e.g., running or swimming, during which the exercising muscles alternate between contraction and relaxation). The increase in muscle blood flow is much smaller during static exercise (e.g., pushing hard against an immovable object, during which the exercising muscles maintain a forceful, steady contraction). The limited increase in muscle blood flow during static exercise is best explained by:
- **a.** Mechanical compression of muscle blood vessels.
- **b.** Metabolic control of muscle blood flow, leading to arteriolar vasodilation.
- **c.** Metabolic control of muscle blood flow, leading to arteriolar vasoconstriction.
- **d.** Autoregulation of blood flow, leading to arteriolar vasoconstriction.
- **e.** The release of histamine and bradykinin by the contracting muscle cells.

8. When a young dog with a PDA attempts vigorous exercise:
- **a.** Arterioles in the exercising skeletal muscle constrict.
- **b.** Oxygen concentration in the skeletal muscle interstitial fluid decreases.
- **c.** Left ventricular output decreases.
- **d.** Right ventricular output decreases.
- **e.** Mean arterial pressure increases to very high levels.

Vet-ebooks.com

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.
Keene BW, Smith FWK, Tilley LP, Hansen B. *Rapid Interpretation of Heart and Lung Sounds: A Guide to Cardiac and Respiratory Auscultation in Dogs and Cats*. 3rd ed. Elsevier; 2014.
Koeppen BM, Stanton BA. *Berne & Levy Physiology*. 7th ed. Elsevier; 2017.
Lavoie J, Hinchcliff KW, eds. *Blackwell's Five-Minute Veterinary Consult: Equine*. 2nd ed. Wiley-Blackwell; 2009.
Orsini JA, Divers TJ, eds. *Equine Emergencies: Treatment and Procedures*. 4th ed. Elsevier; 2013.
Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.
Reed SM, Bayly WM, Sellon DC. *Equine Internal Medicine*. 4th ed. Elsevier; 2017.
Smith BP, ed. *Large Animal Internal Medicine*. 5th ed. Elsevier; 2014.
Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2015.
Sprayberry KA, Robinson NE. *Robinson's Current Therapy in Equine Medicine*. 7th ed. Elsevier; 2014.
Tilley LP, Smith FWK, eds. *Blackwell's Five-Minute Veterinary Consult: Canine and Feline*. 6th ed. Wiley-Blackwell; 2015.

Vet-ebooks.com

25

Neural and Hormonal Control of Blood Pressure and Blood Volume

ROBERT B. STEPHENSON

KEY POINTS

1. Neural-hormonal mechanisms regulate blood pressure and blood volume to ensure adequate blood flow for all body organs.
2. The autonomic nervous system affects the cardiovascular system through the release of epinephrine, norepinephrine, and acetylcholine.
3. The arterial baroreceptor reflex regulates arterial blood pressure.
4. The atrial volume receptor reflex regulates blood volume and helps stabilize blood pressure.
5. The cardiovascular state of conscious subjects is determined by an ongoing and ever-changing mixture of reflex effects and psychogenic responses.

Neural-Hormonal Mechanisms Regulate Blood Pressure and Blood Volume to Ensure Adequate Blood Flow for All Body Organs

The influences of the nervous system and hormones on the cardiovascular system are referred to collectively as the *neural-hormonal mechanisms* of cardiovascular control. The neural-hormonal mechanisms are also called *extrinsic control mechanisms* because they act on organs from outside. As described in Chapter 24, the mechanisms of cardiovascular control that act locally, within individual organs and tissues, are referred to as *intrinsic control mechanisms*. The local, or intrinsic, mechanisms predominate over extrinsic mechanisms in the control of blood flow to the "critical" organs, which include the heart (i.e., coronary circulation), the brain, and working (exercising) skeletal muscle. In contrast, neural-hormonal, or extrinsic, control mechanisms predominate over the intrinsic mechanisms in the control of blood flow to the "noncritical" organs, which include the kidneys, the splanchnic organs, and resting skeletal muscle. In this context, "noncritical organs" are those that can withstand temporary reductions in blood flow (and metabolism) to make extra blood flow available for the critical organs, whose optimal function on a moment-to-moment basis may be necessary for survival (e.g., in a life-threatening situation involving "fight or flight").

Neural-hormonal mechanisms also control the heart rate and cardiac contractility. This allows cardiac output to be adjusted to provide adequate blood flow for all the systemic organs, or at least for the critical organs. An important distinction is that cardiac muscle is under neural-hormonal control, whereas the coronary blood vessels are primarily under local control. When neural-hormonal mechanisms increase the heart rate and cardiac contractility, the cardiac metabolic rate also increases. The increased metabolic rate acts via local metabolic control mechanisms to dilate coronary arterioles, which increases coronary blood flow.

To appreciate the importance of neural-hormonal control mechanisms, consider what would happen in their absence. For example, what would occur during exercise if all the body organs relied on local control mechanisms to adjust their blood flow? At the onset of exercise, metabolic control mechanisms would cause vasodilation in the exercising skeletal muscles. Vascular resistance would decrease in the exercising muscles, and the blood flow through these muscles would increase. However, decreasing the vascular resistance in skeletal muscles would lower the *total peripheral resistance* (TPR). As a consequence, arterial blood pressure would decrease. This would decrease the perfusion pressure for all the systemic organs, and blood flow would therefore decrease below normal levels in the brain, kidneys, splanchnic organs, and so forth. The decreased blood flow in these organs would trigger autoregulatory responses, and these organs would vasodilate. However, that vasodilation would lower the TPR even further, which would reduce arterial pressure even more. This in turn would limit the increase in skeletal muscle blood flow. The end result would be some increase in blood flow in the exercising muscle and decreased blood flow elsewhere, but none of the organs (including skeletal muscle) would receive sufficient blood flow to meet their metabolic needs. Arterial pressure would be dangerously low, and the animal would exhibit profound exercise intolerance.

Neural-hormonal control mechanisms allow an animal to avoid these complications. As exercise begins, cardiac output is increased sufficiently to meet the increased need for blood flow in the exercising muscle (and in the coronary circulation) while keeping all the other organs supplied with a normal blood flow. If cardiac output cannot be increased sufficiently to meet all these needs (e.g., during severe exercise), the control mechanisms take the additional step of temporarily reducing blood flow in the noncritical organs and making this extra flow available to the critical organs.

How do the neural-hormonal control systems "know" when cardiac output is sufficiently high to meet the needs of all the organs and when to initiate vasoconstriction in the noncritical organs? An indirect strategy is used: cardiac output is increased enough to keep arterial pressure at a normal level. As long as arterial pressure is maintained at the normal level, local metabolic control mechanisms can successfully match blood flow to metabolic

Vet-ebooks.com

need in each individual organ. If cardiac output cannot be sufficiently increased to keep arterial pressure from falling, neural-hormonal mechanisms initiate vasoconstriction in the noncritical organs. In effect, neural-hormonal control mechanisms will deprive noncritical organs of an ideal level of blood flow if the critical organs are in need of more blood flow than can be supplied by the heart.

There are many important neural-hormonal control mechanisms, but four are emphasized in the following presentation. The first two of these are *cardiovascular reflexes.* The *arterial baroreceptor reflex* works to regulate arterial pressure through the continual adjustment of cardiac output and vascular resistance (in the noncritical organs). The *atrial volume receptor reflex* works in conjunction with the arterial baroreceptor reflex to regulate arterial pressure and to adjust cardiac preload. The other two neural-hormonal mechanisms described in this chapter are the *defense-alarm reaction* (the "fight or flight response") and *vasovagal syncope* (the "playing dead" reaction). These responses exemplify *psychogenic influences* on the cardiovascular system.

The Autonomic Nervous System Affects the Cardiovascular System Through the Release of Epinephrine, Norepinephrine, and Acetylcholine

The autonomic nervous system is the "neuro" arm of neural-hormonal control. Sympathetic and parasympathetic neurons influence the cardiovascular system through the release of the neurotransmitters norepinephrine and acetylcholine. In addition, sympathetic nerves affect the cardiovascular system by stimulating the release of epinephrine and norepinephrine from the adrenal medulla. The adrenal secretions enter the bloodstream as hormones and circulate throughout the body. Chapter 13 contains additional, basic information about the autonomic nervous system.

Whether acting as neurotransmitters or as hormones, epinephrine, norepinephrine, and acetylcholine exert their cardiovascular effects by activating receptor proteins located in the membranes of cardiac muscle cells, or of smooth muscle cells (or in some cases the endothelial cells) in the walls of blood vessels. The receptors activated by epinephrine and norepinephrine are called *adrenergic receptors* (named after the *adrenal gland*). There are two major types: *α-adrenergic receptors* and *β-adrenergic receptors.* The α-adrenergic receptors are subdivided into α_1 and α_2. There are three subtypes of β-receptors: β_1, β_2, and β_3, with the first two of these being important in cardiovascular control.

Acetylcholine activates *cholinergic receptors.* There are two major types: *muscarinic cholinergic receptors* and *nicotinic cholinergic receptors.* The main cardiovascular effects of acetylcholine are mediated through muscarinic cholinergic receptors located on cardiac muscle, smooth muscle, or endothelial cells. Of five subtypes of muscarinic receptors, the M_2 and M_3 receptor subtypes have the greatest cardiovascular importance.

Table 25.1 summarizes the cardiovascular consequences of the activation of adrenergic and cholinergic receptors. α-adrenergic receptors (both α_1 and α_2) are located in the cell membranes of the smooth muscle cells of the arterioles in all organs and in the smooth muscle cells of the abdominal veins. These adrenergic receptors are innervated by postganglionic sympathetic neurons, which release the neurotransmitter norepinephrine. Circulating epinephrine or norepinephrine can also activate the adrenergic receptors. Activation of these α-adrenergic receptors leads to constriction of the arterioles or the veins.

Arteriolar vasoconstriction increases vascular resistance and decreases the blood flow through an organ. If one or more major body organs are vasoconstricted, the TPR increases. TPR (along with cardiac output) determines arterial blood pressure, so widespread α-adrenergic vasoconstriction in the body leads to an increase in arterial blood pressure. The increase in arterial pressure increases the driving force for blood flow in all organs of the systemic circulation. In effect, the sympathetic nervous system can vasoconstrict some organs and thereby direct more blood flow to other nonvasoconstricted organs.

The major role of veins is to act as reservoirs for blood. *Venoconstriction* displaces venous blood toward the central circulation, which increases central venous pressure, right ventricular preload, and (by the Starling mechanism) stroke volume. Venoconstriction in the abdominal organs (including the spleen) is particularly effective in increasing central venous pressure, because these organs typically hold a large volume of venous blood. Venoconstriction causes a relatively small increase in the resistance to blood flow through an organ because the veins, whether dilated or constricted, offer much less resistance to blood flow than do the arterioles.

Sympathetic control of the heart is exerted through the β_1-adrenergic receptors, which are found on every cardiac muscle cell. These beta receptors are activated by norepinephrine or epinephrine. Chapters 19 and 21 discuss the effects of activation of the cardiac β-adrenergic receptors. In brief, β-activation increases pacemaker rate, speeds cell-to-cell propagation of action potentials, and decreases the refractory period. Also, contractility is increased, so the cardiac contractions are quicker and stronger. The overall effect is increased heart rate and increased stroke volume.

β_2-adrenergic receptors are found on the arterioles, particularly in the coronary circulation and in skeletal muscles. The activation of arteriolar β_2-adrenergic receptors causes relaxation of the vascular smooth muscle and dilation of the arterioles. However, these β_2-adrenergic receptors are not innervated by the sympathetic nervous system, so they are not activated directly by sympathetic nerves. Instead, they respond to circulating epinephrine and norepinephrine (released from the adrenal medulla). The adrenal medulla releases epinephrine and norepinephrine in situations that involve trauma, fear, or anxiety. Dilation of arterioles in the coronary circulation and in skeletal muscles is appropriate in such "fear, fight, or flight" situations because the dilation results in an anticipatory increase in blood flow to the heart and skeletal muscle. Appropriately for its role in emergency situations, β_2-adrenergic vasodilation can overpower α-adrenergic vasoconstriction in the coronary circulation and in skeletal muscles.

Parasympathetic effects on the heart are mediated via the neurotransmitter acetylcholine, which activates cholinergic muscarinic receptors of the M_2 type. Cardiac muscle cells of the sinoatrial and atrioventricular nodes are densely innervated by postganglionic parasympathetic neurons. Atrial cells also receive strong parasympathetic innervation. In these parts of the heart, activation of cardiac M_2 receptors has effects basically opposite to those of the activation of β_1-adrenergic receptors. Specifically, parasympathetic activation powerfully slows the cardiac pacemakers, slows cell-to-cell propagation of action potentials, and increases the refractory period. Curiously, ventricular muscle cells receive very little direct parasympathetic innervation. Therefore parasympathetic activation has only a minor, direct effect on ventricular contractility. However, parasympathetic neurons do exert an interesting, indirect effect on ventricular muscle cells. Most parasympathetic neurons in the

Vet-ebooks.com

TABLE 25.1 Receptors Involved in Autonomic Control of the Cardiovascular System

Receptor Type	Location	Usual Activator	Effect of Activation	Function
α Adrenergic				
α_1 and α_2	Arterioles (all organs)	Norepinephrine from sympathetic neurons, or circulating epinephrine and norepinephrine	Vasoconstriction	Decreases blood flow to organs; increases total peripheral resistance **(major effect)**
	Veins (abdominal organs)	Norepinephrine from sympathetic neurons, or circulating epinephrine and norepinephrine	Venoconstriction	Displaces venous blood toward heart
β Adrenergic				
β_1	Heart (all cardiac muscle cells)	Norepinephrine from sympathetic neurons, or circulating epinephrine and norepinephrine	Increased pacemaker rate; faster speed of conduction; decreased refractory period; quicker, stronger contractions	Increases heart rate, stroke volume, and cardiac output **(major effects)**
β_2	Arterioles (coronary and skeletal muscle)	Circulating epinephrine and norepinephrine [β_2 receptors not innervated]	Vasodilation	Increases coronary blood flow; increases skeletal muscle blood flow
Muscarinic Cholinergic				
M_2	Heart (all cardiac muscle cells, but sparse direct innervation of ventricular muscle cells)	Acetylcholine from parasympathetic neurons	Opposite of β_1	Decreases heart rate and cardiac output **(major effect)**
	Sympathetic nerve endings at ventricular muscle cells	Acetylcholine from parasympathetic neurons	Inhibition of norepinephrine release from sympathetic neurons	Decreases magnitude of sympathetic effects on ventricular muscle cells
M_3	Arterioles (coronary)	Acetylcholine from parasympathetic neurons	Vasodilation (mediated via nitric oxide)	Increases coronary blood flow (minor effect)
	Arterioles (genitals)	Acetylcholine from parasympathetic neurons	Vasodilation (mediated via nitric oxide)	Causes engorgement and erection
	Arterioles (skeletal muscle)	Acetylcholine from specialized sympathetic neurons	Vasodilation (mediated via nitric oxide)	Increases muscle blood flow (in anticipation of exercise)
	Arterioles (most other organs)	[Receptors not innervated; normal activator unknown]	Vasodilation (mediated via nitric oxide)	Function unknown

ventricles release their acetylcholine onto sympathetic neuron terminals rather than directly onto ventricular muscle cells. This acetylcholine activates muscarinic cholinergic receptors on the sympathetic neuron terminals, which inhibits the release of norepinephrine from the terminals and thus weakens the effects of sympathetic activity on ventricular cells. By decreasing the heart rate and opposing sympathetic effects on ventricular contractility, parasympathetic activation can profoundly decrease cardiac output.

Cholinergic muscarinic receptors of the M_3 type are found on the endothelial cells and also on the smooth muscle cells of most arteries and arterioles. Activation of M_3 receptors on smooth muscle cells causes them to contract. However, this *vasoconstrictor* effect is usually overridden by a *vasodilatory* effect of activating the M_3 receptors on the vascular endothelial cells. In this strange arrangement, activation of M_3 receptors on endothelial cells causes the synthesis of *nitric oxide,* which then diffuses out of the endothelial cells and into the nearby smooth muscle cells, where it causes vasodilation. The vasodilatory effect of stimulating the M_3 receptors on endothelial cells is stronger than the vasoconstrictor effect of stimulating the M_3 receptors on smooth muscle cells.

The M_3 receptors on vascular endothelial cells are innervated in three organs. Parasympathetic neurons innervate vascular M_3 receptors in the coronary circulation, where the effect of parasympathetic activation is vasodilation. This vasodilatory effect is minor, however, and the function of this innervation is unclear. In the blood vessels of the external genital organs, parasympathetic neurons release both acetylcholine and nitric oxide. The acetylcholine activates M_3 receptors on the endothelial cells to stimulate the release of additional nitric oxide from endothelial cells. The nitric oxide relaxes vascular smooth muscle, which causes vasodilation, engorgement of the organs with blood, and therefore erection. The third tissue in which vascular M_3 receptors are innervated is skeletal muscle. In some species (e.g., cats and dogs) but not in others (e.g., primates), the M_3 receptors of skeletal muscle blood vessels are innervated by special postganglionic sympathetic neurons that release acetylcholine as a neurotransmitter (rather than the usual norepinephrine). These *sympathetic cholinergic neurons* appear to be activated specifically in anticipation of muscular exercise and during the "fear, fight, or flight" (defense-alarm) reaction. The resulting vasodilation increases blood flow through the skeletal

Vet-ebooks.com

muscle just before and during the initiation of exercise. Although primates do not have sympathetic cholinergic vasodilatory nerves, they can bring about an anticipatory vasodilation of skeletal muscle arterioles through activation of β_1-adrenergic receptors by circulating epinephrine and norepinephrine, as mentioned earlier.

To summarize, arteries and arterioles throughout the body have M_3 adrenergic receptors, and these blood vessels dilate when exposed to acetylcholine (with nitric oxide serving as the mediator). But acetylcholine-releasing autonomic neurons only innervate the blood vessels in the heart, the external genitalia, and (in some species) skeletal muscle. The functional significance of the M_3 receptors on arteries and arterioles in other organs is unclear because no neurons (either sympathetic or parasympathetic) appear to innervate them, and neither acetylcholine nor any other muscarinic receptor agonist normally circulates in the bloodstream.

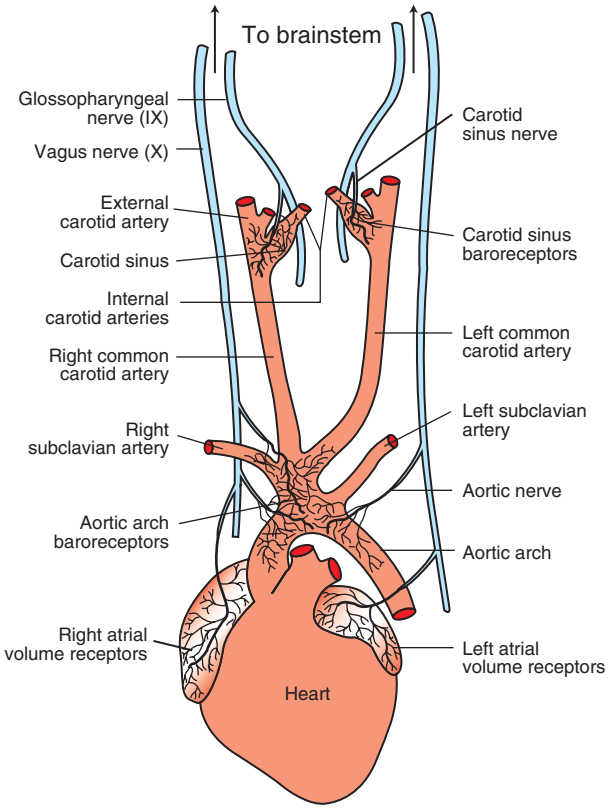
Of all the autonomic influences on the cardiovascular system just discussed, three stand out as most important. The first is α_1- and α_2-adrenergic vasoconstriction in the arterioles of all body organs, which is brought about by the sympathetic nervous system. The second is β_1-adrenergic excitation of cardiac muscle, which is brought about by the sympathetic nervous system and results in an increased heart rate and stroke volume. The third is a decrease in heart rate brought about by parasympathetic activation of cardiac M_2 receptors.

The Arterial Baroreceptor Reflex Regulates Arterial Blood Pressure

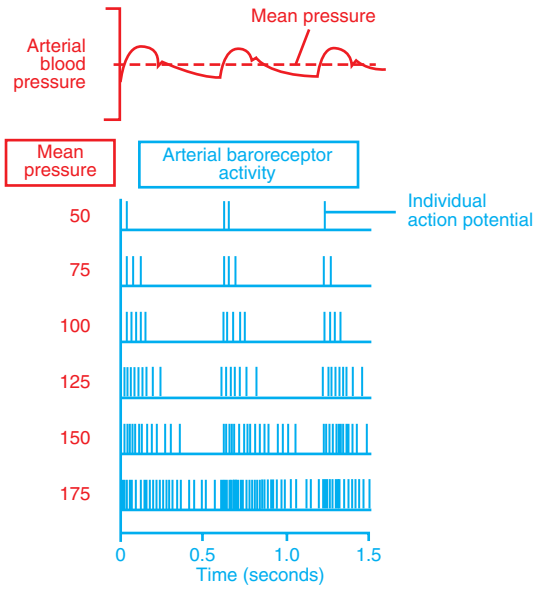
Arterial blood pressure is monitored by pressure-sensitive nerve endings known as *baroreceptors.* The baroreceptors send afferent impulses to the central nervous system (CNS), which reflexively alters cardiac output and vascular resistance (in noncritical organs) to keep blood pressure at a set point. The reflex is called the *arterial baroreceptor reflex.*

The arterial baroreceptors are specialized nerve endings that are embedded in the walls of the carotid arteries and aortic arch (Fig. 25.1). The baroreceptors are concentrated at the origin of each internal carotid artery in enlarged parts of the arteries called the *carotid sinuses.* Similar nerve endings are found in the wall of the aortic arch, especially at the origin of its major branches. These nerve endings are sensitive to stretch (*distention*) of the arterial wall. In effect, they sense arterial pressure because blood pressure is the natural force that distends these arteries. Therefore these nerve endings are called *baroreceptors* (literally, "pressure sensors") even though the actual physical factor being sensed is not pressure but rather stretch.

With every systolic ejection from the heart, blood distends the aorta and arteries, including the carotid sinuses, which causes the baroreceptors to initiate neural impulses (action potentials). Fig. 25.2 illustrates that the frequency of these action potentials is proportional to the arterial blood pressure. The tracing on the top shows the pulsatile arterial pressure on three successive heartbeats. The mean level of arterial pressure is indicated by the dashed line. The tracings below depict the typical patterns of action potentials that would be seen in a baroreceptor afferent neuron for various levels of *mean arterial pressure* (MAP). When MAP is lower than normal (e.g., 50 mm Hg), there are only one or two action potentials with each heartbeat. These action potentials occur during the rapid upstroke of the pressure wave, because the baroreceptors are sensitive to the rate of change of pressure as well as to mean pressure. When MAP is at a higher level (e.g., 75 mm Hg), more action potentials

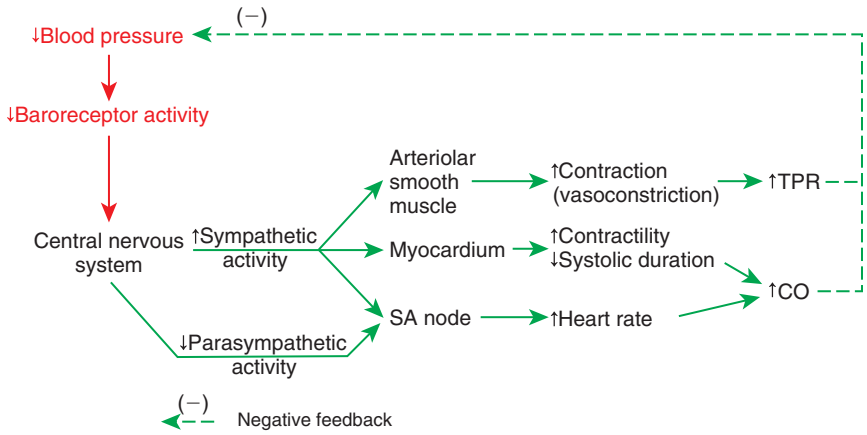


• **Fig. 25.1** Arterial baroreceptors are located in the walls of the carotid sinuses and in the walls of the aortic arch and its major branches. The atrial volume receptors are located in the walls of the right and left atria. See text for a description of the neural pathways followed by the baroreceptor and volume receptor afferents.



• **Fig. 25.2** Each arterial pressure pulse (*top trace, in red*) causes action potentials to be generated in baroreceptor afferent neurons (*lower traces, in blue*). The number of action potentials generated per heartbeat increases dramatically with increases in mean arterial pressure.

Vet-ebooks.com



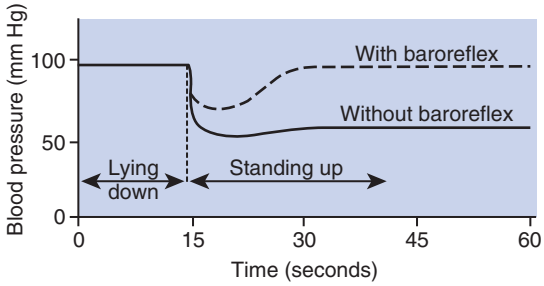
• **Fig. 25.3** The arterial baroreceptor reflex responds to decreases in blood pressure (*top left*) by increasing cardiac output (*CO*), total peripheral resistance (*TPR*), or both (*far right*). These reflex effects offset the initial fall in blood pressure *(dashed line). SA,* Sinoatrial.

are formed during each heartbeat, but the action potentials still tend to occur during the rapid pressure increase at the beginning of the cardiac ejection. The higher the MAP, the more action potentials are formed in each heartbeat. Thus the arterial baroreceptors signal increases in pressure by increasing their action potential frequency. Because the baroreceptors are active when arterial pressure is normal (MAP near 100 mm Hg), they can also signal a decrease in arterial pressure by decreasing their action potential frequency.

The afferent neurons from the aortic arch baroreceptors join the vagus nerves (see Fig. 25.1). In some species, the aortic baroreceptor afferents form a distinct bundle within the vagal nerve sheath, called the *aortic depressor nerve.* The stretch receptors in the carotid sinuses have their afferents in the carotid sinus nerves (Hering's nerves), which join the glossopharyngeal (ninth cranial) nerve. By way of these afferent neurons, the brain receives beat-by-beat information about the level of arterial blood pressure.

Fig. 25.3 summarizes the reflex consequences of a decrease in blood pressure, which decreases afferent baroreceptor activity. The brain responds to a decrease in the afferent activity from the baroreceptors by increasing sympathetic activity. In the heart, sympathetic activation results in increased stroke volume and heart rate, which increases cardiac output. The increase in cardiac output helps restore blood pressure toward normal. The sympathetically driven increase in heart rate is augmented by a simultaneous reduction in parasympathetic activity to the sinoatrial node. Thus the baroreceptor reflex uses reciprocal changes in sympathetic and parasympathetic activity to control heart rate. Sympathetic activity is also increased to the arterioles of all organs, but the consequent vasoconstriction is most pronounced in the noncritical organs (kidney, splanchnic organs, and resting skeletal muscle) because these are the organs in which neural-hormonal control of arterioles predominates over local (metabolic) control. Vasoconstriction in the noncritical organs increases the resistance to blood flow through these organs and therefore increases TPR. The increase in TPR helps restore arterial blood pressure back toward its normal level. In effect, the increased resistance in the noncritical organs helps preserve adequate blood flow in the critical organs.

To fully understand the function of the baroreceptor reflex, it is important to recognize that the reflex does not *reverse* disturbances that alter blood pressure but only *minimizes* their effects on blood pressure. Also, it is important to distinguish between cause and effect when thinking about the baroreflex. What *causes* blood pressure to decrease *below normal* is a decrease *below normal* in cardiac output, TPR, or both. *There is no other way to lower blood pressure.* If the initiating disturbance is that cardiac output falls *below normal* (e.g., following a hemorrhage), the *compensatory responses* of the baroreflex are: (1) to increase TPR *above normal* by initiating sympathetic vasoconstriction in the noncritical organs and (2) to minimize the decrease in cardiac output by initiating increased sympathetic (and decreased parasympathetic) activation of the heart. After compensation, cardiac output is still *below* normal but not as far below normal as in the uncompensated state. Blood pressure is still below normal but not as far below normal as in the uncompensated state.



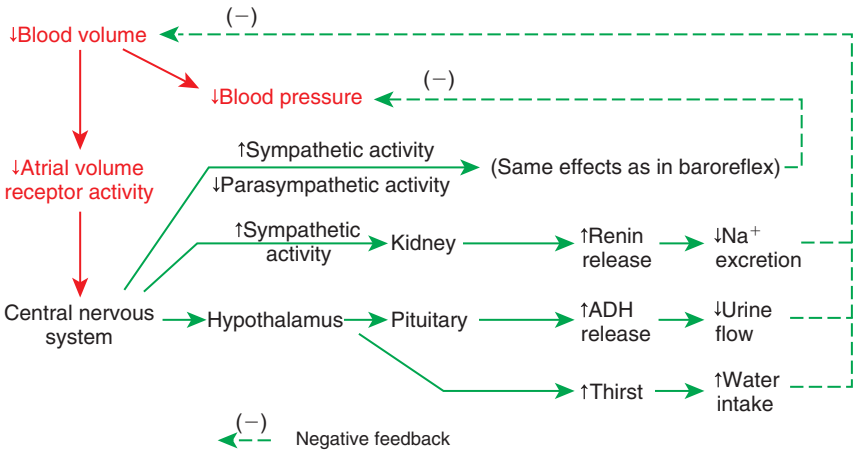
• **Fig. 25.4** The baroreflex is essential for normal, moment-to-moment stability of blood pressure. Dogs in which the baroreflex has been eliminated exhibit much larger swings in blood pressure in response to postural changes than do dogs with an intact baroreflex.

Similarly, if the initiating disturbance is that TPR falls below normal, the compensatory responses of the baroreceptor reflex are to increase cardiac output *above* normal and to restore the low TPR *toward* normal. Blood pressure is still below normal but not as far below normal as in the uncompensated state.

All the reflex responses just described for a decrease in arterial blood pressure occur in reverse in response to an increase in arterial blood pressure above its normal level. Thus the *baroreflex* acts to counteract and minimize both decreases and increases in blood pressure.

The baroreflex responds quickly, initiating compensations for disturbances in blood pressure within 1 second. The reflex is also very powerful. For example, a hemorrhage that would decrease blood pressure by 40 to 50 mm Hg if there were no baroreflex decreases blood pressure by only 10 to 15 mm Hg in an animal with intact baroreflexes. The baroreflex also acts to maintain blood pressure close to normal during changes in posture or activity, as shown in Fig. 25.4. In a dog without baroreflexes, postural changes are accompanied by large, uncontrolled variations in blood pressure. By minimizing fluctuations in blood pressure, the baroreflex helps

Vet-ebooks.com



• **Fig. 25.5** The atrial volume receptor reflex responds to a decrease in blood volume by decreasing sodium and fluid loss in the urine, and by increasing oral water intake. The reflex also helps support blood pressure by increasing cardiac output and total peripheral resistance (similar to baroreflex, as depicted in Fig. 25.3). *ADH,* Antidiuretic hormone.

ensure an adequate driving force for blood flow to the critical organs.

Although the baroreceptor reflex is essential for the moment-to-moment stability of blood pressure, it does not appear to be the major mechanism responsible for setting the long-term level of arterial blood pressure, because baroreceptors slowly adapt or *reset* to the prevailing level of arterial pressure. In other words, the baroreceptors come to accept whatever the prevailing blood pressure is as if it were the normal pressure. For example, in an animal or a human who has been hypertensive for a few days or weeks, the baroreflex functions to regulate blood pressure at the elevated level rather than restore blood pressure toward its normal level. Also, the baroreflex can become reset in a downward direction during a period of sustained hypotension. For example, in chronic heart failure, in which arterial pressure may be below normal for days or weeks, the baroreflex appears to regulate blood pressure at a depressed level rather than push it back toward its normal level.

In summary, the baroreflex responds quickly and powerfully to counteract sudden changes in blood pressure, but it has little influence on the long-term level of blood pressure over days or weeks.

The Atrial Volume Receptor Reflex Regulates Blood Volume and Helps Stabilize Blood Pressure

The *atrial volume receptor reflex* is initiated by specialized sensory nerve endings located primarily in the walls of the left and right atria (see Fig. 25.1). These nerve endings are activated by stretch, but they are called *volume receptors* because the volume of blood in each atrium determines how much the atrial wall is stretched. For example, the loss of blood volume through hemorrhage or dehydration results in a decrease in the amount of blood in the major veins and in the atria. When atrial volume decreases, atrial pressure decreases, as does the stretch on the atrial walls. This decreases the frequency of action potentials generated in the atrial stretch receptors. Conversely, increases in blood volume result in increased atrial stretch and an increased frequency of action potentials generated in the atrial stretch receptors. In effect, these atrial stretch receptors are sensitive detectors of atrial blood volume and, indirectly, of total blood volume. Additional stretch-sensitive nerve endings, which act in concert with the atrial volume receptors, are located in the walls of the pulmonary veins.

Fig. 25.5 summarizes the reflex consequences of a decrease in blood volume, which decreases atrial volume receptor activity. The CNS responds reflexively to the decrease in afferent activity from the atrial volume receptors by increasing sympathetic efferent activity to the heart and systemic arterioles, and decreasing parasympathetic efferent activity to the heart. In this respect, the atrial volume receptor reflex and the baroreceptor reflex exert synergistic effects; that is, a decrease in blood volume leads (via the atrial volume receptor reflex) to the same responses that are triggered by the baroreflex in response to a decrease in arterial blood pressure. In both cases, the reflex responses include increased cardiac contractility, decreased systolic duration, increased heart rate, and arteriolar vasoconstriction in the noncritical organs. By initiating these responses, the atrial volume receptor reflex helps combat the decrease in arterial blood pressure that would otherwise result from a decreased blood volume. In this regard, the atrial volume receptor reflex augments the effectiveness of the baroreceptor reflex as a regulator of blood pressure.

The atrial volume receptor reflex acts in three additional ways to help restore lost blood volume (see Fig. 25.5). First, the reflex acts through the hypothalamus to increase the sensation of thirst. If water is available, the animal drinks. This provides the fluid necessary to increase blood volume back toward normal. Second, the atrial volume receptor reflex acts through the hypothalamus and pituitary gland to increase the release of *antidiuretic hormone* (ADH, also known as *arginine vasopressin*). ADH is synthesized in hypothalamic neurons, which transport it to the posterior pituitary gland. From there, ADH is released into the bloodstream (see Chapter 33). ADH acts on the kidneys to decrease urine production. The third way in which the atrial volume receptor reflex helps restore lost blood volume is to stimulate the release of the hormone *renin* from the kidneys. Renin acts to increase the production of the hormone *angiotensin II,* which acts to increase production of the hormone *aldosterone,* which acts to decrease the amount of sodium excreted by the kidneys; that is, activation of the *renin-angiotensin-aldosterone system* causes the body to conserve available sodium.

The combination of decreased sodium excretion (by the actions of renin) and decreased urine flow (by the actions of ADH) results in the conservation of body fluid. The conservation of body fluid, combined with an increased water intake, eventually restores blood volume back toward normal.

Although not diagrammed in Fig. 25.3, the arterial baroreceptor reflex also responds to decreases in arterial pressure by increasing thirst, ADH release, and renin release. An increase in arterial pressure

Vet-ebooks.com

above normal initiates the opposite effects. Thus the arterial baroreceptor reflex and the atrial volume receptor reflex are synergistic partners in the interrelated tasks of regulating arterial pressure and blood volume.

The Cardiovascular State of Conscious Subjects Is Determined by an Ongoing and Ever-Changing Mixture of Reflex Effects and Psychogenic Responses

The baroreceptor reflex and the atrial volume receptor reflex are only two of several important cardiovascular reflexes that affect blood pressure and blood volume. These two reflexes illustrate several properties common to all cardiovascular reflexes. First, these reflexes originate from changes detected by peripheral sensory receptors. Second, the reflexes occur at a subconscious level through neural pathways that primarily involve cardiovascular centers in the brainstem. Consequently, cardiovascular reflexes persist in unconscious and anesthetized subjects, although the strength and character of the reflexes are altered by anesthesia. Finally, the reflexes use sympathetic and parasympathetic neurons as well as hormonal and behavioral responses to bring about cardiovascular changes.

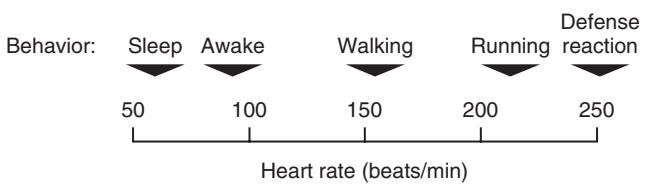
In conscious subjects, neural-hormonal control of the cardiovascular system involves both cardiovascular reflexes and psychogenic effects. Psychogenic responses originate from conscious perceptions or emotional reactions. They are eliminated by unconsciousness or general anesthesia. They involve neural pathways originating in the forebrain, including the limbic system and cerebral cortex, that project to brainstem nuclei involved in cardiovascular control. Psychogenic responses are often triggered by sensory stimuli. For example, the sights, sounds, and smells of a veterinary clinic may trigger perceptions and emotions that cause increases in heart rate and blood pressure in both animal patients and in their human companions. Psychogenic responses can also occur without any obvious sensory triggers. For example, anxiety about a future event can increase heart rate and blood pressure, at least in humans. Cardiovascular reflexes and psychogenic reactions use the same sympathetic and parasympathetic neurons and some of the same hormones to bring about cardiovascular changes.

Two important psychogenic responses are the defense-alarm reaction and vasovagal syncope (the "playing dead" reaction). The *defense-alarm reaction* ("fear, fight, or flight" response) is a collection of emotional, behavioral, and autonomic neural-hormonal responses to a threatening situation, physical injury, or trauma. The cardiovascular component of the defense-alarm reaction involves increased sympathetic activity and decreased parasympathetic activity. Typically, the sympathetic activation is sufficiently strong to cause the release of epinephrine and norepinephrine from the adrenal medulla. The cardiovascular responses during a defense-alarm reaction therefore include an increased heart rate, increased stroke volume, vasoconstriction in noncritical organs (kidneys, splanchnic organs, resting skeletal muscle), vasoconstriction in skin, vasodilation in coronary vessels and in working skeletal muscle, and increased blood pressure. The cardiovascular responses during the defense reaction are enhanced by other circulating hormones, including ADH and angiotensin II, which cause arteriolar vasoconstriction in the non-critical organs, in addition to their effects on salt and water excretion by the kidneys. The resulting elevated blood pressure helps ensure adequate blood flow for the critical organs (exercising skeletal muscles, heart, and brain).

During a defense-alarm reaction, the baroreceptor reflex is reset by the CNS so that it regulates blood pressure at an elevated level rather than acting to oppose the increased pressure. This is analogous to resetting the cruise control on a car so that it regulates speed at an elevated level rather than acting to oppose an increased speed. Thus it is more accurate to say that the baroreceptor reflex regulates blood pressure at a variable set point (set by the CNS) than to say that the baroreflex regulates arterial pressure at any single "normal" pressure.

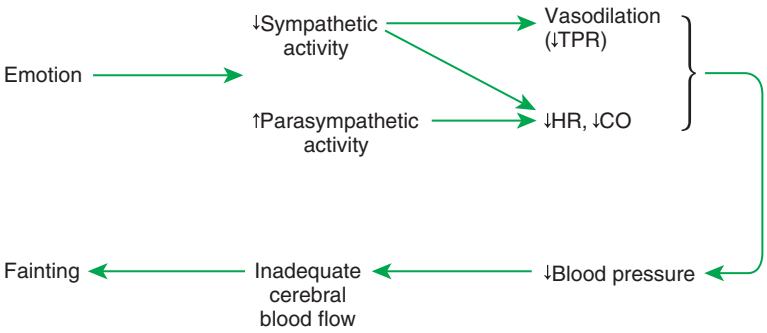
It is important to recognize that the defense-alarm reaction is simply the extreme form of a continuum of states of emotional arousal. Sleep is at the opposite end of this cardiovascular and emotional continuum. In quiet rest or sleep, sympathetic activity is minimal and parasympathetic activity is maximal. During a full-blown defense-alarm reaction, sympathetic activity is maximal and parasympathetic activity is minimal. Between these extremes lie all the levels of emotional arousal experienced by animals and humans, from moment to moment, during ordinary and extraordinary daily activities. Cardiovascular variables, such as heart rate and blood pressure, are sensitive to these changes in emotional state (Fig. 25.6). For example, a large dog may normally have a heart rate of 70 beats/min while resting at home, but it would be entirely normal for the same dog to have a heart rate of 120 beats/min while "resting" in a veterinary clinic, if the dog is apprehensive in that setting. Another important point for the clinician to remember is that emotional responses are subjective. Situations that severely agitate one animal may cause only a mild alerting response in another animal. The clinician must evaluate heart rate, blood pressure, and other cardiovascular signs with respect to the particular patient's emotional state.

Vasovagal syncope is another psychogenic response that may be encountered in veterinary practice. This response is also called "playing dead" or "playing possum" (i.e., behaving like an opossum). In response to certain threatening or emotional situations, some humans and animals experience a psychogenic *decrease* in blood pressure and may faint. In many ways, this response is the opposite of the defense-alarm reaction. As shown in Fig. 25.7, vasovagal syncope involves a decrease in sympathetic activity and an increase in parasympathetic activity. These neural changes bring about a vasodilation in the noncritical organs, with a consequent decrease in TPR. Heart rate and cardiac output also decrease, so there is a large drop in arterial blood pressure. The expected compensatory reflex responses do not take place because emotional state appears to override the baroreceptor reflex in this case. If blood pressure falls so low that there is inadequate cerebral blood flow, the patient faints. The term *vasovagal syncope* denotes *vaso*dilation, *vagal* (parasympathetic) activation, and *syncope* (fainting). It is not clear why some animals respond to a threatening situation with a defense-alarm reaction, whereas others exhibit vasovagal syncope.



• **Fig. 25.6** The defense-alarm reaction is simply the extreme on a continuum of emotional and physical arousal. Cardiovascular variables (e.g., heart rate, plotted here for a typical large dog) respond sensitively to every change along this arousal scale.

Vet-ebooks.com



• **Fig. 25.7** Vasovagal syncope ("playing dead" reaction) is an emotional response that involves decreases in sympathetic activity and increases in parasympathetic activity. See text for full description. *CO,* Cardiac output; *HR,* heart rate; *TPR,* total peripheral resistance.

CLINICAL CORRELATIONS

Intraoperative Hemorrhage

History

Four hours after abdominal surgery for a splenic sarcoma, a 30-kg, 9-year-old male Labrador retriever is observed to be severely lethargic and recumbent. An abnormally large amount of blood had been lost during the surgical removal of the spleen because the dog had a hereditary blood-clotting defect (von Willebrand's disease).

Clinical Examination

The dog's gums are pale, and his capillary refilling time is abnormally prolonged (3 s). His extremities are cool to the touch. The femoral pulse is rapid and weak. An electrocardiogram indicates sinus tachycardia at a rate of 185 beats/min. The hematocrit (packed cell volume) is 38%, and the plasma protein concentration is 5.6 g/dL; both values are below normal. A jugular catheter is inserted, and central venous pressure is measured and found to be –1 mm Hg (normal, +1 to +3 mm Hg). Despite the intravenous administration of 600 mL of lactated Ringer's solution during surgery, the dog has not produced any urine. About 100 mL of blood-tinged fluid is removed from the abdomen by *abdominocentesis.*

Comment

This case illustrates the typical clinical signs following hemorrhage. Most of the blood in a dog is in the systemic veins, so most of the blood missing after hemorrhage is missing from the veins. The result is an abnormally low central venous pressure as observed in this dog. The decreased central venous pressure causes a decreased ventricular preload and a decreased ventricular end-diastolic volume. This leads to decreases in stroke volume (Starling's law of the heart), cardiac output, and arterial blood pressure. Inadequate cardiac output and blood pressure lead to behavioral depression.

Neural-hormonal compensation for the hemorrhage is initiated by the atrial volume receptor reflex and the arterial baroreceptor reflex. Heart rate is increased by the combination of increased sympathetic activation and decreased parasympathetic activation. The elevated sympathetic activity would also be increasing cardiac contractility. Without this sympathetic drive, stroke volume would be even lower than it is. The combination of high heart rate and low stroke volume accounts for the rapid but weak (low pulse pressure) femoral pulse. Sympathetic activity also causes vasoconstriction in the mucous membranes, resting skeletal muscle, splanchnic organs, and kidneys (noncritical organs and tissues). The reduced blood flow in these tissues accounts for the pale gums, the slow refilling of capillaries, the cool limbs, and the lack of urine production by the kidneys. Urine formation by the kidneys is also being depressed by the combined hormonal effects of ADH and the renin-angiotensin-aldosterone system.

Hemorrhage does not directly reduce either the hematocrit or the plasma protein concentration, because whole blood is being lost. However, two factors caused the hematocrit and plasma protein concentration to decrease in this dog. First, the fluid administered intravenously during surgery (lactated Ringer's solution) contained neither red blood cells nor plasma proteins, so the cells and proteins remaining in the bloodstream were diluted by the addition of the fluid. Second, the hemorrhage reduced not only venous and arterial pressures but also capillary hydrostatic pressure, which changed the balance of hydrostatic and oncotic forces (Starling forces) across the capillary walls in favor of reabsorption. The interstitial fluid reabsorbed into the bloodstream contained no red blood cells and almost no plasma proteins. This caused further dilution of the cells and proteins in the blood.

Treatment

Therapy for this dog involves measures to stop ongoing blood loss and to restore the lost blood volume. In this dog, the hemorrhage is predominantly seepage from small intraabdominal vessels at the surgical site as a result of the coagulation defect. Transfusions of donor blood or plasma, or concentrated preparations of clotting proteins, would promote clotting and limit subsequent hemorrhage. After measures are taken to promote clotting, additional crystalloid solutions (e.g., lactated Ringer's solution) can be infused into this dog because the hematocrit and plasma protein concentration are not dangerously low. However, if crystalloid solutions are administered, the hematocrit and plasma protein concentration should be monitored closely to avoid the hypoxia that results from too much dilution of the red blood cells or the edema that results from too much dilution of the plasma proteins. Renal function should be monitored closely because the combination of hypoxia and reflex vasoconstriction can lead to ischemic damage of kidney tissue, resulting in renal failure.

Three-Day-Old Heifer Calf With Diarrhea

History

A dairy producer calls you about a 3-day-old Jersey heifer calf that began to have diarrhea yesterday. The diarrhea has gotten progressively worse today, and the calf is lethargic and laying down frequently. The calving was observed, and the newborn calf seemed normal. The producer thinks the calf received colostrum. The calf has continued to nurse, although with decreasing frequency. To boost the calf's fluid intake, the producer gave it a supplemental bottle of replacement milk twice yesterday and once today. He started the calf on antibiotics this morning.

Physical Examination

The calf's temperature is below normal. Both heart rate and respiratory rate are markedly decreased. Mucous membranes are injected and capillary refill time (CRT) is prolonged, indicating sluggish circulation. Distal extremities are cool to the touch, and peripheral pulses are decreased, providing further evidence of poor perfusion and suggesting low arterial pressure. After being pinched, the calf's skin remains "tented", which indicates dehydration. The calf's gastrointestinal motility is increased, and she has some watery diarrhea while you are doing your exam.

Comment

The calf appears to be septic. Most likely the calf inhaled or ingested bacteria, and the bacteria and their toxic products then invaded both the intestinal tract and the bloodstream. The resulting intestinal inflammation

Continued

Vet-ebooks.com

CLINICAL CORRELATIONS—cont'd

(*enteritis*) caused diarrhea, which led to a substantial loss of body fluid volume (*dehydration*). The loss in body fluid led to a decrease in blood volume and consequently decreased central venous pressure, decreased cardiac filling, and decreased cardiac output. In this respect, the effects of dehydration are much like those of hemorrhage (see preceding Clinical Correlation on "Intraoperative Hemorrhage"). However, in hemorrhage the fluid (plasma) lost from the body has the same electrolyte composition as normal extracellular fluid, whereas the electrolyte composition of diarrhea fluid differs from that of normal body fluid. In consequence, the remaining body fluid not only has less volume than normal but also comes to have abnormal concentrations of Na^+, K^+, Ca^{++}, Cl^-, HCO_3^-, and H^+. For example, in regard to H^+, the consequence is acidosis. These electrolyte abnormalities have effects throughout the body, but in terms of cardiovascular physiology, the main effect is the depression of cardiac performance. Moreover, the bacteria and bacterial toxins in the blood (*septicemia*) have a direct, depressing effect on cardiac muscle. Therefore the cardiac output in this calf is imperiled not only by low preload (as in the case of hemorrhage) but also by myocardial failure secondary to electrolyte abnormalities and septicemia. The result is a severely limited cardiac output, low arterial pressure, and poor perfusion (ischemia) throughout the systemic circulation. The condition is called *septic shock*. The eventual consequence is multiorgan failure.

The initial neural-hormonal compensations for diarrheal dehydration are initiated by the atrial volume receptor reflex and the arterial baroreceptor reflex, just as in the case of hemorrhage (see preceding Clinical Correlation). The problem in septic shock is that the ability of these reflexes to drive the heart to increased rate and contractility is overwhelmed by the direct cardiac-depressing effects of the electrolyte abnormalities and septicemia. The calf has low heart rate and low cardiac contractility not because of the neural-hormonal compensations but in spite of them.

Treatment

While instituting treatment as outlined in the following, blood tests would be done on this calf to help ascertain the animal's electrolyte status (e.g., basic metabolic panel), immunological state (e.g., immunoglobulin level), character and severity of infection (e.g., differential white blood cell count), severity of inflammation (e.g., increased fibrinogen), and kidney function (e.g., creatinine or blood urea nitrogen). Antibiotic therapy should be continued. Intravenous (i.v.) fluids (e.g., lactated Ringers) should be given to help hydrate the calf and to correct electrolyte deficits, including acidemia. Administration of plasma will improve the oncotic pressure; also, plasma has globulins and nutrients to help combat the sepsis. The calf could also be treated with epinephrine to help improve heart rate and cardiac output, and to increase coronary blood flow and skeletal muscle blood flow, through β-adrenergic receptor activation. Nursing should be maintained if possible. Supplemental bottles of milk or milk replacer may be warranted, depending what how much she is drinking. If she is too depressed to drink, a feeding tube should be placed and maintained. Close clinical observation is essential. Her prognosis is guarded.

Practice Questions

1. Neural-hormonal control of blood flow is most directly responsible for which of the following responses in a skeletal muscle?
- **a.** Active hyperemia
- **b.** Reactive hyperemia
- **c.** Autoregulation
- **d.** Decreased blood flow during a period of mechanical compression
- **e.** Increased blood flow in anticipation of exercise

2. The normal reflex responses to a sudden increase in arterial blood pressure include an increase in:
- **a.** Heart rate.
- **b.** Renin release from the kidney.
- **c.** Parasympathetic activity directed to the heart.
- **d.** Sympathetic activity directed to blood vessels in kidneys, resting skeletal muscles, and splanchnic organs.
- **e.** Sympathetic activity directed to blood vessels in brain, coronary circulation, and exercising skeletal muscles.

3. Vasovagal syncope:
- **a.** Involves decreased blood pressure and heart rate.
- **b.** Involves increased sympathetic activity.
- **c.** Involves decreased cardiac parasympathetic activity.
- **d.** Involves constriction of splanchnic arterioles.
- **e.** Prepares an animal for "fight or flight."

4. Which of the following would prevent the dilation of arterioles in an exercising muscle?
- **a.** Surgically cutting all the autonomic nerves that innervate the muscle.
- **b.** Administering a drug that stimulates the release of epinephrine and norepinephrine from the adrenal glands.
- **c.** Administering a drug that blocks muscarinic cholinergic receptors.
- **d.** Administering a drug that blocks β-adrenergic receptors.
- **e.** None of the above is correct.

5. A drug is injected intravenously into a dog and causes a transient increase in mean arterial pressure and a transient decrease in heart rate. The baroreceptor nerves are then cut and the drug is reinjected. This time, the drug causes a larger increase in blood pressure but no change in heart rate. These results are most consistent with the primary action of the drug being to:
- **a.** Activate the muscarinic cholinergic (M_3) receptors of arterioles.
- **b.** Activate the α-adrenergic receptors of arterioles.
- **c.** Activate β_1-adrenergic receptors of the pacemaker cells of the SA node.
- **d.** Increase the synthesis of nitric oxide in arterioles.
- **e.** Decrease the activity of arterial baroreceptors.

Vet-ebooks.com

6. A dog has suffered a substantial hemorrhage. The dog's heart rate is increased above normal, and the skin is cold. The mucous membranes are pale, and capillary refill time is prolonged. In this situation (compared with normal):
- **a.** The baroreceptor nerves are firing at an increased rate.
- **b.** The sympathetic nerves to the heart are firing at a decreased rate.
- **c.** The sympathetic nerves to the blood vessels of the skin and mucous membranes are firing at an increased rate.
- **d.** The parasympathetic nerves to the blood vessels are firing at an increased rate.
- **e.** The release of renin by the kidney is decreased.

7. During a defense-alarm reaction (fight-or-flight response), heart rate increases because of:
- **a.** Increased circulating acetylcholine released from the adrenal glands.
- **b.** Increased activation of the arterial baroreceptors.
- **c.** Increased arterial blood pressure.
- **d.** Increased sympathetic nerve activity.
- **e.** Increased parasympathetic activity.

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.

Chapleau MW, Abboud FM. *Neuro-Cardiovascular Regulation: From Molecules to Man*. New York: New York Academy of Sciences; 2001.

Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. Elsevier; 2017.

Katz AM. *Physiology of the Heart*. 5th ed. Baltimore: Lippincott, Williams & Wilkins; 2010.

Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.

Silverstein D, Hopper K. *Small Animal Critical Care Medicine*. St Louis: Saunders Elsevier; 2009.

Smith B. *Large Animal Internal Medicine*. 4th ed. Mosby Elsevier; 2009.

Smith FWK, Tilley LP, Oyama M, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2015.

Thrall MA, Weiser G, Allison R, Campbell T, eds. *Veterinary Hematology and Clinical Chemistry*. 2nd ed. Wiley-Blackwell; 2012.

Tilley LP, Smith FWK, eds. *Blackwell's Five-Minute Veterinary Consult: Canine and Feline*. 6th ed. Wiley-Blackwell; 2015.

Wehrwein EA, Orer HS, Barman SM. Overview of the anatomy, physiology, and pharmacology of the autonomic nervous system. *Compr Physiol*. 2016;6:1239–1278.

Weiss DJ, Wardrop KJ. *Schalm's Veterinary Hematology*. 6th ed. Ames, Iowa: Wiley-Blackwell; 2010.

Zachary JF, ed. *Pathologic Basis of Veterinary Disease Expert Consult*. 6th ed. Elsevier; 2016.

Vet-ebooks.com

26

Integrated Cardiovascular Responses

ROBERT B. STEPHENSON

KEY POINTS

1. Both Starling's mechanism and the arterial baroreflex help compensate for heart failure.
2. Serious complications secondary to heart failure include exercise intolerance, edema, salt and water retention, kidney failure, uremia, septic shock, and decompensation.
3. The immediate cardiovascular effects of hemorrhage are minimized by compensations initiated by the atrial volume receptor reflex and the arterial baroreceptor reflex.
4. The blood volume lost in hemorrhage is restored through a combination of capillary fluid shifts and hormonal and behavioral responses.
5. In large animals, the transition from a recumbent to a standing posture elicits the same cardiovascular responses as hemorrhage.
6. During the initiation of exercise, local and neural mechanisms interact to increase cardiac output and to deliver more blood flow to exercising muscle.

Chapters 18 to 25 describe the various elements of cardiovascular function and control. An understanding of these individual elements is not sufficient, however, to provide a basis for the diagnosis and treatment of cardiovascular dysfunction. The veterinary clinician must understand the *interaction* of these elements in both normal and abnormal situations. Therefore this chapter discusses three fundamentally important, *integrated* cardiovascular responses: (1) the response to heart failure, (2) the response to hemorrhage, and (3) the response to exercise. In addition to elucidating important, integrated responses, this discussion provides a review and summary of key concepts of cardiovascular physiology.

Both Starling's Mechanism and the Arterial Baroreflex Help Compensate for Heart Failure

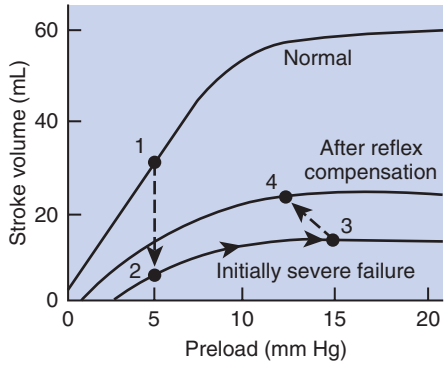
There are many types and causes of *heart failure.* Some clinicians use the term very broadly to refer to any condition in which a problem in the heart limits its ability to deliver enough cardiac output to support the metabolic activity in the body tissues. Such conditions would include various valve defects, arrhythmias, and heartworm infestation. A more restrictive definition, and one favored by physiologists, is that *heart failure* is any condition in which a depressed *cardiac contractility* limits the ability of the heart to deliver an adequate cardiac output. The broader definition of heart failure encompasses virtually any problem with the heart as a pump; a common synonym is *pump failure.* The more restrictive definition, as used in this chapter, equates heart failure with *myocardial failure,* a depressed contractility of the heart muscle itself.

Causes of depressed cardiac contractility include ischemia of the heart muscle, myocarditis, toxins, drug effects, or electrolyte imbalances. If the decrease in contractility affects both sides of the heart, the condition is called *bilateral heart failure.* In other circumstances, failure may be restricted primarily to either the left ventricle or the right ventricle and thus is called *left-sided heart failure* or *right-sided heart failure.*

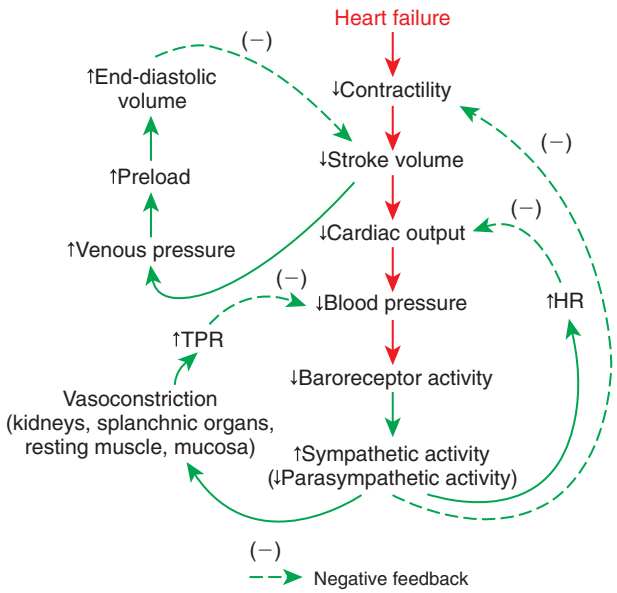
Ventricular function curves are useful for envisioning the consequences of heart failure and the compensations for heart failure. In Fig. 26.1, the curve labeled *Normal* indicates the relationship between stroke volume and preload for a normal ventricle (for a review, see Fig. 21.3C). The curve labeled *Initially severe failure* shows that a ventricle in failure has a depressed contractility (i.e., a smaller stroke volume for any given preload). If a normal heart suddenly goes into severe failure, stroke volume changes from its normal value (shown by point 1) to the low value (shown by point 2). For purposes of illustration, imagine that these curves define the function of the left ventricle and that the left ventricle is the one that fails. A decrease in left ventricular stroke volume causes a decrease in left ventricular output, which results in a decrease in mean arterial blood pressure. Unless there is adequate compensation for this fall in blood pressure, severe exercise intolerance is certain, inadequate perfusion of the critical organs is likely, and death is a strong possibility. However, several mechanisms react rapidly, within seconds to minutes, to compensate for heart failure and to minimize its adverse effects.

One compensation for heart failure is *Starling's mechanism.* If the left ventricle suddenly decreases its stroke volume, the right ventricle (at least for a few heartbeats) maintains a higher stroke volume than the failing left ventricle. The excess blood pumped by the right ventricle has to "go somewhere," and most of the excess accumulates in the pulmonary veins and left atrium. In effect, blood dams up behind (i.e., upstream of) the left ventricle. The resulting increase in left atrial pressure creates an increase in left ventricular preload, which leads to an increase in left ventricular end-diastolic volume and, by Starling's mechanism, a partial correction of the low stroke volume. This improvement in stroke volume is depicted in Fig. 26.1 as a transition from point 2 to point 3. The sequence of events, whereby an increase in preload helps to correct for the initial fall in stroke volume, is also diagrammed in Fig. 26.2 (*top left loop*). Note that the compensation by Starling's mechanism does not return stroke volume to its normal value because contractility remains severely depressed; however, without this compensation, severe heart failure would quickly be fatal.

Vet-ebooks.com



• **Fig. 26.1** Ventricular function curves depicting the consequences and compensations for heart failure in terms of changes in preload (end-diastolic ventricular pressure) and stroke volume. See text for details.



• **Fig. 26.2** Consequences (red arrows) and compensations (green arrows) for heart failure. The changes described here include those presented graphically in Fig. 26.1. See text for details. *HR,* Heart rate; *TPR,* total peripheral resistance.

The arterial baroreflex is another mechanism that reacts rapidly to compensate for heart failure. Even after compensation by Starling's mechanism, left ventricular output remains below normal, as does arterial blood pressure. Therefore baroreceptor activity is below normal. The central nervous system (CNS) responds reflexively by increasing sympathetic efferent activity to the heart and blood vessels, and by decreasing parasympathetic activity to the heart.

The sympathetic effect on the heart increases ventricular contractility. Contractility is not restored to normal but is brought to a higher level than existed in the absence of reflex compensation. Graphically, the effect of the baroreflex is to move the failing ventricle to a function curve that is intermediate between the *Normal* curve and the curve of *Initially severe failure* (see point 4 in Fig. 26.1). Note that the increase in contractility also brings stroke volume closer to (but still not reaching) its normal level.

The reflexive increase in sympathetic activity raises the heart rate above normal and decreases systolic duration (as described in Chapter 21); these changes also help restore cardiac output toward normal despite the persistently depressed stroke volume. Finally, sympathetic activation causes vasoconstriction, particularly in the noncritical organs, which increases total peripheral resistance (TPR) above normal, which helps return arterial pressure toward its normal level, even though cardiac output still remains below normal.

The net effect of the compensations by way of Starling's mechanism and the baroreflex is that arterial blood pressure can be maintained near its normal level, at least when the animal is at rest, despite a severe ventricular failure. Fig. 26.2 summarizes these reflex effects. Note that contractility, stroke volume, and cardiac output will remain at least somewhat below normal, even after compensation by Starling's mechanism and the baroreflex. By contrast, preload, sympathetic activity, heart rate, and TPR are above normal. Arterial pulse pressure will be depressed due to the combination of low stroke volume and high heart rate. The low pulse pressure will cause baroreceptor activity to remain below normal, even if mean arterial pressure (MAP) is restored to, or nearly to, its normal level. This decreased baroreceptor activity helps sustain the reflex increases in sympathetic activity.

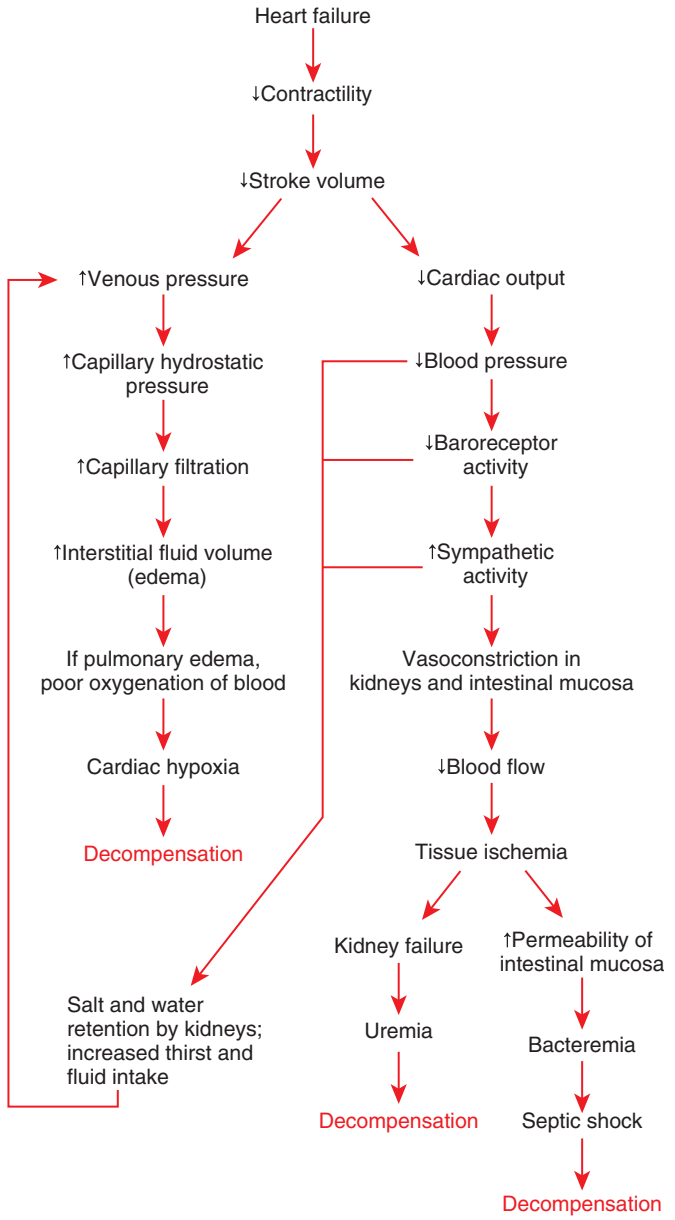
Serious Complications Secondary to Heart Failure Include Exercise Intolerance, Edema, Salt and Water Retention, Kidney Failure, Uremia, Septic Shock, and Decompensation

Even though Starling's mechanism and the baroreflex can compensate to a remarkable degree for severe heart failure, important secondary complications often develop. These complications make heart failure a serious clinical problem, even in cases where compensatory mechanisms maintain MAP at a near-normal level when the animal is at rest.

Heart failure causes *exercise intolerance.* A normal animal can increase cardiac output substantially during exercise by activating sympathetically mediated increases in stroke volume and heart rate. In a patient with heart failure, however, sympathetic activation is being harnessed even at rest to restore the depressed cardiac output toward normal. Therefore the patient has a limited "sympathetic reserve" available (i.e., limited ability to further increase cardiac sympathetic activity) during exercise. In consequence, the failing heart cannot provide the increased cardiac output needed to meet the blood flow requirements of exercising skeletal muscle. That is, metabolic vasodilation in the exercising muscle causes such a large decrease in TPR that arterial pressure decreases. Blood flow then becomes inadequate in all the organs, including the exercising muscle. The patient exhibits lethargy and weakness; even mild exercise leads quickly to exhaustion.

Edema is another serious complication secondary to heart failure. As noted, blood dams up in the atrium and veins behind a failing ventricle. In the case of left ventricular failure, left atrial pressure increases, as does pressure in the pulmonary veins and pulmonary capillaries. The increase in pulmonary capillary hydrostatic pressure leads to an increase in the filtration of capillary fluid into the interstitial spaces of the lungs. *Pulmonary edema* develops. The excess of interstitial fluid slows the transfer of oxygen from the lung alveoli into the lung capillaries and can result in inadequate oxygenation of the blood (*hypoxemia*). In extreme cases, interstitial fluid leaks into the intrapleural space (*pleural effusion*) and into the alveolar air spaces, which causes a further reduction in lung function. The resulting hypoxia in critical organs can be fatal. In a patient with right ventricular failure, the increase in venous pressure occurs in the systemic circulation. Therefore the resulting

Vet-ebooks.com



• **Fig. 26.3** Heart failure leads to exercise intolerance (see text for details). Additional, life-threatening complications secondary to heart failure are diagrammed here, including edema, salt and water retention and increased fluid intake, kidney failure and uremia, and septic shock. Vicious cycles develop in which the effects of heart failure make the heart failure worse (*Decompensation*).

edema occurs in the systemic organs, particularly in dependent extremities and in the abdomen. The cause-and-effect sequence by which heart failure leads to edema is summarized in Fig. 26.3 (*top left*).

Whether the edema is in the lungs or in the systemic circulation, its degree is limited by the three previously discussed safety factors (see Fig. 23.5). These safety factors would probably keep the edema of heart failure well controlled were it not for an additional factor that exaggerates the elevation of venous pressure in heart failure. As long as arterial pressure remains subnormal in a patient with heart failure, the baroreceptor reflex and some additional mechanisms involving the kidneys work to raise blood volume above normal. These volume-increasing mechanisms include increased thirst (which raises fluid intake), increased release of antidiuretic hormone (ADH) from the pituitary (which decreases the amount of fluid lost in the urine), and activation of the renin-angiotensin-aldosterone system (which decreases sodium loss in the urine). These effects of the baroreflex were mentioned briefly in Chapter 25; the additional mechanisms involving the kidneys are described in more detail in Chapters 41 and 43.

The point for now is that the patient with severe heart failure experiences a substantial and persistent increase in blood volume. The excess blood accumulates particularly in the veins upstream from the failing ventricle, which exaggerates the increases in venous and capillary pressure and in capillary filtration. The normal safety factors against edema may be overwhelmed. For this reason, one of the main goals in the clinical treatment of heart failure is to counteract the buildup of excessive blood volume and interstitial fluid volume. *Diuretic drugs* are the main therapies used for this purpose (see Chapter 43).

Severe, persistent heart failure leads to several additional adverse effects. The baroreceptor reflex responds to an abnormally low arterial pressure in heart failure by initiating arteriolar vasoconstriction, primarily in the kidneys, splanchnic organs, and resting skeletal muscle (the noncritical organs). In severe heart failure, the skin and mucous membranes are also vasoconstricted. Vasoconstriction in these organs helps compensate for heart failure by permitting the available cardiac output to be routed to the critical organs (brain, heart, and working skeletal muscle). However, persistent vasoconstriction leads to the additional complications of kidney failure, uremia, and septic shock.

Vasoconstricted kidneys cannot form urine in a normal manner and therefore do not rid the body of the excess volume of blood and interstitial fluid that accumulates in heart failure. Moreover, persistent vasoconstriction damages kidney tissue and leads to a buildup of nitrogenous and acidic waste products in the body. The condition is called *uremia* (literally "urine in the blood"). To make matters worse, after a prolonged period of intense vasoconstriction, damage to the kidney tissue becomes irreversible. At this stage, uremia, acidosis, and salt and water retention may persist even if clinical treatment is temporarily successful in returning cardiac output and blood pressure close to normal. For this reason, *renal failure* often is the terminal event in chronic heart failure.

Intense and prolonged vasoconstriction in the splanchnic circulation can also have lethal consequences. The mucosa of the gastrointestinal tract is particularly susceptible to ischemic damage. Normally, the intestinal mucosa creates a barrier between the contents of the intestinal lumen and the bloodstream. Ischemic damage to the intestinal mucosa allows bacteria and bacterial toxins (endotoxins) to pass into the bloodstream or the peritoneum. The resulting bacteremia or peritonitis can cause septic shock and death. The causes and consequences of renal and splanchnic ischemia are summarized in Fig. 26.3 (*bottom right*).

Cardiac decompensation is an additional (and frequently terminal) complication secondary to heart failure. The basic concept of decompensation is that, when heart failure reaches a certain degree of severity, the body's attempted compensations for heart failure end up making the heart failure worse. Vicious decompensatory cycles develop and can lead to death within a few hours unless there is vigorous medical intervention.

The specific mechanisms of the *decompensatory cycles* are very complex, but three examples illustrate the concept. As previously explained, in the case of left ventricular failure, the damming up of blood in the left atrium is compensatory because it increases left ventricular preload, which helps boost stroke volume back

Vet-ebooks.com

toward normal. However, the increased left ventricular preload leads to the secondary complication of pulmonary edema. If severe, pulmonary edema interferes with the oxygenation of blood. Cardiac muscle is critically dependent on an adequate supply of oxygen; hypoxia depresses the contractility of cardiac muscle. Thus a vicious cycle can develop: severely depressed ventricular contractility → severe pulmonary edema → inadequate oxygenation of blood → hypoxia of the left ventricular muscle → further depression of ventricular contractility.

For a second example of a vicious, decompensatory cycle, consider again the effects of the baroreflex on the kidneys. In a resting animal, renal vasoconstriction is compensatory for heart failure in that it helps increase TPR, which helps bring the low arterial pressure back toward normal, which helps keep perfusion pressure high enough to deliver adequate blood flow to the critical organs. However, as already mentioned, intense and prolonged renal vasoconstriction leads to kidney failure and an accumulation of acidic and nitrogenous waste products in the blood (uremia). Uremia depresses cardiac contractility. Thus another vicious cycle can develop: severe ventricular failure → intense and prolonged renal vasoconstriction → damage to kidney tissues → uremia → metabolic waste products accumulate in cardiac muscle → further depression of ventricular contractility.

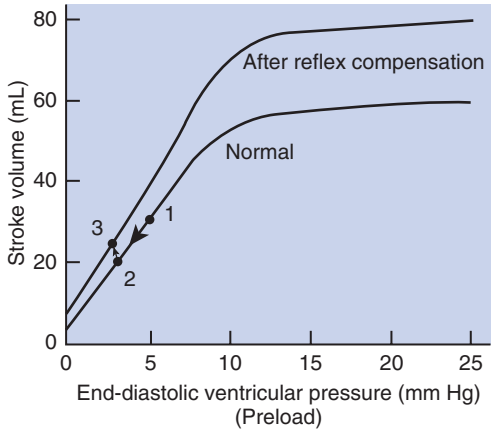
A third vicious, decompensatory cycle results from the fact that septic shock depresses cardiac contractility. The cycle is: severe ventricular failure → intense and prolonged splanchnic vasoconstriction → ischemic damage to intestinal mucosa → bacteria and endotoxins pass through the damaged mucosa, from intestines into bloodstream → bacteremia causes further depression of ventricular contractility.

Other decompensatory cycles develop in cases of severe, prolonged heart failure, but these three examples (which are illustrated in Fig. 26.3) show why decompensation is such a serious and frequently fatal development.

Careful clinical diagnosis and prompt treatment of heart failure are imperative, even if compensatory mechanisms have maintained blood pressure near its normal level when the patient is at rest. In evaluating the severity of heart failure and the extent of compensation, it is clinically useful to group the signs of heart failure into two categories. The first category is referred to as *backward heart failure*. The signs of backward heart failure include the changes in the circulation *upstream* from the failing ventricle: increased atrial pressure, increased venous pressure, excessive capillary filtration, edema, and the functional changes secondary to edema (e.g., respiratory failure). The category *forward heart failure* refers to the consequences of heart failure *downstream* from the failing ventricle: decreased cardiac output, decreased arterial blood pressure, exercise intolerance, and the consequences of prolonged vasoconstriction in the systemic organs (especially the kidneys and intestines).

The Immediate Cardiovascular Effects of Hemorrhage Are Minimized by Compensations Initiated by the Atrial Volume Receptor Reflex and the Arterial Baroreceptor Reflex

Figs. 26.4 and 26.5 summarize the cardiovascular responses to hemorrhage. The curve labeled *Normal* in Fig. 26.4 shows that the maintenance of a normal stroke volume depends on the maintenance of a normal level of ventricular preload. When



• **Fig. 26.4** Direct effect of hemorrhage is to decrease ventricular preload, which decreases stroke volume (transition from point 1, which is normal, to point 2). A reflex increase in sympathetic activity increases ventricular contractility above normal (upper curve), which restores stroke volume toward, but not to, normal (transition from point 2 to point 3).

hemorrhage occurs, blood is lost from the whole cardiovascular system but particularly from the veins, which are the blood reservoirs of the body. Hemorrhage therefore decreases venous volume, venous pressure, atrial pressure, ventricular preload, and ventricular end-diastolic ventricular volume. In the absence of any compensation, ventricular stroke volume decreases from normal (point 1 in Fig. 26.4) to point 2.

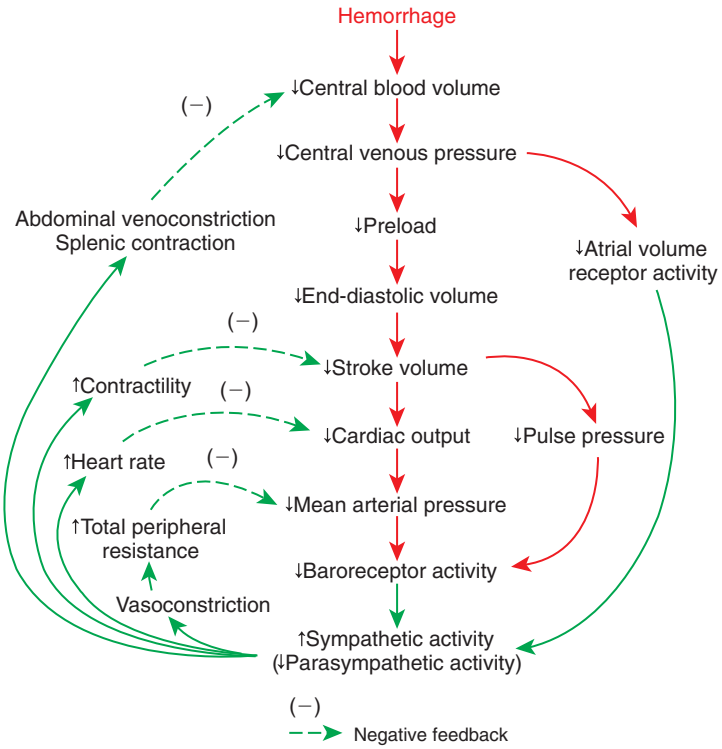
Note that no specification was made in the preceding paragraph about whether the itemized sequence of events affects the right heart or the left heart. The distinction is irrelevant because the right and left hearts are part of a series circuit. Therefore the volumes of blood pumped by the right and left ventricles must always come into balance within a few beats. Specifically, if hemorrhage lowers right ventricular preload (and therefore right ventricular output), the consequence will be decreased venous return to the left heart, which will decrease left ventricular preload (and therefore left ventricular output).

Fig. 26.4 shows that the normal ventricular function curve is rather steep to the left of point 1 (the normal operating point). Therefore a 40% hemorrhage results in approximately 40% reductions in venous pressure, atrial pressure, ventricular preload, and stroke volume. In the absence of compensations, cardiac output and MAP would also decrease by 40%. MAP would then be inadequate to sustain normal function in the critical organs, and the animal would die. With intact compensatory mechanisms, however, a normal animal can withstand a 40% hemorrhage without death and have only about a 10% decrease in MAP.

The arterial baroreflex and the atrial volume receptor reflex initiate the immediate compensations for hemorrhage. Hemorrhage decreases MAP, which decreases the activity of arterial baroreceptors. The reflexive response is to increase sympathetic activity and to decrease parasympathetic activity. The increased sympathetic activity acts on the heart to increase cardiac contractility. This helps restore stroke volume back toward normal, despite a persistent, subnormal preload and end-diastolic volume. The effect of this sympathetic compensation is diagrammed in Fig. 26.4 as point 3. Although stroke volume is returned toward normal, it remains low; after compensation for a 40% hemorrhage, the stroke volume may remain 25% below normal.

Additional compensations help restore MAP closer to normal despite the persistence of low stroke volume. First, heart rate

Vet-ebooks.com



• **Fig. 26.5** Summary of the consequences of hemorrhage (red arrows) and the rapid compensations initiated by the arterial baroreflex and atrial volume receptor reflex (green arrows). The changes described here include those portrayed graphically in Fig. 26.4. See text for further details.

increases above normal, which brings cardiac output back to within about 20% of its normal level. In addition, sympathetic vasoconstriction in the noncritical organs raises TPR above normal, resulting in a MAP that remains within approximately 10% of its normal level, despite a persistent 20% drop in cardiac output. You are urged to review the compensations described thus far by locating them on Fig. 26.5.

You may wonder why the compensatory actions of the baroreflex continue even though MAP is returned most of the way toward normal. The reason is that the baroreceptors are responsive to changes in pulse pressure as well as to changes in MAP, and pulse pressure remains low. Two factors cause the subnormal pulse pressure: (1) the persistent decrease in stroke volume and (2) the increase in heart rate above normal. Thus even if MAP is returned substantially toward normal after compensation for a hemorrhage, baroreceptor activity (action potential frequency) remains below normal.

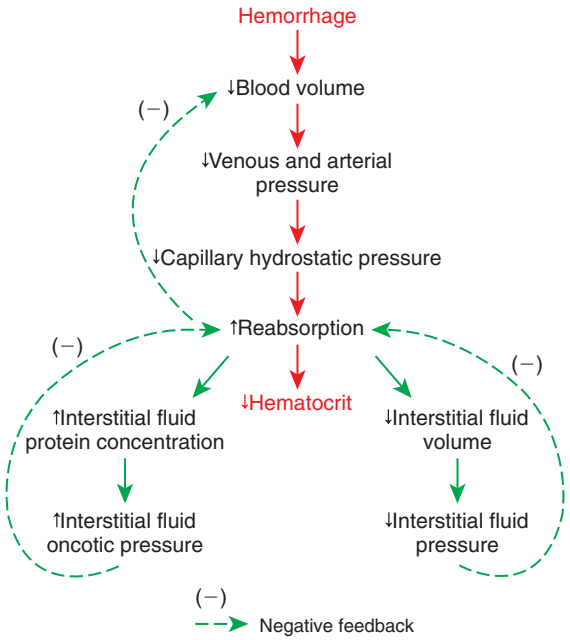
The atrial volume receptor reflex also contributes to the sustained increase in sympathetic activity (and decrease in parasympathetic activity) after hemorrhage. Hemorrhage leads to a persistent decrease in central venous pressure and atrial pressure. Therefore the activity of the atrial volume receptors is decreased below normal. The CNS responds to this decreased afferent activity from atrial volume receptors by elevating sympathetic efferent activity and by decreasing cardiac parasympathetic efferent activity. Thus as illustrated in Fig. 26.5, the atrial volume receptor reflex and the arterial baroreflex work synergistically to compensate for hemorrhage.

In severe hemorrhage. the reflex increases in sympathetic activity affect not only the heart and resistance vessels but also the veins. The abdominal veins in particular are constricted when sympathetic activation is intense. Sympathetic venoconstriction displaces blood from the abdominal veins and moves it toward the central circulation, which helps restore the low central venous pressure, atrial pressure, and preload back toward normal (see Fig. 26.5, *left side*). Sympathetic activation also constricts the blood vessels within the spleen and the muscular capsule around the spleen. Some of the blood sequestered in the spleen is expelled into the abdominal veins, and from there it moves toward the heart. In species that have large spleens (e.g., dog and horse), splenic contraction can mobilize a volume of blood equal to 10% of the total blood volume. An additional, adaptive feature of the blood sequestered in the spleen is that it has a higher-than-normal hematocrit. The mobilization of these sequestered red blood cells helps offset the fall in hematocrit that is a normal consequence of interstitial fluid reabsorption after hemorrhage (as described next).

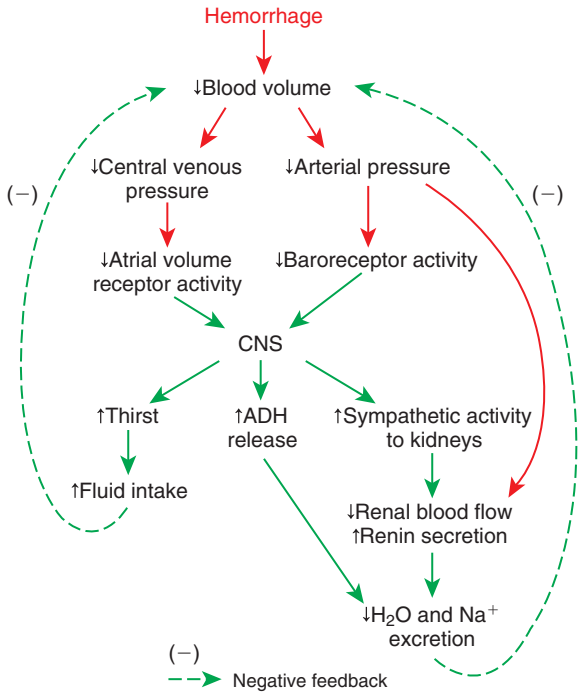
The arterial baroreceptor reflex and the atrial volume receptor reflex act within a few seconds to restore blood pressure toward its normal level after a hemorrhage. Other compensations come into play in the minutes and hours after hemorrhage to restore the lost fluid volume.

The Blood Volume Lost in Hemorrhage Is Restored Through a Combination of Capillary Fluid Shifts and Hormonal and Behavioral Responses

Hemorrhage causes both venous and arterial pressures to fall below normal, so capillary hydrostatic pressure also falls below normal. This alters the balance of hydrostatic and oncotic pressures acting on water in a direction that favors reabsorption of interstitial fluid back into the capillaries (Fig. 26.6). The volume of interstitial fluid that can be reabsorbed by this process in 1 hour is approximately 10% of the volume lost in the hemorrhage. However, the rate of reabsorption of interstitial fluid becomes limited after 3 to 4 hours. As interstitial fluid is reabsorbed, there is a decrease in interstitial fluid hydrostatic pressure (it becomes even more negative



• **Fig. 26.6** During the first 3 to 4 hours after a hemorrhage, interstitial fluid is reabsorbed into the bloodstream, which helps compensate for the lost blood volume. A complication is that the hematocrit decreases. Reabsorption is limited by decreases in interstitial fluid hydrostatic pressure and by increases in interstitial fluid oncotic pressure.



• **Fig. 26.7** Behavioral and hormonal responses after hemorrhage include increased fluid intake and decreased salt and water loss in the urine, which lead to the eventual restoration of the blood volume lost in the hemorrhage. *ADH,* Antidiuretic hormone; *CNS,* central nervous system.

than normal), and this opposes further reabsorption. Also, as interstitial fluid is reabsorbed, the interstitial fluid protein concentration increases because proteins in the interstitial fluid are not reabsorbed. The resulting increase in interstitial fluid oncotic pressure also opposes further reabsorption. Despite these limits, the reabsorption of interstitial fluid is an important compensation for hemorrhage in the first few hours.

The interstitial fluid that is reabsorbed into the bloodstream after a hemorrhage contains neither plasma proteins nor blood cells. Therefore the proteins and cells remaining in the bloodstream after the hemorrhage become diluted as interstitial fluid is reabsorbed. As a consequence, the concentration of plasma proteins in blood decreases and so does the hematocrit. This is why a decreasing hematocrit over a few hours in an otherwise normal patient is presumptive evidence that a hemorrhage has occurred recently or is continuing to occur. In the absence of an obvious hemorrhage, such a patient should be examined for evidence of internal bleeding.

Fig. 26.7 illustrates how the atrial volume receptor reflex and the arterial baroreceptor reflex participate in the eventual, complete restoration of blood volume after a hemorrhage. As mentioned earlier, hemorrhage leads to a decrease in the activity of both atrial volume receptors and arterial baroreceptors. One reflexive response to the decreased receptor activity is activation of sympathetic nerves. Some of the effects of the sympathetic activation have already been described (see Fig. 26.5). Sympathetic activity (coupled with a decrease in arterial pressure) also acts on the kidneys to increase their release of the hormone renin. As described in Chapter 25, renin works through the renin-angiotensin-aldosterone system to decrease sodium excretion by the kidneys. Moreover, decreased activity of the baroreceptors and atrial volume receptors also triggers an increase in ADH secretion from the pituitary gland. ADH circulates to the kidneys, where it reduces urine formation. The overall effect of renal vasoconstriction, the renin-angiotensin-aldosterone system, and ADH is to decrease the excretion of salt and water excretion in the urine. Note that these actions *conserve* the available body fluid volume after hemorrhage, but they do not *restore* it to normal. The actual restoration of body fluid volume after hemorrhage requires increased fluid intake. The baroreceptor reflex and the atrial volume receptor reflex, together with the hormone angiotensin, act through the hypothalamus to increase the sensation of thirst. If water is available, fluid intake increases until the lost body fluid volume is restored to normal. This may take 1 to 2 days.

The final compensations for hemorrhage involve the restoration of the lost plasma proteins and blood cells. The liver synthesizes the plasma proteins, and the blood cells are produced by the bone marrow. The time required for complete restoration may be several days for the plasma proteins and a few weeks for the blood cells.

The preceding discussion focused on the effects of severe hemorrhage. All the same compensations occur to a lesser degree after mild hemorrhage. For example, when a human donates blood, about 10% of the blood volume (0.5 L) is removed. All the compensations just described are evident after this 10% hemorrhage.

In Large Animals, the Transition From a Recumbent to a Standing Posture Elicits the Same Cardiovascular Responses as Hemorrhage

Gravity acts on all parts of the body, including the blood contained within blood vessels. Gravity increases the distending pressure in the dependent vessels (those below heart level). In an upright animal or human, the gravitational effect is most significant in

the leg and abdominal vessels. The increase in distending pressure does not cause much accumulation of blood in the arteries and arterioles because these vessels are not easily distensible (i.e., they have low compliance). However, the gravitational effect significantly distends dependent veins because of their much greater compliance. The extra blood that pools in the dependent veins is blood that would otherwise have returned to the central circulation. Therefore when an animal or human transitions from lying down to standing up, there is a decrease in both central blood volume and central venous pressure, just as there would be after hemorrhage. In a normal human, the assumption of an upright posture decreases central venous pressure by the same amount as does a 10% hemorrhage, and it triggers all the compensatory responses already described for hemorrhage. In small animals, the gravitational effect of standing is negligible. In some large animals, such as horses and cattle, the volume of blood that pools in the leg veins is minimized by the relatively small size of veins in the extremities.

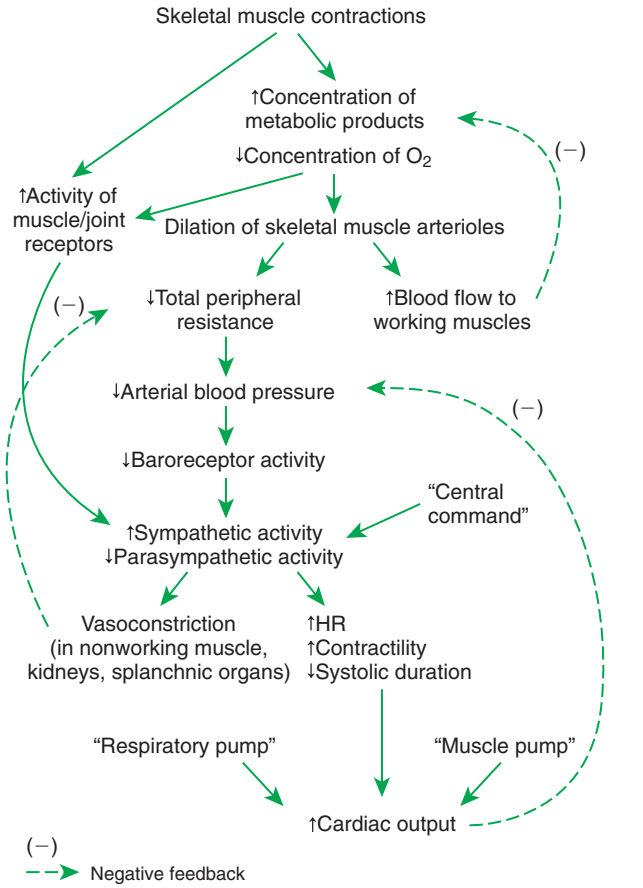
During the Initiation of Exercise, Local and Neural Mechanisms Interact to Increase Cardiac Output and to Deliver More Blood Flow to Exercising Muscle

As discussed in Chapter 24 and summarized in Fig. 26.8 (*top*), local metabolic control mechanisms dilate skeletal muscle arterioles during exercise. As exercise commences and muscle metabolism increases, metabolic products accumulate in exercising muscle, and the local oxygen concentration decreases. The metabolic products and hypoxia both cause dilation of the arterioles within the exercising muscle. This vasodilation is a local response, not dependent on nerves or hormones. The result is an increased blood flow to the exercising muscle (*active hyperemia*). The increased blood flow delivers more oxygen and removes some of the accumulated metabolic vasodilating products. A new steady-state balance is reached when the increase in blood flow matches the increase in muscle metabolism.

The metabolic (local) control mechanism can increase muscle blood flow enough to match the increased metabolic rate during exercise only if arterial blood pressure is maintained at a level sufficient to provide the needed additional blood flow. This necessitates a substantial increase in cardiac output and, in extreme exercise, vasoconstriction in the noncritical organs (which makes more blood flow available for the exercising muscle and other critical organs). These adjustments in cardiac output and blood flow through the noncritical organs are brought about by three neural mechanisms: central command, the exercise reflex, and the arterial baroreflex.

Central command is a psychogenic effect. In preparation for exercise (and continuing during exercise), the CNS increases sympathetic activity to the heart and blood vessels, and decreases parasympathetic activity to the heart. The sympathetic and parasympathetic changes are graded, depending on the intensity of the exercise. In effect, central command represents a "guess" by the brain as to the levels of sympathetic and parasympathetic activity that will be needed during the exercise to match cardiac output to the needs of the systemic organs.

The *exercise reflex* is the second mechanism that helps set the level of sympathetic and parasympathetic activity during exercise. The exercise reflex is initiated by specialized nerve endings within



• **Fig. 26.8** Cardiovascular responses to exercise involve a complex interplay of local (metabolic) control mechanisms with central command, reflexes, and the blood-pumping effects of muscle contraction and respiration. The overall result is increased blood flow in exercising muscle, decreased total peripheral resistance, vasoconstriction and decreased blood flow in the noncritical organs, substantially increased cardiac output, and maintenance of arterial blood pressure at or above its normal resting level. *HR,* Heart rate.

muscles and joints. An increase in muscular work and in the movement of the body joints activates these muscle and joint receptors. The resulting increase in afferent neural activity initiates a reflexive increase in sympathetic (and decrease in parasympathetic) efferent nerve activity. Although the mechanism for excitation of the muscle and joint receptors is not completely understood, it is clear that the activation of these receptors is necessary to keep blood pressure from falling during exercise.

The *arterial baroreceptor reflex* is the third major controller of sympathetic and parasympathetic activity during exercise. The baroreflex serves to fine-tune autonomic drive to the heart and arterioles to stabilize arterial pressure. If central command and the exercise reflex do not raise sympathetic activity to a sufficiently high level during a particular bout of exercise, arterial pressure will fall below normal. The arterial baroreceptors detect this low pressure, and the baroreflex responds by increasing sympathetic activity. Conversely, if central command and the exercise reflex raise sympathetic activity too high for the level of exercise, arterial pressure will rise above normal. The response of the baroreflex is to "dial back" on the increase in sympathetic activity. The importance of the baroreflex in stabilizing blood pressure during exercise is revealed by the large, poorly controlled swings in arterial pressure

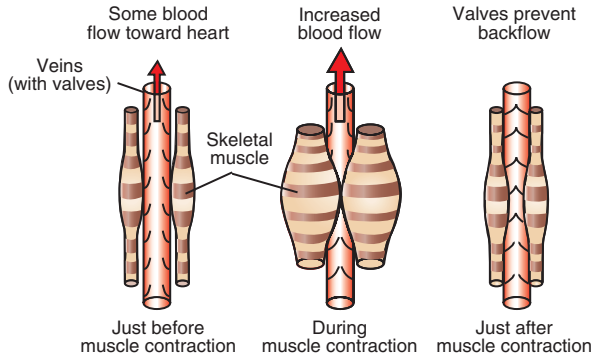
that occur during attempted exercise in animals whose arterial baroreceptors have been experimentally disabled.

In effect, central command and the exercise reflex initiate the autonomic adjustments for exercise, and the arterial baroreflex fine-tunes this autonomic activity to keep arterial pressure from falling (see Fig. 26.8).

An important additional feature of blood pressure control during exercise is that the arterial pressure the CNS accepts as "normal" is not fixed but varies with the type and intensity of exercise. In general, the more intense the exercise, the more the target or "set point" arterial pressure is increased above its normal resting level. With regard to the type of exercise, the set point for arterial pressure is significantly elevated during dynamic, rhythmic muscle contractions (e.g., a trotter or pacer during harness racing) but much more dramatically increased during exercise that involves powerful, sustained muscle contractions (e.g., a draft horse straining to pull an extremely heavy load). During exercise, the CNS resets the baroreflex so that it acts to regulate and stabilize blood pressure at the elevated target or set point level. This is analogous to resetting the cruise control of a car to regulate speed at an elevated level when leaving a country road and entering a freeway. Thus (as already mentioned in Chapter 25) it is more accurate to say that the baroreceptor reflex regulates blood pressure at a variable set point (set by the CNS) than to say that the baroreflex regulates arterial pressure at any single "normal" pressure.

Two additional, nonneural mechanisms also help increase the cardiac output during exercise. The first of these is the *muscle pump* (Fig. 26.9). When skeletal muscles contract, they squeeze the blood vessels contained within them. In consequence, during a sustained contraction, a muscle will tend to restrict its own blood flow (see Chapter 24). If the muscle contractions are rhythmical, however, each contraction causes blood to be expelled out of the muscle veins and thus toward the central circulation. During the subsequent muscle relaxation, there is minimal backflow of blood (from the central circulation backwards into the veins) because the veins have one-way valves within them. Thus by alternately squeezing and releasing the veins, exercising muscles exert a pumping action that displaces venous blood toward the central circulation and increases central venous pressure. The consequence is an increase in ventricular preload above the level that would otherwise exist. The increase in preload enhances ventricular filling, which (by Starlings law of the heart) increases stroke volume and therefore cardiac output.

The second nonneural mechanism that helps increase cardiac output during exercise is the *respiratory pump*. Vigorous exercise involves an increase in the rate and the depth of respiration. During each inspiration, a subatmospheric pressure is generated within the intrapleural space. This negative pressure distends the airways of the lungs and expands them. It also increases the distending pressure on the central (intrathoracic) veins and the heart. Distention



• **Fig. 26.9** During dynamic exercise, venous blood is "massaged" back toward the central circulation by the rhythmical, alternating contraction and relaxation of skeletal muscles. This so-called *muscle pump* helps increase central venous pressure in an exercising animal.

of the central veins and heart helps promote the flow of blood from the abdominal veins into the central veins and heart. In addition, the diaphragm muscle moves caudally during inspiration and compresses the abdominal organs. The resulting increase in intraabdominal pressure "squeezes" blood out of the abdominal veins and toward the central veins. Overall, the respiratory pumping action helps to increase venous return, central venous volume, and ventricular preload during exercise.

Cardiac output in well-conditioned humans and many animal species can increase to six times its resting level during vigorous exercise as a result of the combined effects of sympathetic and parasympathetic responses, the muscle pump, and the respiratory pump. Note, however, that the success of the mechanisms that increase cardiac output during exercise depends on the heart's ability to respond normally both to increased sympathetic drive and to increases in preload. As mentioned earlier, during heart failure the autonomic mechanisms available to increase cardiac contractility and heart rate are invoked simply to maintain a normal cardiac output at rest. Therefore the autonomic nervous system in a patient with heart failure has a limited ability to bring about further increases in cardiac output during the initiation of exercise. For this reason, patients with heart failure typically exhibit exercise intolerance.

Maximal exercise ability in normal humans and animals appears to be limited by cardiac output. That is, the respiratory system can oxygenate as much blood as the heart can deliver to the lungs, and skeletal muscles can take up and metabolize as much oxygen as the heart can deliver to them. Therefore oxygen transport from the lungs to the skeletal muscles reaches a maximum when cardiac output is maximized. This sets the upper limit to the level of exercise that can be sustained.

CLINICAL CORRELATIONS

Exercise Intolerance Secondary to Congestive Heart Failure

History

An 8-year-old female Great Dane has been diagnosed previously with idiopathic dilative cardiomyopathy. Severe, generalized cardiac enlargement was evident on thoracic radiographs. The dog has been losing weight and is unable to complete daily walks with her owners.

Clinical Examination

Femoral pulses are weak but regular at 140 beats/min. The mucous membranes are pale, and the capillary refill time is prolonged. Heart sounds are muffled, and a murmur is heard on the left side over the atrioventricular valve. Respiratory rate is higher than normal (45 breaths/min). Auscultation reveals increased bronchovesicular (respiratory) sounds. The abdomen is

Continued

CLINICAL CORRELATIONS—cont'd

distended, and the abdominal organs are difficult to palpate. The electrocardiogram shows sinus tachycardia with broad, high-voltage QRS complexes. Thoracic radiography reveals a greatly enlarged heart and moderate pulmonary edema. Echocardiography reveals dilation of all four cardiac chambers. Ejection fraction is below normal, and there is mitral regurgitation.

Additional diagnostic tests are conducted to help assess the degree of complications secondary to the heart failure. In arterial blood, the percentage saturation of hemoglobin with oxygen is 78% (normal, 95% to 100%), the difference in oxygen content between arterial and venous blood is 8.5 mL of O_2 per deciliter of blood (normal, 4 to 6 mL per deciliter), the serum creatinine concentration is 3 mg/dL (normal, <1 mg/dL), urine specific gravity is 1.036 (high normal), and central venous pressure is 14 mm Hg (normal, 0 to 3 mm Hg).

When persuaded to exercise, the dog appears to become tired after walking less than 1 minute. Her legs begin to tremble, and then she collapses. Her pulse rate increases to 180 beats/min, and her mucous membranes become dark and cyanotic (blue).

Comment

Chronic heart failure secondary to cardiomyopathy is common in large dogs older than 4 years of age. Often the cardiomyopathy is idiopathic (of unknown cause). The case presented here is fairly typical of advanced heart failure. All the clinical findings are either direct consequences of the heart failure or consequences of attempts by the body to compensate for the heart failure (see Figs. 26.1, 26.2, and 26.3). In brief, ventricular failure (decreased contractility) leads to decreased stroke volume, cardiac output, and blood pressure.

Compensations for the heart failure involve reflex decreases in parasympathetic activity, increases in sympathetic activity, and increases in the release of ADH and renin. Heart rate is increased, which helps raise cardiac output back toward normal. Pulse pressure, judged by palpation of the femoral pulse, is reduced (because stroke volume is low and heart rate is high). The pale color and slow refill time of the mucous membranes indicate vasoconstriction and decreased blood flow, and this strongly suggests that the dog's splanchnic organs, kidneys, and resting skeletal muscles are also vasoconstricted. The reduced blood flow to these "noncritical organs" helps support arterial pressure and reserves the available cardiac output for the heart and brain. Renal vasoconstriction slows the rate of urine formation. Urinary loss of salt and water is further reduced by the actions of ADH and the renin-angiotensin-aldosterone system. The urine that does form has a high solute concentration (high specific gravity). Metabolic products (e.g., creatinine) that are normally eliminated by the kidneys accumulate in the blood. The resulting uremia, if severe, can further depress cardiac function and initiate the vicious cycle of decompensation. Splanchnic vasoconstriction and low splanchnic blood flow may be impairing digestive function and contributing to the dog's weight loss.

Salt and water retention increases blood volume above normal. Most of the excess blood volume is in the veins, so venous and atrial pressures are above normal. The elevated atrial pressure (preload) increases ventricular end-diastolic volume above normal, which helps the failing heart to restore the low stroke volume back toward (not to) normal. However, the excessive volume and pressure of blood in the veins also cause systemic edema (including ascites, as evidenced by the distended abdomen) and also pulmonary edema (visible on the radiograph). Pulmonary edema impairs the ability of the lungs to oxygenate blood. Therefore the hemoglobin saturation and the oxygen content of arterial blood are both below normal in this dog. The tissues of the body respond to the low rate of oxygen delivery by unloading as much oxygen as they can from the blood as it flows through the tissue. This makes the arteriovenous difference in oxygen content greater than normal. The general inadequacy of cardiovascular transport and consequent metabolic stresses on all the tissues of the dog also contribute to weight loss.

Despite many compensatory mechanisms, this dog is unable to deliver a normal amount of well-oxygenated blood to the body tissues, even at rest. When the dog tries to exercise, cardiac output increases very little. Therefore when exercise-induced vasodilation occurs in the exercising muscles and TPR decreases, blood pressure falls dramatically. There is a further decrease in blood flow in the tissues of the systemic circulation that were already vasoconstricted (e.g., mucous membranes), and these tissues become hypoxic and cyanotic. Inadequate blood flow in the exercising skeletal muscles leads to hypoxia and acidosis, and the dog collapses.

Treatment

The ideal treatment strategy for this dog is to improve the contractile performance of the myocardium. Theoretically, β-adrenergic agonists or cardiac glycosides could be administered to increase cardiac contractility. However, currently available drugs are either ineffective or only mildly effective in dogs with severe, chronic heart failure. One reason is that dogs in heart failure have already engaged adrenergic drive to the heart through activation of their sympathetic nervous systems. Therefore instead of trying to increase cardiac contractility as the primary goal of treatment, a common strategy is to limit the adverse effects of venous and pulmonary congestion and to support cardiac output in various ways. Diuretic or venodilator drugs limit the increases in venous pressures and are usually effective in controlling signs of congestion (venous distention and edema). Such drugs must be used cautiously, however, because they create the risk of lowering preload and therefore exacerbating the low cardiac output. An appropriate initial treatment for this dog includes a diuretic (furosemide) and a cardiac glycoside (digitalis). If digitalis fails to improve cardiac contractility and cardiac output (likely in this advanced stage of cardiomyopathy), an arteriolar vasodilator (hydralazine) or a mixed vasodilator-venodilator (enalapril) can be added to the furosemide/digitalis regimen. Arteriolar vasodilators can augment the output of a failing heart by reducing the afterload (arterial pressure) against which the heart must eject blood.

Despite therapy, the prognosis for a dog with such severe, chronic heart failure is poor.

Cow With "Hardware Disease"

History

A 4-year-old, pregnant Holstein cow is presented for lethargy, poor appetite, and edema. She is due to calve in 2 months. The producer noticed that over the last few weeks the cow has seemed progressively more lethargic and reluctant to move. He observed swelling below her jaw and in her brisket. She has lost 75 to 125 pounds.

Clinical Examination

The cow appears depressed. She is dehydrated. Her mucous membranes are dark (indicating poor perfusion), and capillary refill time is prolonged. She has marked brisket and submandibular edema. Her jugular veins are prominent. She grunts when she moves. Her temperature, pulse, and respiratory rates are all increased. Her heart sounds are muffled (as if heard through fluid), and she has a murmur ("washing machine" murmur). She has increased bronchovesicular (respiratory) sounds dorsally, but the sounds are muffled ventrally. Peripheral pulses are weak. Rumen contractions are decreased (one every 1 to 3 minutes). Feces are scant. Blood is submitted for a complete blood count and chemistry profile. Results indicate that the white blood cell count is low and serum creatinine concentration is increased. Fibrinogen, globulins, and total protein are all increased. Calcium and potassium levels are low.

An electrocardiogram reveals decreased amplitude of QRS complexes and ST segment elevation. Echocardiography reveals excessive fluid and gas in the pericardial space. Fibrin tags are also present. The right atrium and right ventricle appear to collapse during diastole, which is consistent with *cardiac tamponade* (excessive pericardial fluid pushing in on the heart). The left ventricle also contracts less forcefully and less completely than normal during systole (*decreased left ventricular free wall motion*).

With guidance from the echocardiogram, a sample of pericardial fluid is obtained. The fluid is reddish in color (rather than clear) and has a distinct, bad odor. Laboratory analysis reveals elevated protein concentration and an

CLINICAL CORRELATIONS—cont'd

elevated count of white blood cells (primarily neutrophils) in the pericardial fluid. Culture reveals that both aerobic and anaerobic bacteria are present.

Comment

This cow has traumatic reticuloperitonitis with pericarditis. *Traumatic reticuloperitonitis* (TRP), or "hardware disease," is common in cattle. Cattle are indiscriminate eaters, and they accidentally swallow sharp metal objects that get mixed into their feed. Metal objects settle in the reticulum of the rumen. Contractions of the reticulum may push sharp objects through the wall of the reticulum and into the peritoneum. Bacteria follow and cause peritonitis. Subsequently, the sharp object may penetrate the diaphragm, which is located just cranial to the reticulum, and may then move on to penetrate the pericardium. The consequence is pericarditis (inflammation of the pericardium). Sequelae include formation of scar tissue (seen as fibrin tags), pericardial bacterial infection, and accumulation of inflammatory fluid in the pericardium. The pericardial fluid presses on the cardiac chambers, restricting their filling during diastole, and this causes pump failure.

Evidence of congestive pump failure in this cow includes poor perfusion (weak pulses, dark mucous membranes, and prolonged capillary refill time), cardiac abnormalities (subnormal right atrial and ventricular filling, decreased left ventricular motion), elevated heart rate, distended jugular veins, edema, and lethargy.

Treatment

Prognosis is poor in this case because of the combination of pericardial infection and congestive pump failure. The producer could try to treat the infection, in hopes of delivering a live calf. Because the cow is already in pump failure with very limited cardiac output, however, it is likely that the calf is not receiving sufficient blood flow and oxygen. The calf could die in utero and could be aborted by the cow.

Practice Questions

1. What goes in the blank in the following sequence of cause and effect?

Myocardial failure → Decreased cardiac contractility → _____ → Decreased stroke volume.

a. Increased end-diastolic volume
b. Decreased end-diastolic volume
c. Increased end-systolic volume
d. Decreased heart rate
e. Increased mean aortic pressure

2. Which of the following is most likely to be the cause of dependent edema (swelling in regions of the body below the level of the heart) in a dog with chronic heart failure secondary to hypertrophic cardiomyopathy?

a. Increased lymph flow
b. Increased capillary hydrostatic pressure
c. Increased interstitial fluid hydrostatic pressure
d. Increased plasma colloid osmotic (oncotic) pressure
e. Increased interstitial fluid colloid osmotic (oncotic) pressure

3. During experimental trials on a new artificial aortic valve, a 20-kg dog is anesthetized and placed on cardiac bypass for 1 hour (i.e., a heart-lung machine is substituted for the dog's own heart and lungs). After successful installation of the artificial valve, the dog is taken off bypass, and the normal circulation is restored. One hour later, the dog's central venous pressure is 20 mm Hg, mean arterial pressure is 75 mm Hg, and heart rate is 140 beats/min. The cardiac output is not measured, but the surgeon suspects that it is too low, and therefore the dog's tissues are not being adequately supplied with blood. Which of the following procedures would be most likely to improve the delivery of blood to the dog's tissues?

a. Administration of a drug that increases cardiac contractility (e.g., digitalis)
b. Administration of an α-adrenergic agonist drug (e.g., phentolamine)
c. Administration of a β-adrenergic antagonist drug (e.g., propranolol)
d. Increasing the heart rate by electrical pacing
e. Transfusion with 500 mL of whole blood

4. What goes in the blank in the following sequence of cause and effect?

Hemorrhage → Decreased central venous pressure → Decreased ventricular preload → _____ → Decreased stroke volume.

a. Decreased ventricular contractility
b. Decreased systolic duration
c. Decreased arterial baroreceptor activity
d. Decreased diastolic filling time
e. Decreased end-diastolic volume

5. Which of the following is most likely to be below normal in a dog that is beginning to compensate for a significant hemorrhage?

a. Activity of the arterial baroreceptors
b. Release of ADH from the pituitary gland
c. Sensation of thirst
d. Sympathetic nerve activity to the systemic arterioles
e. Sympathetic nerve activity to the heart

6. A dog suffers a sudden, severe hemorrhage. One hour later, the dog is resting quietly, but his arterial pulse pressure, mean pressure, and hematocrit are all below normal. His mucous membranes are pale, and capillary refill time is longer than normal. Which of the following is most likely?

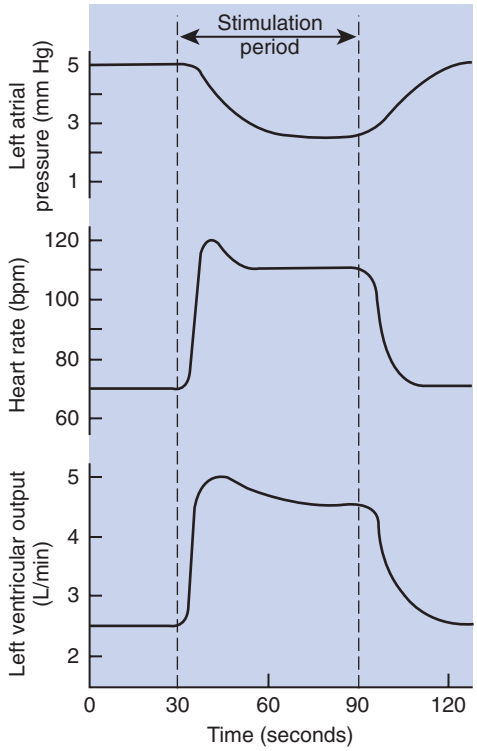
a. The diminished pulse pressure results from an abnormally low heart rate.
b. The diminished mean pressure results from an abnormally low total peripheral resistance.
c. The diminished hematocrit results from reabsorption of interstitial fluid into the bloodstream.
d. The pale mucous membranes result from abnormally low plasma protein concentration
e. The prolonged capillary refill time results from decreased sympathetic nerve activity to the blood vessels.

7. When a sheep is held in a vertical, head-up position, cardiac output decreases because:
   a. Sympathetic nerve activity causes the spleen to contract and abdominal veins to constrict.
   b. Valves in the leg veins prevent the return of blood to the heart.
   c. The respiratory pump promotes movement of abdominal venous blood into the thorax.
   d. Central blood volume increases.
   e. Right atrial pressure decreases.

8. What goes in the blank in the following sequence of cause and effect?
   Initiation of skeletal muscle exercise → Increased metabolic rate in exercising muscles → _____ → Increased blood flow in exercising muscles.
   a. Decreased cardiac output
   b. Decreased arterial pressure
   c. Vasodilation in exercising muscle
   d. Increased total peripheral resistance
   e. Increased delivery of oxygen to exercising muscle

9. During the transition from rest to exercise in a normal dog, which of the following will decrease?
   a. The external work (pumping work) of the heart
   b. The heart muscle's total work (total energy consumption)
   c. Blood flow through the lungs
   d. Systolic duration
   e. Velocity with which action potentials are conducted through the AV node

10. One of the nerves leading to a dog's heart is stimulated while left atrial pressure, heart rate, and left ventricular output are measured. The dog's responses are graphed in Fig. 26.10. Which of the following is a correct interpretation of these data?
    a. During the stimulation, venous return to the left atrium exceeds left ventricular output.
    b. The increase in left ventricular output at the beginning of stimulation can be explained by Starling's law of the heart.
    c. Stroke volume is lower after 15 seconds of stimulation than before stimulation.
    d. The effects of the nerve stimulation are similar to those caused by sympathetic activation of the heart.
    e. The effects of the nerve stimulation are similar to those that would result from activation of muscarinic cholinergic receptors in the heart.



• **Fig. 26.10** Cardiovascular data for Practice Question 10.

11. When a horse is running, the deep breathing and rhythmic contraction/relaxation of skeletal muscles help to maintain an elevated cardiac output by increasing:
    a. Venous return and central venous pressure
    b. Venous distending pressure in the legs
    c. Cardiac contractility
    d. Total peripheral resistance
    e. Intrathoracic (intrapleural) pressure

12. The responses to myocardial failure, hemorrhage, and exercise have in common that _____ will be elevated above its normal, resting level.
    a. Total peripheral resistance
    b. Heart rate
    c. Cardiac contractility
    d. Stroke volume
    e. Central venous pressure

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Elsevier; 2017.

Constable PD, Hinchcliff KW, Done SH, Grunberg W. *Veterinary Medicine—A Textbook of the Diseases of Cattle, Horses, Sheep, Pigs and Goats*. 11th ed. Elsevier; 2017.

Cote E. *Clinical Veterinary Advisor, Dogs and Cats*. 3rd ed. Elsevier; 2015.

Ettinger SJ, Feldman EC, Cote E. *Textbook of Veterinary Internal Medicine, Expert Consult*. 8th ed. Elsevier; 2017.

Hinchcliff KW, Geor RJ, Kaneps AJ, eds. *Equine Exercise Physiology: The Science of Exercise in the Athletic Horse*. Elsevier; 2008.

Katz AM, Konstam MA. *Heart Failure: Pathophysiology, Molecular Biology, and Clinical Management*. 2nd ed. Baltimore: Lippincott, Williams & Wilkins; 2008.

Kruger W. *Acute Heart Failure: Putting the Puzzle of Pathophysiology and Evidence Together in Daily Practice*. 2nd ed. Springer International Publishing; 2017.

Nelson RW, Couto CG. *Small Animal Internal Medicine*. 5th ed. Elsevier; 2014.
Noble A, Johnson R, Thomas A, Bass P. *The Cardiovascular System: Basic Science and Clinical Conditions*. 2nd ed. New York: Churchill Livingstone Elsevier; 2010.
Pappano AJ, Wier WG. *Cardiovascular Physiology*. 10th ed. Mosby; 2013.
Smith BP. *Large Animal Internal Medicine*. 5th ed. Elsevier; 2015.
Smith FWK, Tilley LP, Oyama MA, Sleeper MM. *Manual of Canine and Feline Cardiology*. 5th ed. Elsevier; 2016.
Xiang L, Hester RL, Granger DN. *Cardiovascular Responses to Exercise*. 2nd ed. Morgan & Claypool Life Sciences; 2017.
Zachary JF, ed. *Pathologic Basis of Veterinary Disease Expert Consult*. 6th ed. Elsevier; 2016.

27 Regulation of Gastrointestinal Function

THOMAS H. HERDT

KEY POINTS

1. An independent, intrinsic enteric nervous system lies within the wall of the gut.
2. The enteric nervous system contains receptors, sensory neurons, interneurons, and motor neurons.
3. The gut receives extrinsic innervation from the autonomic nervous system.
4. The gut contains afferent neurons that relay information to the central nervous system.
5. The gastrointestinal system has an intrinsic endocrine system.
6. The immune system participates in regulation of gastrointestinal activity.
7. The regulation of gastrointestinal function is integrated by the interaction of many regulatory molecules acting on multiple cell types in the gut.

The gastrointestinal (GI) system is regulated in an integrated manner from two control systems. One level of control is applied by the central nervous and endocrine systems, and is exerted in a manner similar to other organ systems. The second level of control is unique to the GI system and is exerted by intrinsic nervous and endocrine components located within GI organs. This intrinsic level of control allows the gut to regulate its functions autonomously based on local conditions, such as the amount and type of food contained in the lumen. Coordination of GI function with the rest of the body is achieved by integration of intrinsic (within the gut) and extrinsic (outside of the gut) influences (Fig. 27.1).

An Independent, Intrinsic Enteric Nervous System Lies Within the Wall of the Gut

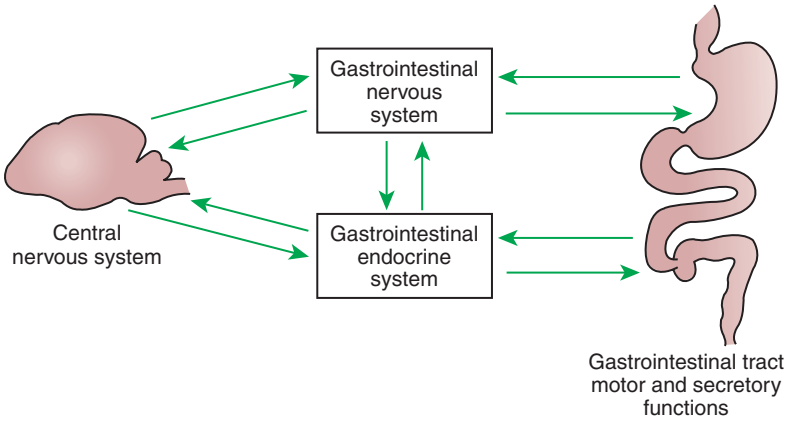
The enteric nervous system (ENS) is extensive and highly sophisticated, containing about as many neurons as the spinal cord. The ENS consists of cell bodies and their associated neurons, all of which lie within the gut wall. Anatomical characteristics of the GI wall are illustrated in Fig. 27.2, which depicts the small intestine as an example. Within the gut wall, cell bodies of the ENS are arranged into two systems of ganglia: the myenteric (Auerbach) plexus and the submucosal (Meissner) plexus. (See Chapter 13 for a more general discussion of ganglia.) The myenteric plexus consists of ganglia located between the circular and longitudinal muscle layers. The submucosal plexus has its ganglia in the submucosal layer. Axons from the cell bodies project in rich networks near the ganglia. A thick net of neurons runs in the plane between the circular and longitudinal muscle layers, connecting the ganglia of the myenteric plexus. Individual neurons leave the neuronal network to innervate structures within the gut wall and to intercommunicate between the myenteric and submucosal plexuses (Fig. 27.3). Interneuronal connections within the myenteric plexus are extensive and traverse long segments of gut, whereas interneuronal connections are limited within the submucosal plexus. The intricacy of the ENS has earned it the name "little brain" within the gut.

The Enteric Nervous System Contains Receptors, Sensory Neurons, Interneurons, and Motor Neurons

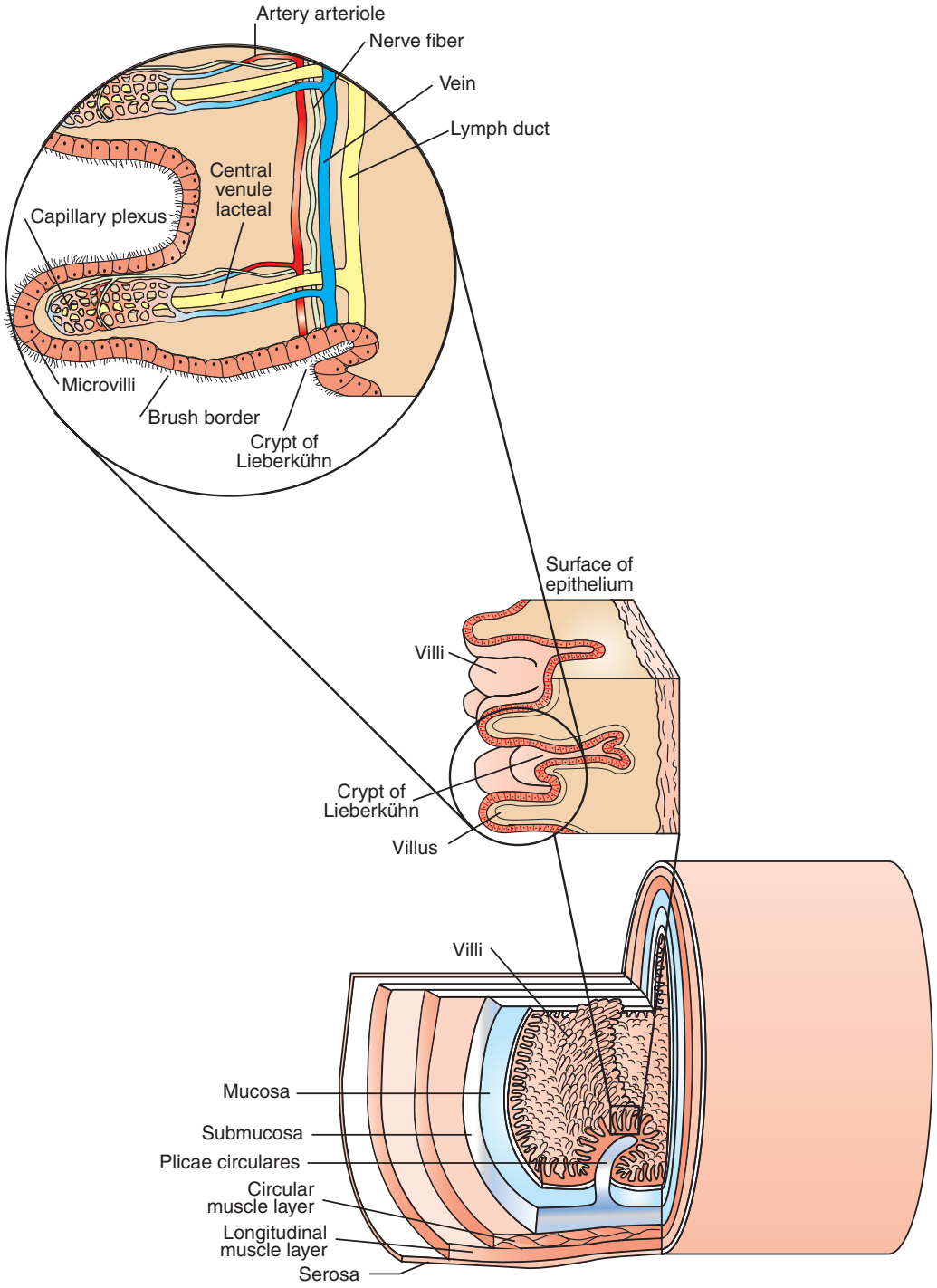
The plexuses of the ENS contain sensory (afferent) neurons, interneurons, and motor (efferent) neurons. Sensory input comes from mechanoreceptors within the muscular layers and chemoreceptors within the mucosa. Mechanoreceptors monitor distention of the gut wall, whereas chemoreceptors in the mucosa monitor chemical conditions in the gut lumen (Fig. 27.4).

Enteric motor nerves supply vascular muscle, gut muscle, and glands within the gut wall. Motor innervation of gut structures is less intimate than that found, for example, in skeletal muscle; no direct synaptic type of junction exists between enteric nerve endings and the structures they innervate. Rather, axons end in arborizations that contain many vesicular structures called *varicosities* (see Fig. 27.4). The varicosities contain regulatory substances collectively known as neurocrines that are secreted in response to action potentials and affect the activities of nearby muscle or glandular cells. Efferent neurons of the ENS may be stimulatory or inhibitory, and the nature of their action is largely determined by the type of neurocrine substance they secrete. Many of the neurocrine substances secreted by the ENS are peptides, and many are identical in chemical structure to regulatory molecules secreted by the endocrine and paracrine cells of the gut. These cells and their secretions are discussed in the following section.

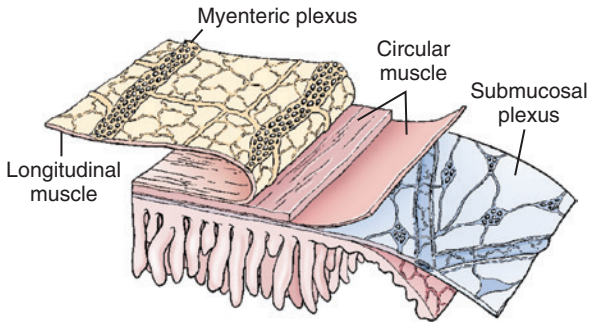
Many of the stimulatory, or excitatory, neurons are cholinergic, having acetylcholine as their neurocrine transmitter substance. A peptide known as substance P is another common excitatory neurocrine, and other minor excitatory neurocrines may exist. Inhibitory enteric neurons contain different neurocrine transmitters,



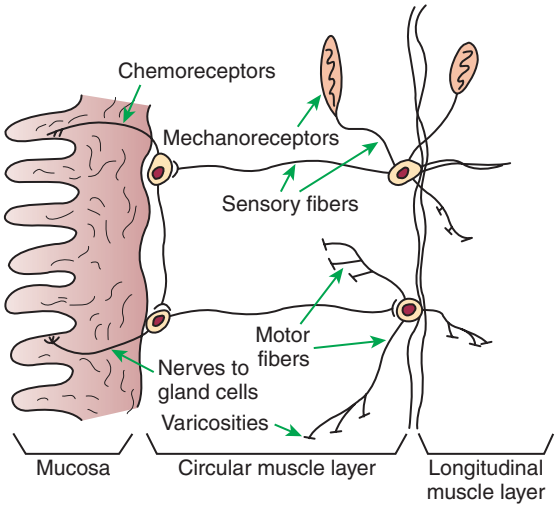
• **Fig. 27.1** Gut function is under direct regulation of the enteric nervous system (ENS) and the gastrointestinal (GI) endocrine system. Most central nervous system influence on the gut is mediated through indirect effects on the ENS and GI endocrine systems.



• **Fig. 27.2** Cross-sectional anatomy of the gut wall.



• **Fig. 27.3** Organization of the enteric nervous system. Note the arborizations of nerve fibers that run between the individual ganglia of the myenteric and the submucosal plexus. (Modified from Furness JB, Costa M. Types of nerves in the enteric nervous system. *Neuroscience* 1980; 5:1.)



• **Fig. 27.4** Arrangement of nerve fibers and receptors within the enteric nervous system. Varicosities release regulatory substances known as neurocrines in the vicinity of muscle fibers.

most of which are peptides. These inhibitory neurocrines vary widely and include such peptides as somatostatin and pituitary adenylate cyclase–activating peptide (PACAP), as well as nonpeptide neurocrines such as nitric oxide (NO) and adenosine triphosphate (ATP). Some neurocrines, such as vasoactive intestinal peptide (VIP), are inhibitory to gut muscle but stimulatory to the secretion of mucosal glands. The substances mentioned here are major neurocrines in the GI tract; the complete list of neurocrines and other regulatory substances in the gut is long and still expanding. Many of their names, such as calcitonin gene-related peptide (CGRP) and PACAP, are obscure, nondescriptive, and probably relate to historical events in their discovery rather than to their function in GI physiology. Box 27.1 illustrates the multiplicity and complexity of neurocrines and other regulatory molecules in the gut. The physiological basis for this complex and seemingly redundant system of regulatory molecules is discussed later after the GI endocrine and immune systems.

The Gut Receives Extrinsic Innervation From the Autonomic Nervous System

The parasympathetic and sympathetic nervous systems form the link between the central nervous system (CNS) and the ENS.

• BOX 27.1 Neurohumoral Regulatory Molecules of the Gut[a]

Peptide

Calcitonin gene-related peptide (CGRP)
Cholecystokinin
Enkephalins
Enteroglucagon
Gastrin-releasing peptide (GIP)
Gastrin
Galanin
Glucose-dependent insulinotropic peptide (GIP)
Motilin (I)
Neuropeptide Y
Neurotensin (I or S, depending on region of gut)
Peptide histidine isoleucine (PHI) (I)
Peptide YY
Pituitary adenylate cyclase–activating peptide (PACAP) (I)
Secretin
Somatostatin (I)
Substance K
Substance P (E)
Vasoactive intestinal polypeptide (VIP) (I)

Nonpeptide

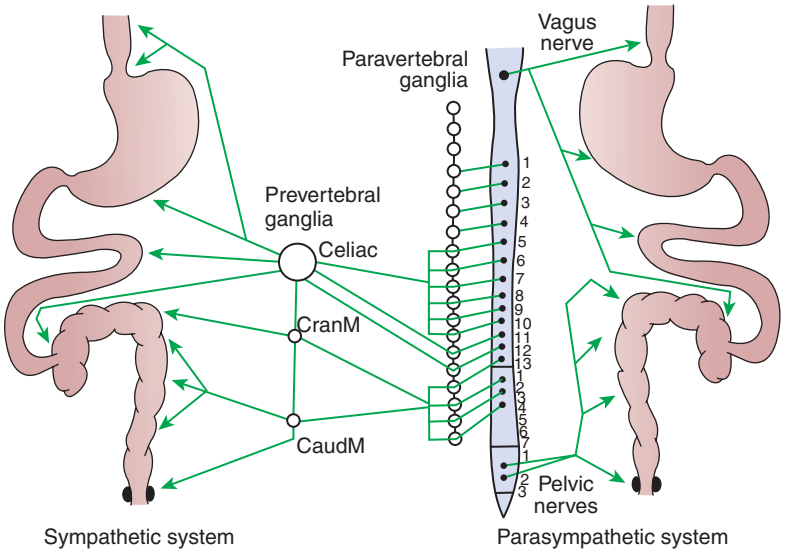
Acetylcholine (Ach) (E)
Nitric oxide (NO) (I)
Adenosine triphosphate (ATP) (E)
5-Hydroxytryptamine (5-HT, or serotonin) (I or E, depending on receptor type)

[a]*Most of these substances can act as neurocrines, endocrines (hormones), or paracrines, depending on their site of synthesis and mechanism of distribution. Those substances known to function as hormones are italicized. Their actions on smooth muscle activity are indicated as inhibitory (I) or excitatory (E).*

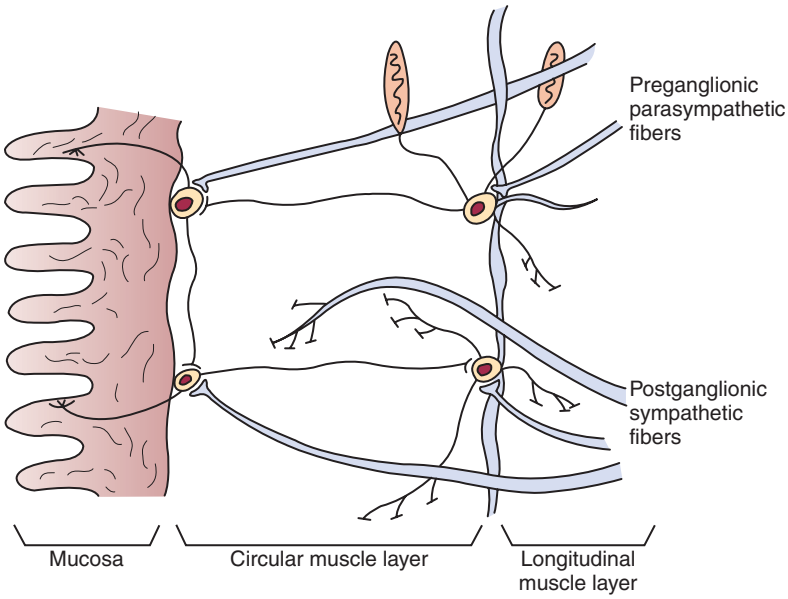
Most of the GI tract receives parasympathetic innervation by way of the vagus nerve, except the terminal portions of the colon, which receives parasympathetic innervation from the sacral cord through the pelvic nerve (Fig. 27.5). By classic description, the parasympathetic nervous system is composed of preganglionic and postganglionic fibers (see Chapter 13). However, this division of parasympathetic fiber types is not clearly defined for the gut, because extrinsic preganglionic fibers of the parasympathetic system become integrated with fibers of the ENS. Parasympathetic, preganglionic fibers reach the gut and synapse on cell bodies of the ENS, and thus the enteric ganglia of the gut serve as peripheral autonomic ganglia of the parasympathetic system. However, the ENS is much more than postganglionic parasympathetic neurons. Unlike typical postganglionic parasympathetic neurons, the ENS neurons receive input from sources other than preganglionic parasympathetic fibers. These include afferent neurons and interneurons of the ENS, as well as humoral (chemical) influences from other cells in the gut (see next section). Thus the classic description of the organization of the parasympathetic nervous system does not strictly apply to the gut (Fig. 27.6).

In contrast to the extrinsic parasympathetic fibers, extrinsic sympathetic fibers that enter the gut are primarily postganglionic. Postganglionic sympathetic fibers arise from cells in the prevertebral ganglia (see Chapter 13) and follow the splanchnic nerves and vascular arteries into the gut wall (see Fig. 27.5). Some sympathetic fibers synapse on neurons of the ENS, whereas others exert a direct effect on GI muscles and glands.

Neurons of the sympathetic and parasympathetic nervous system influence GI function through the release of neurocrines. In general,



• **Fig. 27.5** Distribution of autonomic nerve fibers to the gut. The spinal cord is represented in the center, with the sympathetic system extended to the left and the parasympathetic system to the right. *CranM,* Cranial mesenteric ganglion; *CaudM,* caudal mesenteric ganglion.



• **Fig. 27.6** Interface between the autonomic and enteric nervous systems. Note that parasympathetic fibers reaching neurons of the enteric system are preganglionic, whereas sympathetic fibers are postganglionic.

parasympathetic neurocrines are stimulatory (i.e., they increase gut blood flow, motility, and glandular secretions), whereas those of the sympathetic system are inhibitory.

The Gut Contains Afferent Neurons That Relay Information to the Central Nervous System

The gut contains afferent neurons that course to CNS centers through nerves associated with the autonomic nervous system. Vagal afferent nerves of the GI tract are associated with both mechanoreceptors and chemoreceptors and thus provide input to the CNS about changes in gut muscle tension and chemical conditions in the gut lumen. These signals allow the CNS to function in a coordinated manner with the ENS in the regulation of gut function. Chapters 28 to 32 discuss numerous examples of parasympathetic modulation and reinforcement of GI activities. These parasympathetic inputs depend on sensory signals received at the level of the brainstem. These signals and actions are part of the normal or physiological functions of the gut, and most of the signals do not reach the conscious awareness of the animal.

There are also afferent neurons from the gut that connect with the CNS through the splanchnic nerves, which also contain sympathetic efferent fibers. These afferent neurons convey signals arising from mechanoreceptors and chemoreceptors in the gut. In contrast to the vagal afferent nerves, however, the splanchnic afferents usually signal the presence of pathological conditions, such as overdistention of the gut wall, inflammation, or the presence of noxious chemicals or substances in the gut lumen. These nerves are distributed into the serosal surfaces and mesentery of the gut, in addition to the muscular and mucosal layers. These splanchnic afferent nerves convey to the CNS the conscious perception of pain caused by abnormal pressure within the gut lumen, inflammation, the presence of noxious stimuli in the gut lumen, or the stretching of mesenteric attachments. Such nerve signals result in the intense pain response characteristic of such conditions as equine colic and gastric distention and volvulus in dogs. Painful stimuli

from the gut usually evoke a sympathetic motor response, inhibiting gut motility and many glandular secretions.

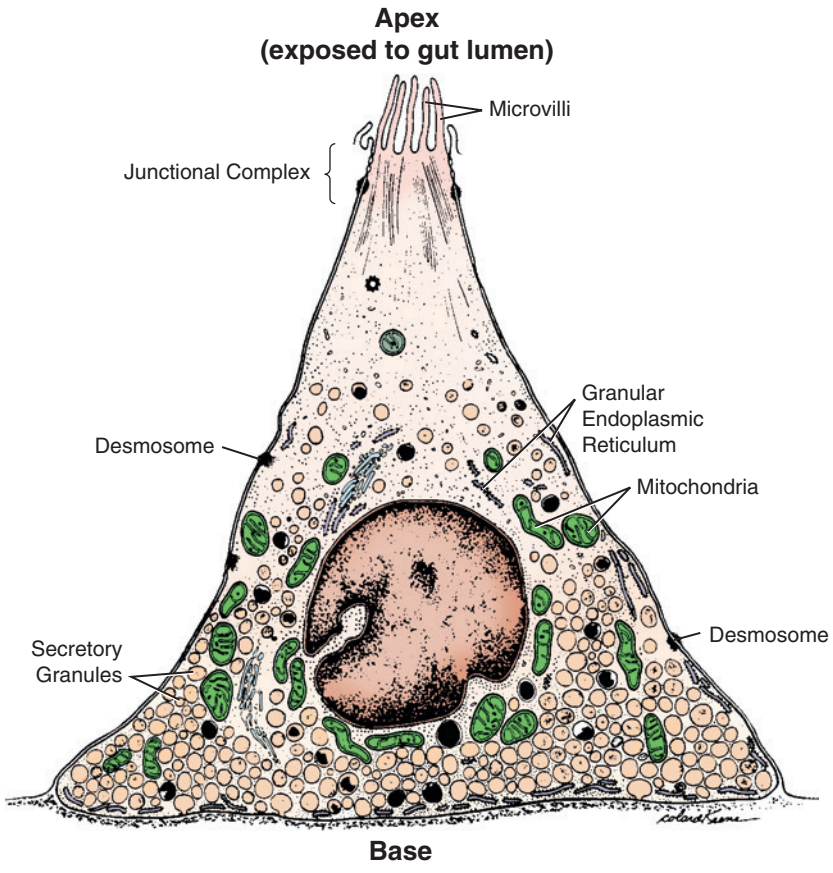
The Gastrointestinal System Has an Intrinsic Endocrine System

The GI system has an extensive number and variety of endocrine and paracrine cells, which are distributed diffusely throughout the gut epithelium. Endocrine cells are those that produce true hormones, which by definition are molecules that travel through the blood from their site of synthesis to their site of action. Paracrine substances, on the other hand, are molecules secreted by one cell that exert their effect locally, traveling by diffusion to nearby target cells. Typically, both the GI endocrine cells and the GI paracrine cells are columnar with a broad base and a narrow apex (Fig. 27.7). They are positioned individually among the other mucosal cells in both the secretory and the absorptive areas of the mucosa. The narrow apex of these cells is exposed to the lumen of the gut, allowing them to "sample" or "taste" the luminal contents. This anatomical arrangement provides a mechanism for the cells to sense changes in gut luminal contents and to respond by releasing hormones and other regulatory substances. The base of these cells contains secretory granules, which are storage forms of hormones and paracrine substances. Endocrine cells secrete near blood vessels, whereas paracrine cells secrete into the interstitial fluid. Some paracrine cells have long basal appendages that direct their secretions to the vicinity of target cells.

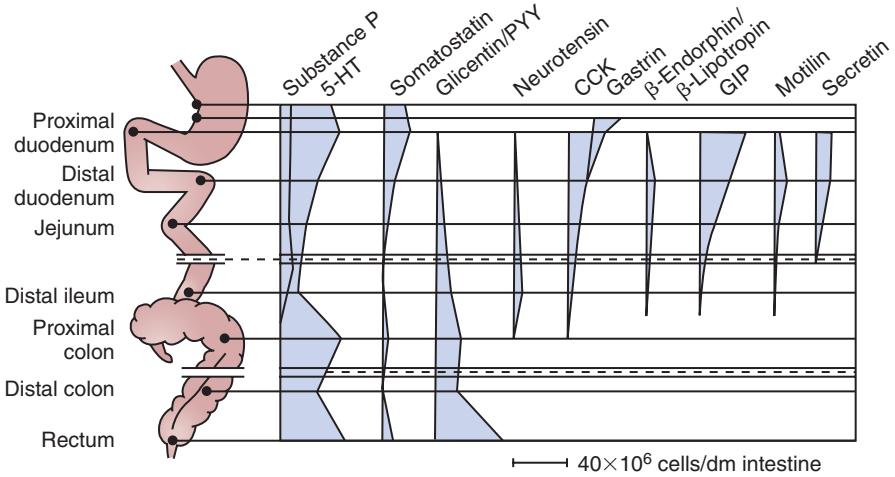
It is important for students to appreciate that secretory molecules of these cells are regulatory, not digestive. The products of these cells are secreted into either the blood or the interstitial fluid, not into the lumen of the gut. The many types of endocrine and paracrine cells in the gut epithelium are morphologically similar. However, despite their similar appearance, many distinct populations of cells produce a wide variety of hormones and other regulatory substances. Many of these substances are peptides in structure and can be referred to collectively as regulatory peptides, an inclusive term that may include molecules of neurocrine, endocrine, or paracrine origin (see previous definition of neurocrine). There are at least 28 peptides that have, or that are suspected to have, gut regulatory functions. Not all regulatory molecules in the gut are peptides, however, and several important nonpeptide paracrine and neurocrine substances exist. See Box 27.1 for a listing of peptide and nonpeptide neurocrine, paracrine, and endocrine molecules of importance in the regulation of gut function. For convenience and clarity, all these substances are frequently referred to collectively as *neurohumoral regulatory molecules.*

Each type of endocrine or paracrine cell has a characteristic distribution within the GI tract. For example, gastrin-producing cells, called G cells, are found primarily in the distal portion of the stomach, and few are found elsewhere in the gut. Cholecystokinin-producing cells are found in the small intestine, especially in the proximal region. Thus although endocrine cells are distributed throughout the GI tract, the production of individual regulatory molecules may be confined to specific areas. This is not always the case, however, as with somatostatin- and serotonin-producing cells, which are distributed along the entire length of the gut (Fig. 27.8). Table 27.1 lists the sites of production and actions of the major GI hormones.

A subset of GI paracrine cells is known as enterochromaffin cells because of their particular histological staining characteristics. These cells are found extensively throughout the gut, although



• **Fig. 27.7** Gastrointestinal (GI) endocrine cell. All the GI endocrine cells have a similar structure, but each cell produces only one type of hormone. Note the narrow apex that is exposed to the intestinal luminal contents and the broad base for storage of secretory granules. (From Johnson LR, Christensen J, Jacobsen ED, et al, eds. *Physiology of the Gastrointestinal Tract,* vol 1. 2nd ed. New York: Raven Press; 1987.)



• **Fig. 27.8** Distribution of gastrointestinal endocrine and paracrine cells throughout the gut, based on the hormone or paracrine substance they secrete. (Modified from Brown DR, ed. Gastrointestinal regulatory peptides. In *Handbook of Experimental Pharmacology,* vol 106, Berlin: Springer-Verlag; 1993.)

TABLE 27.1 Major Gastrointestinal Hormones

	Site of		
Hormone	Production	Action	Release Stimulus
Gastrin	Distal stomach	Primary: Stimulates acid secretion from stomach glands Secondary: Stimulates gastric motility, growth of stomach epithelium	Protein in stomach; high gastric pH; vagal stimulation
Secretin	Duodenum	Primary: Stimulates bicarbonate secretion from pancreas Secondary: Stimulates biliary bicarbonate secretion	Acid in duodenum
Cholecystokinin (CCK)	Duodenum to ileum, with highest concentration in duodenum	Primary: Stimulates enzyme secretion from pancreas Secondary: Inhibits gastric emptying	Proteins and fats in small intestine
Gastric inhibitory polypeptide (GIP)	Duodenum and upper jejunum	Primary: Inhibits gastric motility and secretory activity Secondary: Stimulates insulin secretion provided sufficient glucose is present; may be most important action in many species	Carbohydrate and fat in small intestine
Motilin	Duodenum and jejunum	Primary: Probably regulates motility pattern of the gut in period between meals Secondary: May regulate tone of lower esophageal sphincter	Acetylcholine

their density varies from region to region (see Fig. 27.8). These mucosal cells are similar in structure to the endocrine and paracrine cells previously discussed, being exposed to the luminal contents of the gut through their apical membranes. These cells secrete a regulatory molecule known as serotonin or 5-hydroxytryptamine (5-HT). This regulatory substance is particularly important in signaling neurons that have excitatory influence on gut muscle, as well as in generating sensory signals from the gut mucosa. The enterochromaffin cells thus serve as important chemoreceptors within the gut with serotonin serving as the signaling molecule linking these chemosensory cells to afferent neurons of the ENS.

The GI regulatory peptides influence various gut functions and in many cases form part of regulatory feedback loops, many of which take the form of negative feedback loops. The thermostat and furnace in your house form an example of a negative feedback loop. A drop in temperature is detected by the thermostat, stimulating the thermostat, so to speak, to send a regulatory signal to the furnace. In response, the furnace ignites, raising the temperature of the room. The increase in temperature removes the "stimulus" to the thermostat, subsequently removing the signal to the furnace to ignite, which then stops the increase in room temperature. Thus room temperature is modulated within a controlled range. Numerous

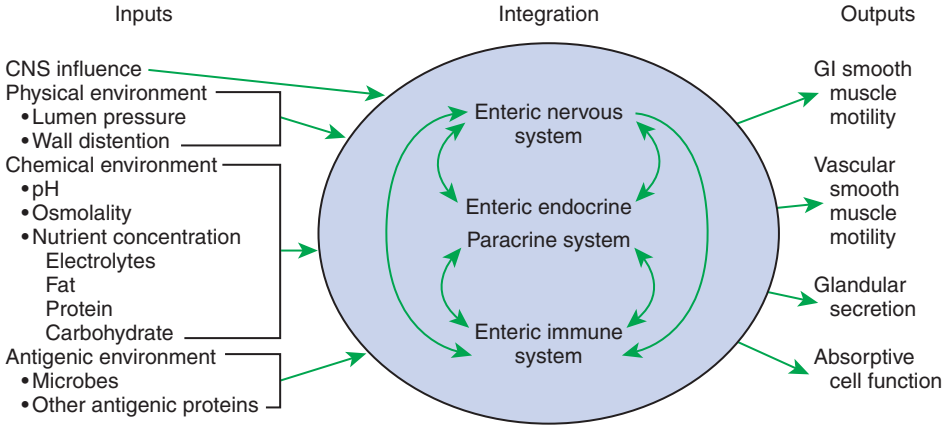
negative feedback loops exist in the gut whereby a response elicited by a regulatory molecule has a suppressive effect on the secretion of that molecule. For example, the hormone gastrin stimulates gastric acid production, lowering gastric pH. Low gastric pH in turn suppresses the secretion of gastrin, thus modulating stomach pH in the same manner that room temperature was modulated by the thermostat and furnace. Numerous examples of GI neurohumoral regulatory feedback loops will be discussed in subsequent chapters in the context of the actions they regulate.

The Immune System Participates in Regulation of Gastrointestinal Activity

The ENS and GI endocrine/paracrine systems monitor the physical (wall distention and pressure) and chemical environment within the gut. Another important aspect of the gut environment, however, is the antigenic milieu. It is the function of the GI immune system to monitor the antigenic environment of the gut. The mucosal surface of the gut is exposed to a vast number of microorganisms and other antigens, and the many immune cells in the gut mucosa provide the means necessary to control their numbers and their access to the body. In fact, the majority of the immune system cells in the body reside in the gut mucosa. These cells span the entire range of immune cell types. On one hand, the immune cells of the gut respond to antigenic stimulation similar to immune cells in other body sites: antigenic memory is created; neutralizing and opsonizing antibodies are synthesized; and killer cells are recruited (see Chapters 54 and 55). On the other hand, inflammatory mediators such as prostaglandins, histamine, and cytokines, the products of immune cells, can interact directly with the ENS and GI endocrine/paracrine cells to modulate gut activities. For example, if a microbe begins to invade an area of the gut, sensitized immune cells will secrete prostaglandins, cytokines, and other immune mediator substances. These immune mediators can interact directly with cells of the ENS and the GI endocrine/paracrine system, eliciting a response of increased fluid secretion and motility in the stimulated area of gut. The result is that the pathogenic microbe is washed downstream in the gut and eventually passed in the feces, thus protecting the gut and ridding the animal of the pathogen.

The Regulation of Gastrointestinal Function Is Integrated by the Interaction of Many Regulatory Molecules on Multiple Cell Types in the Gut

You should now appreciate that the physical, chemical, and antigenic environments of the gut are monitored by multiple systems, and that responses to various stimuli are mediated by a bewildering and seemingly redundant array of neurohumoral regulatory molecules. From a clinical standpoint, it is most important to realize that all the regulatory molecules of the gut form a highly integrated scheme for the overall control of gut function (Fig. 27.9). This integration is achieved in a milieu of stimulatory and inhibitory molecules that can arise from nervous or glandular (endocrine/paracrine) elements of the gut. All these regulatory molecules act by occupying ligand sites on target cells. These target cells may be glands, muscles, and vascular muscle cells of the GI tract, but they may also be other regulatory cells, such as neurons and endocrine/paracrine cells. Thus nerve cells can influence endocrine/paracrine cells, and vice versa. This provides a system of positive and negative feedback controls that can orchestrate a fine degree of regulation over gut function. The overall effect is



• **Fig. 27.9** Stimulus-response coupling of gut reflexes and reactions at their most basic level. Despite its apparent complexity, the gut receives a limited array of afferent information and can respond with a similarly limited number of reactions. Afferent information supplied to the gut includes signals from the central nervous system, mechanical stimulation generated by wall distention or luminal pressure, chemical signals generated by luminal contents, and antigenic stimulation provided by microbial products or ingested antigens. This information is received by three integrating systems: the enteric nervous system, the endocrine/paracrine system, and the immune system of the gut. Among these systems there is considerable cross-communication that aids in integrating information and refining the efferent signals to the response elements. The responses of the gut are equally limited. Smooth muscle of the gastrointestinal wall can contract or relax; vascular smooth muscle can adjust blood supply; and absorptive and secretory cells can become more or less active. Although the gut is indeed complex, all functions may be reduced to these elements. *CNS*, Central nervous system; *GI*, gastrointestinal.

that the balance of inhibitory versus excitatory neurohumoral regulatory molecules determines the activity of muscles, glands, and blood vessels in a given area of the gut.

From a clinical standpoint, it is important to understand that specific neurohumoral regulatory molecules provide opportunities for pharmacological intervention. Several highly effective drugs that either mimic or block the actions of GI neurohumoral regulatory molecules are now available for the treatment of GI diseases in animals.

Acknowledgment

We would like to thank Dr. Ayman Sayegh for authoring an earlier edition of this chapter.

Practice Questions

1. Which statement is *the most accurate* anatomical description of the enteric nervous system (ENS) of the gut?
- **a.** Intrinsic neuronal fibers and their cell bodies are diffusely spread throughout the length and thickness of the stomach and intestine.
- **b.** Intrinsic neuronal fibers traverse the length of the stomach and intestine in discrete nerve bundles.
- **c.** Intrinsic neuronal cell bodies are aggregated in a discrete "gut brain" that is positioned near the pylorus.
- **d.** Intrinsic neuronal cell bodies lie in discrete planes within the thickness of the gut wall and are diffusely distributed throughout its length.
- **e.** Intrinsic neuronal fibers exist only in the longitudinal muscle layer of the stomach and intestine.

2. Which statement is *true* with regard to the parasympathetic fibers that innervate the cells of the ENS?
- **a.** The fibers exit the central nervous system from lumbar segments of the spinal cord.
- **b.** The fibers have vasoactive intestinal peptide as a neurotransmitter.
- **c.** The fibers are inhibitory.
- **d.** The fibers are preganglionic.
- **e.** There are no parasympathetic fibers that innervate cells of the ENS.

3. Which statement is *true* with regard to the GI endocrine cells?
- **a.** Secretory activity is influenced by the luminal contents of the gut.
- **b.** Hormones are secreted directly into the gut lumen and affect the activity of glands "downstream" from the point of secretion.
- **c.** The secretory products of the GI endocrine cells are steroid hormones.
- **d.** Each GI endocrine cell can produce many different types of hormones, depending on the stimulus applied to it.

4. Which of the following statements concerning the neurohumoral regulatory molecules of the gut is *true*?
- **a.** They are all peptides.
- **b.** They all are excitatory and increase the rate of gut motility and secretion.
- **c.** GI neurohumoral regulatory molecules and their analogs and antagonists will probably become important agents of therapy in GI disease.
- **d.** All the above.

5. Which of the following conditions in the gut would NOT provide direct sensory input to the ENS?
- **a.** Changes in luminal pressure resulting in changes in tension of gut wall musculature.
- **b.** Changes in luminal pH.
- **c.** Changes in luminal osmotic pressure.
- **d.** Changes in the rate of ingesta flow.

6. Which of the following neurocrine transmitter molecules is most consistently excitatory relative to gut functions?
- **a.** Somatostatin
- **b.** Acetylcholine
- **c.** Vasoactive intestinal peptide (VIP)
- **d.** ATP

7. All neurocrine transmitter molecules are peptides.
- **a.** True
- **b.** False

8. The conscious sensation of pain due to excessive distention of a segment of the gut arises from afferent impulses traveling to the brain through:
- **a.** The myenteric plexus
- **b.** The submucosal plexus
- **c.** The vagus nerve
- **d.** The splanchnic nerves

9. Which anatomical description best applies to the organization of the GI endocrine cells?
- **a.** Cells arranged into small islets in the intestinal submucosa.
- **b.** Cells distributed diffusely throughout the mucosa of the GI tract.
- **c.** Cells arranged into discrete glands situated between the circular and longitudinal muscle layers.
- **d.** Cells distributed diffusely throughout the submucosa of the GI tract.

10. The influence of the GI mucosal immune system on gut functions is mediated by:
- **a.** The CNS.
- **b.** The direct action of cytokines on gut smooth muscle and glands.
- **c.** The interaction of cytokines with the ENS and GI endocrine/paracrine cells.
- **d.** The direct action of T cells on gut smooth muscle and glands.

Bibliography

Entry-Level Supplements

Barrett KE. Neurohumoral regulation of gastrointestinal function in gastrointestinal physiology. In: Barrett KE, eds. *Gastrointestinal Physiology*. 2nd ed. McGraw-Hill Education; 2014.

Johnson LR. *Gastrointestinal Physiology*. 8th ed. Philadelphia: Mosby; 2014.

Advanced Reading

Bellono NW, Bayrer JR, Leitch DB, et al. Enterochromaffin cells are gut chemosensors that couple to sensory neural pathways. *Cell*. 2017; 170(1):185.

Brierley SM, Hughes P, Harrington A, Blackshaw LA. Innervation of the gastrointestinal tract by spinal and vagal afferent nerves. In: Johnson LR, Ghishan FK, Kaunitz JD, et al, eds. *Physiology of the Gastrointestinal Tract*. 5th ed. Elsevier Science; 2012.

Chao C, Hellmich MR. Gastrointestinal peptides: gastrin, cholecystokinin, somatostatin, and ghrelin. In: Johnson LR, Ghishan FK, Kaunitz JD, et al, eds. *Physiology of the Gastrointestinal Tract*. 5th ed. Elsevier Science; 2012.

Furness JB. *The Enteric Nervous System*. 2nd ed. Malden, Massachusetts: Blackwell; 2006.

Genton L, Kudsk KA. Interactions between the enteric nervous system and the immune system: role of neuropeptides and nutrition. *Am J Surg*. 2003;186:253.

Kunze WA, Furness JB. The enteric nervous system and regulation of intestinal motility. *Annu Rev Physiol*. 1999;61:117.

Stevens CE, Hume ID. *Comparative Physiology of the Vertebrate Digestive system*. 2nd ed. Cambridge, UK: Cambridge University Press; 1996.

Wood JD. Integrative functions of the enteric nervous system. In: Johnson LR, Ghishan FK, Kaunitz JD, et al, eds. *Physiology of the Gastrointestinal Tract*. 5th ed. Elsevier Science; 2012.

28

Motility Patterns of the Gastrointestinal Tract

THOMAS H. HERDT

KEY POINTS

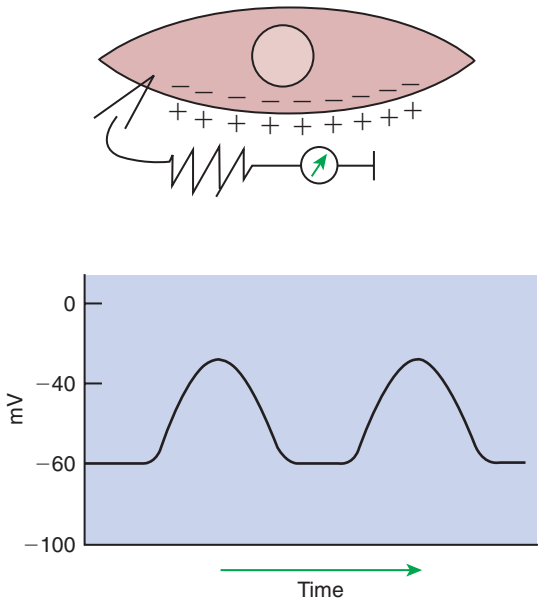
1. Slow waves of electrical depolarization are a unique feature of gut smooth muscle.
2. When slow waves reach sensitized smooth muscle cells, action potentials and contraction result.
3. Coordinated motility enables the lips, tongue, mouth, and pharynx to grasp food and propel it down the gastrointestinal tract.
4. Motility of the esophagus propels food from the pharynx to the stomach.
5. The function of the stomach is to process food into a fluid consistency and release it into the intestine at a controlled rate.
6. The proximal stomach stores food awaiting further gastric processing in the distal stomach.
7. The distal stomach grinds and sifts food entering the small intestine.
8. Control of gastric motility differs in the proximal and distal stomach.
9. The rate of gastric emptying must match the small intestine's rate of digestion and absorption.
10. Between meals, the stomach is cleared of indigestible material.
11. Vomiting is a complex reflex coordinated from the brainstem.
12. Motility of the small intestine has digestive and interdigestive phases.
13. The ileocecal sphincter prevents movement of colon contents back into the ileum.
14. Motility of the colon causes mixing, retropulsion, and propulsion of ingesta.
15. The colon is an important site of storage and absorption in all animals.
16. Despite large anatomical differences in the colons of herbivores compared with omnivores and carnivores, there are similarities in motility.
17. The anal sphincter has two layers with separate innervation.
18. The rectosphincteric reflex is important in defecation.
19. Major differences between avian and mammalian digestive systems include, in birds, both the lack of teeth and the separation of gastric functions into distinct anatomical regions.

The walls of the gastrointestinal (GI) tract, at all levels, are muscular and capable of movement. Movements of the GI muscles have direct actions on ingesta in the gut lumen. GI movements have several functions: (1) to propel ingesta from one location to the next; (2) to retain ingesta at a given site for digestion, absorption, or storage; (3) to break up food material physically and mix it with digestive secretions; and (4) to circulate ingesta so that all portions come into contact with absorptive surfaces.

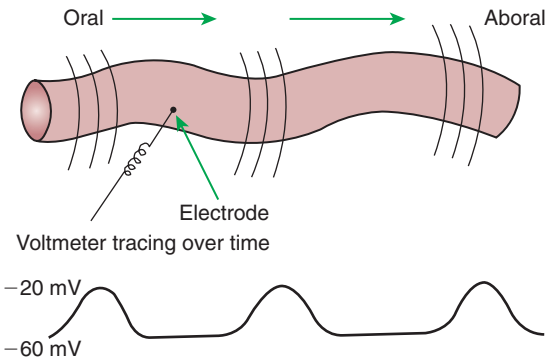
Movement of the gut wall is referred to as *motility,* and motility may be of a propulsive, retentive, or mixing nature. The time it takes material to travel from one portion of the gut to another is referred to as the *transit time.* An increase in propulsive motility decreases the transit time, whereas an increase in retentive motility increases the transit time. Selectively increasing retentive motility and reducing propulsive motility are important aspects of diarrhea therapy.

Slow Waves of Electrical Depolarization Are a Unique Feature of Gut Smooth Muscle

The first level of control of GI motility lies in the intrinsic electrical properties of the smooth muscle mass. These electrical properties consist of spontaneously undulating waves of partial depolarization that sweep over the gut smooth muscle. The origin of this electrical activity is from specialized smooth muscle cells referred to as the *interstitial cells of Cajal* (ICC). The ICC form an interconnecting lattice of cells that surrounds the circular and longitudinal layers of muscle over the entire length of the gut. These cells are very similar in structure and function to the Purkinje cells of the heart. The ICC exhibit rhythmical and spontaneous oscillation in their transmembrane electrical potentials, as illustrated in Fig. 28.1. They are connected to one another and to cells of the general smooth muscle mass by *tight junctions* or *nexuses.* These connections allow for the flow of ions from cell to cell. The resulting ionic



• **Fig. 28.1** Spontaneous changes in membrane polarity of the interstitial cells of Cajal, specialized gastrointestinal (GI) smooth muscle cells responsible for spontaneous electrical rhythmicity in gut muscle. The upper illustration represents a single cell with a voltmeter measuring the transmembrane electrical potential. The graph illustrates spontaneous changes in electrical potential (in millivolts, mV) that would be measured across the cell membrane.



• **Fig. 28.2** Partial membrane depolarizations of GI smooth muscle cells occur in a coordinated manner, creating waves of depolarization that sweep over large segments of muscle. Electrodes placed on or near the surface of the muscle record changes in potential as waves of depolarization pass toward or away from them. Coordinated changes in membrane potential among cells are necessary for these waves to be measured because random changes among cells would cancel each other out, and no changes would be recorded by electrodes placed extracellularly.

movements lead to the propagation of waves of partial cell membrane depolarization across large numbers of cells. Within the ICC, fluctuations in intracellular calcium concentrations appear responsible for the spontaneous changes in membrane polarization. Fig. 28.1 illustrates the concept of a fluctuating membrane potential in a single ICC. The property of spontaneous electrical rhythmicity, in combination with their electrical connection to the smooth muscle mass, imparts to the ICC their role as electrical "pacemakers" of the gut.

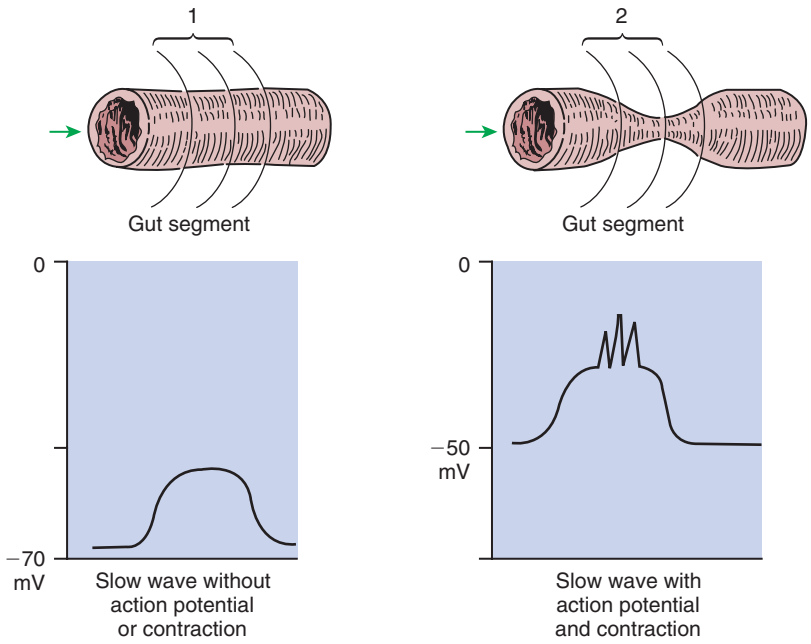
In GI smooth muscle cells, the baseline membrane potential is usually −70 to −60 millivolts (mV). Under the influence of the ICC, the membrane potential fluctuates from this baseline level by as much as 20 to 30 mV. Thus under resting conditions, the depolarization is only partial, and the membrane potential never reaches 0 mV. The smooth muscle cells are connected to the ICC and to each other by nexuses, allowing the changes in membrane potential to be spread, or *propagated,* over large areas of muscle. The ICC initiate these changes and thus determine their origin and direction of propagation. Under normal conditions in the small intestine, changes in membrane potential begin high in the duodenum and are propagated aborally (away from the mouth) along the length of the small intestine (Fig. 28.2). These aborally moving waves of partial depolarization are called *slow waves* or the *basic electrical rhythm* of the gut. In the dog, slow waves occur about 20 times per minute in the small intestine. In the stomach and colon, slow waves occur less frequently, about five times per minute. However, the slow waves are present throughout the smooth muscle portions of the GI tract. The frequency of slow waves varies among domestic species, but their presence does not vary.

The slow waves are an intrinsic property of the GI smooth muscle and associated ICC. The presence of the slow waves depends only on the ICC, whereas the amplitude and, to a lesser extent, the frequency of the slow waves can be modulated by the enteric nervous system (ENS). The link between slow waves and muscle contractions, however, is under control of nervous, endocrine, and paracrine factors, as discussed next.

When Slow Waves Reach Sensitized Smooth Muscle Cells, Action Potentials and Contraction Result

Slow waves have an important relationship with muscle contractions, but they are not the direct stimuli for contractions. Slow waves are constantly passing over GI smooth muscle, whether it is actively contracting or not. GI smooth muscle cells, as with other muscle cells, contract in association with action, or spike, potentials. These potentials are characterized by complete depolarization of the membrane for a short time, in contrast to the slow waves, which are characterized by incomplete depolarization (see Chapter 4). Action potentials in the GI smooth muscle occur only in association with slow waves. Thus the presence of slow waves is necessary but not sufficient to cause muscle contractions. When slow waves pass over an area of smooth muscle without eliciting action potentials, no contractions occur. When slow waves pass over an area of smooth muscle and action potentials are superimposed on the slow waves, gut muscle contracts. Control and coordination of smooth muscle activity is achieved by influencing the likelihood that action potentials will be superimposed on slow waves. Such control is a function of peptides and regulatory substances produced by the ENS and enteric endocrine and paracrine cells.

Smooth muscle control and coordination are achieved by modulation of the baseline electrical potential in the smooth muscle cells. Peptides and other regulatory molecules from ENS neurons or endocrine/paracrine cells are released in the vicinity of the smooth muscle cells, affecting membrane ion channels and influencing the baseline membrane potential (see Chapter 27 for a discussion of gut peptides and other regulatory molecules). Excitatory molecules elevate the baseline (bring it closer to zero), and inhibitory molecules lower the baseline (make it more negative). The position of the baseline influences how close the overall potential will come to



• **Fig. 28.3** *1,* No muscle contraction occurs in the absence of action potentials. *2,* Muscle contracts when the crest of the slow waves reaches a critical point of depolarization, allowing action potentials to occur. The probability of action potentials occurring during the passage of a slow wave over a segment of gut muscle is influenced by the degree of baseline depolarization. Norepinephrine lowers the baseline (increases its absolute value), whereas acetylcholine raises the baseline (decreases its absolute value). *mV,* Millivolts.

0 mV at the crest of a slow wave. When the membrane potential of a smooth muscle comes close to zero, action potentials occur and muscle contracts (Fig. 28.3). Regulatory molecules (neurocrines, paracrines, and hormones) that are excitatory elicit smooth muscle contraction by elevating the baseline, whereas inhibitory substances inhibit muscle contraction by lowering the baseline.

The integrated actions of the slow waves, ENS, and endocrine/paracrine system appear to function to synchronize the contractions of the GI muscle mass. For the muscle to perform efficiently, all or many of the muscle cells in one layer of a segment of gut must be synchronized to contract simultaneously. This can best be visualized by considering the circular muscle layer. The contents of the circle cannot be "squeezed" effectively unless all the muscles of the circumference contract simultaneously; it would have little effect on luminal pressure if one portion of the circle contracted while another portion relaxed. In any discrete area of gut, slow waves pass simultaneously over the entire circumference of the smooth muscle. If that area has been sensitized by an excitatory regulatory molecule, the entire circumference of circular muscle will contract in synchrony.

Muscle contractions can occur at a frequency no higher than the frequency of the slow waves. As an example of frequency modulation, consider the activity of muscle in the stomach of the dog. Slow waves in the canine stomach occur about five times per minute. The crest of each slow wave may or may not be accompanied by action potentials. Therefore during a given minute, the muscle in a localized area may not contract at all or may contract up to five times. If the passing slow waves generate no action potentials, the muscle does not contract at all. In a given minute, if action potentials are associated with one slow wave, the muscle contracts once. Action potentials on two slow waves result in two contractions, and so on, up to a maximum of five contractions per minute, but no more than five, because there are no more slow waves.

The motility patterns of the gut vary in their complexity, as described in the following sections. In the stomach and colon, motility patterns are relatively complex compared with the small intestine. In all cases, motility patterns are programmed into the ENS and coordinated in conjunction with the slow waves.

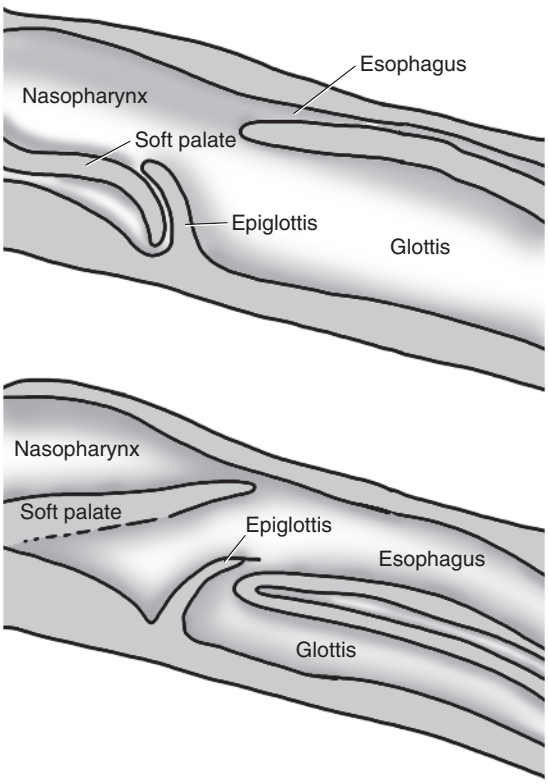
Coordinated Motility Enables the Lips, Tongue, Mouth, and Pharynx to Grasp Food and Propel It Down the Gastrointestinal Tract

Before digestion can begin, food must be directed into the GI tract. To ingest food, quadruped animals must first grasp it with the lips, teeth, or tongue. This involves highly coordinated activity of small, voluntary skeletal muscles. The muscles of the face, lips, and tongue appear to be among the most delicately controlled voluntary muscles of most domestic animals. The exact method of food *prehension* varies greatly among different species. For example, horses use their lips extensively, whereas cattle often use their tongues for grasping food. In all domestic animals, however, prehension is a highly coordinated process involving direct control by the central nervous system (CNS). Problems of prehension may develop because of abnormalities in the teeth, jaws, muscles of the tongue and face, cranial nerves, or CNS. The facial nerve, the glossopharyngeal nerve, and the motor branch of the trigeminal nerve control the muscles of prehension.

Mastication, or chewing, involves the actions of the jaws, tongue, and cheeks, and is the first act of digestion. It serves not only to break food particles down to a size that will pass into the esophagus but also to moisten and lubricate food by thoroughly mixing it with saliva. Abnormalities of the teeth are a common cause of digestive disturbances in animals.

Deglutition, or swallowing, involves voluntary and involuntary stages and occurs after food has been well masticated. In the voluntary phase of swallowing, food is molded into a bolus by the tongue and then pushed back into the pharynx. When food enters the pharynx, sensory nerve endings detect its presence and initiate the involuntary portion of the swallow reflex.

The involuntary actions of the swallow reflex occur primarily within the pharynx and esophagus. The *pharynx* is the common opening of both the respiratory and the digestive tract. The major physiological function of the pharynx is to ensure that air, and only air, enters the respiratory tract and that food and water, and



• **Fig. 28.4** Midline cross-sectional schematics showing the position of the structures of the larynx and pharynx during breathing (*top*) and swallowing (*bottom*).

only food and water, enter the digestive tract. The involuntary portion of the swallow reflex is the action that directs food into the digestive system and away from the upper airway. This reflex involves the following series of highly coordinated actions (Fig. 28.4). Breathing stops momentarily. The soft palate is elevated, closing the pharyngeal opening of the nasopharynx and preventing food from entering the internal openings of the nostrils. The tongue is pressed against the hard palate, closing off the oral opening of the pharynx. The hyoid bone and larynx are pulled forward; this action pulls the glottis under the epiglottis, blocking the laryngeal opening. Concurrently, the arytenoid cartilages constrict, further closing the opening of the larynx and preventing the movement of food into the respiratory system. When all openings to the pharynx are closed, a wave of muscular constriction passes over the walls of the pharynx, pushing the bolus of food toward the opening of the esophagus. As the food reaches the esophagus, the upper esophageal sphincter relaxes to accept the material.

The complex reactions of deglutition are controlled by lower motor neurons located in various centers of the brainstem. Efferent nerve fibers from these centers travel in the facial, vagus, hypoglossal, and glossopharyngeal nerves, and the motor branch of the trigeminal nerve. Clinically, problems with prehension, mastication, and deglutition frequently are related to neurological lesions, either peripherally in the cranial nerves or centrally in the brainstem.

Motility of the Esophagus Propels Food From the Pharynx to the Stomach

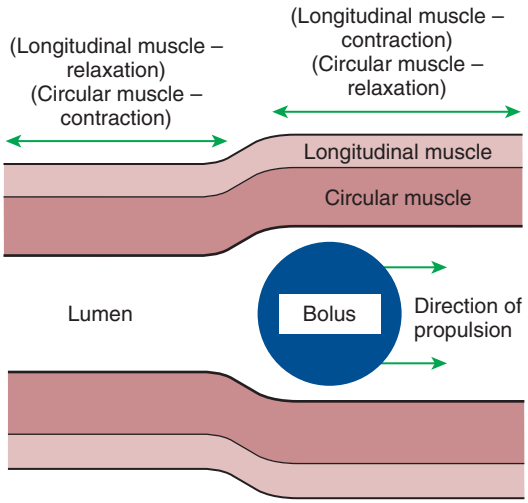
The *esophagus,* as with other tubular portions of the gut, contains an outer longitudinal and inner circular layer of muscle. The esophagus is unique compared with other areas of the gut in that much of its muscular wall is composed of striated skeletal muscle fibers. In most domestic animals, the entire length of esophageal musculature is striated. In horses, primates, and cats, however, a portion of the distal esophagus is smooth muscle. The striated muscle portions of esophagus are under control of somatic (not parasympathetic) motor neurons in the vagus nerve, whereas the smooth muscle portions are under direct control of the ENS and indirect control of the autonomic nervous system. A myenteric plexus exists throughout the entire length of the esophagus. In the area of striated muscle, the myenteric plexus probably serves a sensory function and acts to coordinate the movements of the striated muscle portion with the esophageal smooth muscle segments and stomach.

In terms of motor activity, the esophagus may be viewed as consisting of an upper sphincter, body, and lower sphincter. The upper esophageal sphincter is called the *cricopharyngeal muscle.* This muscle and the upper end of the esophagus are attached to the cricoid cartilage of the larynx. When deglutition is not taking place, the muscle compresses the end of the esophagus against the cartilage of the larynx, tightly closing the upper esophageal opening. During deglutition, the cricopharyngeal muscle relaxes and the larynx is pulled forward. The ventral portion of the upper end of the esophagus is attached to the larynx and the dorsal portion to the cervical spine. Because of these attachments, the forward motion of the larynx in conjunction with the relatively fixed nature of the cervical spine tends to pull open the upper esophageal orifice passively (see Fig. 28.4).

The body of the esophagus serves as a relatively simple conduit, rapidly transferring food from the pharynx to the stomach. Food is propelled through the esophagus by propulsive movements known as *peristalsis.* Peristalsis consists of a moving ring of constriction in the wall of a tubular organ. In the esophagus, these rings start at the cranial end and progress toward the stomach. The rings reduce or obliterate the esophageal lumen, thus pushing the bolus of food ahead of them in much the same manner as a person would push material out of a soft rubber tube by stripping it with the fingers. In addition to the constriction of the circular muscles, there may be some contraction of longitudinal muscles just ahead of, or aboral to, the ring of circular muscle contraction. This longitudinal muscle activity increases the size of the esophageal lumen to accommodate the advancing food bolus (Fig. 28.5). Peristalsis is a universal type of GI propulsive motility that exists at all levels of the gut.

During deglutition, the upper esophageal sphincter relaxes as the pharynx constricts; food is pushed into the upper portion of the esophageal body, and a wave of peristalsis propels the material toward the stomach. As the food bolus reaches the distal end of the esophagus, the lower sphincter relaxes, and the ingested matter enters the stomach. If the esophagus is not cleared of food material by the primary wave of peristalsis, secondary peristaltic waves are generated. One or more secondary waves are almost always adequate for pushing material into the stomach and clearing the esophagus. If food or foreign bodies become lodged in the esophagus, secondary waves of peristalsis may lead eventually to muscle spasms that constrict tightly around the lodged material. These spasms frequently interfere with therapeutic attempts to remove obstructing objects in the esophagus.

When deglutition is not taking place, the body of the esophagus is relaxed, but the upper and lower sphincters remain constantly constricted. The constriction of these sphincters is important because of the differences in external pressure applied to the esophagus at



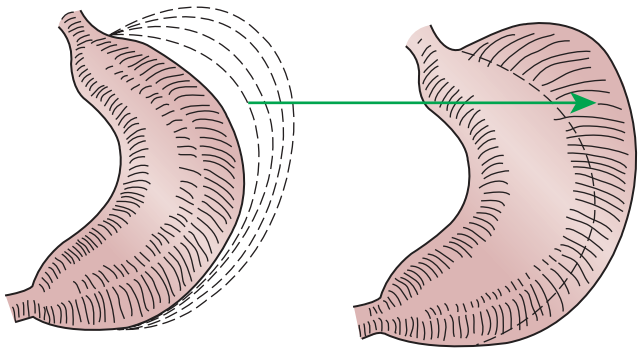
• **Fig. 28.5** Peristalsis consists of a moving ring of luminal constriction preceded by an area of luminal distention. The area of constriction is created by contractions of the circular muscle, whereas the dilation is created by contractions of the longitudinal muscle. The net action is to propel a bolus of ingesta.

different points along its length. During the inspiratory phase of breathing, the portion of the esophagus within the thorax is subjected to less-than-atmospheric pressure. If the two esophageal sphincters were not tightly closed, inspiration would cause aspiration of air from the pharynx and reflux of ingesta from the stomach into the body of the esophagus. Stomach contents would be drawn into the esophagus because inspiratory pressures in the thorax are lower than intraabdominal pressure. It is particularly important that the lower esophageal sphincter remains closed during inspiration because the mucosa of the esophagus is not equipped to resist the caustic actions of gastric contents, thus movement of stomach contents into the esophagus would cause damage to the esophageal mucosa.

In many species, the action of the lower esophageal sphincter is aided by the anatomical nature of the attachment of the esophagus and stomach. The esophagus enters the stomach obliquely, allowing distention of the stomach to block the esophageal opening in a valvelike manner. During deglutition, the longitudinal muscle of the esophagus contracts, shortening the esophagus and opening the valve at the junction with the stomach. This anatomical arrangement, along with the lower esophageal sphincter, is particularly well developed in the horse, making reflux of stomach material into the esophagus extremely rare in this species. In many cases, when the intragastric pressure of the horse is pathologically raised, the stomach ruptures before vomiting or esophageal reflux takes place.

The Function of the Stomach Is to Process Food Into a Fluid Consistency and Release It Into the Intestine at a Controlled Rate

Among animal species, there is tremendous diversity in the anatomy and motility patterns of the stomach. The following discussion applies best to the animals with the simplest stomachs, such as the dog and cat, but is probably also a reasonable description of the activity of the somewhat more complex stomachs of the pig, horse, and rat. The complex motility patterns of the ruminant stomach are discussed in Chapter 31.



• **Fig. 28.6** Adaptive relaxation refers to the stretching of the stomach wall that occurs as the organ fills during eating. This stretching results from muscle relaxation and is accompanied by little or no change in intraluminal pressure.

The function of the stomach is to serve food to the small intestine. There are two important aspects of this function: rate of delivery and consistency of material. The stomach serves both as a storage vat to control the rate of delivery of food to the small intestine and as a grinder and sieve that reduces the size of food particles and releases them only when they are broken down to a consistency compatible with small-intestinal digestion.

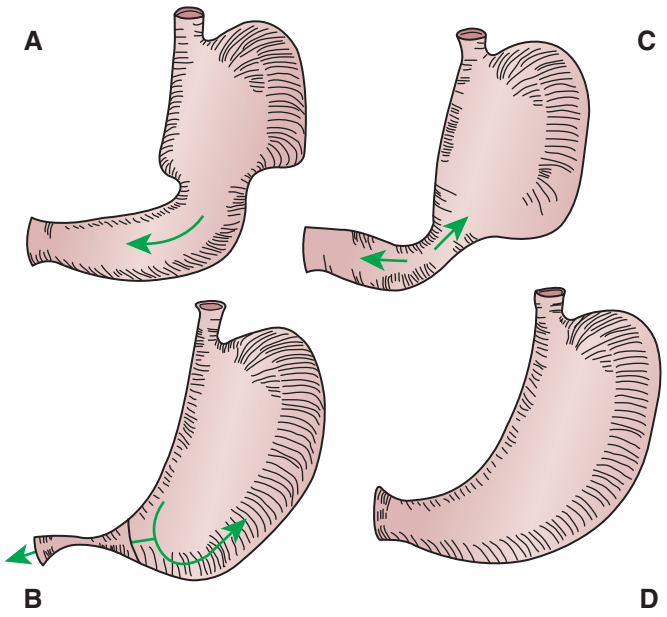
The stomach is divided into two physiological regions, each of which has a different effect on gastric function. The *proximal region,* at the esophageal end of the stomach, serves a storage function, retaining food as it awaits eventual entry into the small intestine. The *distal region* serves a grinding and sieving function, breaking solid pieces of food down into particles small enough for small-intestinal digestion.

The Proximal Stomach Stores Food Awaiting Further Gastric Processing in the Distal Stomach

The major muscular activity in the proximal region of the stomach is of a weak, continuous-contraction nature. These *tonic* contractions tend to shape the gastric wall to its contents and provide gentle propulsion of material into the distal stomach. The major muscular reflex of the proximal stomach is *adaptive relaxation* (Fig. 28.6). This reflex is characterized by relaxation of the muscles as food enters the stomach. Because of this relaxation, the stomach can dilate to accept large quantities of food without an increase in intraluminal pressure. Thus the proximal stomach serves as a food storage area. Because of the rather passive muscular activity of the proximal stomach, little mixing occurs there. In fact, food boluses tend to become layered in the stomach in the order in which they were swallowed. As the stomach empties, tension on the wall of the proximal stomach increases slightly, pushing food distally in the stomach, where it can be processed for transport into the duodenum.

The Distal Stomach Grinds and Sifts Food Entering the Small Intestine

The muscular activity of the distal stomach and *pylorus* (sphincter-like junction between stomach and duodenum) is completely



• **Fig. 28.7** Grinding and churning activity of the distal stomach. (A) Wave of peristalsis begins at the junction of the proximal and distal areas of the stomach and moves toward the pylorus. (B) As the peristaltic wave approaches the pylorus, the pylorus constricts, causing some of the ingesta to be crushed within the peristaltic ring and propelled back toward the proximal stomach. (C) As the peristaltic wave reaches the pylorus, some finely ground and liquefied material passes through into the duodenum, but the majority of material has been propelled back into the stomach. (D) Between contractions, no gross movement of gastric contents occurs. (From Johnson LR, editor. *Gastrointestinal Physiology,* St. Louis, MO: Mosby; 1985).

different from that of the proximal stomach. In the distal stomach, known as the *antrum,* there is intense slow-wave activity, and muscular contractions are frequently present. Strong waves of peristalsis begin at about the middle of the stomach and migrate, with the slow waves, toward the pylorus. As the waves of peristalsis near the pylorus, the pylorus constricts, blocking the gastric exit of all but the smallest particles (Fig. 28.7). Particles leaving the stomach during the digestive phase of activity are less than 2 mm in diameter. Particles too large to pass the pylorus are crushed and ejected back into the antrum by the passing wave of peristalsis. Thus the peristaltic actions of the distal stomach walls serve not only to propel food but also, and perhaps more importantly, to grind and mix it.

Control of Gastric Motility Differs in the Proximal and Distal Stomach

The motility of the stomach, as in other smooth muscle portions of the gut, is under the control of regulatory molecules from the ENS and endocrine/paracrine systems. Fibers of the vagus nerve synapse on nerve cell bodies of the extensive gastric myenteric plexus and thereby exert a high degree of control over the ENS and thus over gastric motility. The effects of vagal stimulation on the proximal and distal regions of the stomach are opposite; in the proximal stomach, vagal activity suppresses muscular contractions and leads to adaptive relaxation, whereas in the distal stomach, vagal stimulation causes intense peristaltic activity. Vagal stimulation of distal antral motility is mediated by acetylcholine, but vagal inhibition of proximal stomach motility is not. The identity of the inhibitory mediator is not well established, but it may be vasoactive intestinal peptide.

Vagal action on the stomach is stimulated by events occurring within the CNS and within the stomach and intestine. The anticipation of food consumption causes vagal stimulation of the stomach and thus primes the stomach to receive a meal. Reactions of the GI tract that originate in the CNS in response to anticipated food intake are often referred to as the *cephalic phase* of digestion. Reactions to the cephalic phase are then augmented as food enters the stomach. In response to food in the stomach, vagal activity increases as sensory receptors in the stomach create a positive feedback loop.

The exact role of hormones in regulation of gastric motility is not completely established. *Gastrin,* which is secreted from cells in the gastric antrum, appears to enhance gastric motility. *Cholecystokinin* (CCK), *secretin,* and *gastric inhibitory peptide* (GIP) appear to suppress gastric motility, at least in the dog. The roles of the various GI hormones are difficult to determine from available information, because many of the experimental results reported have been in response to administration of GI hormones at amounts far greater than those normally occurring.

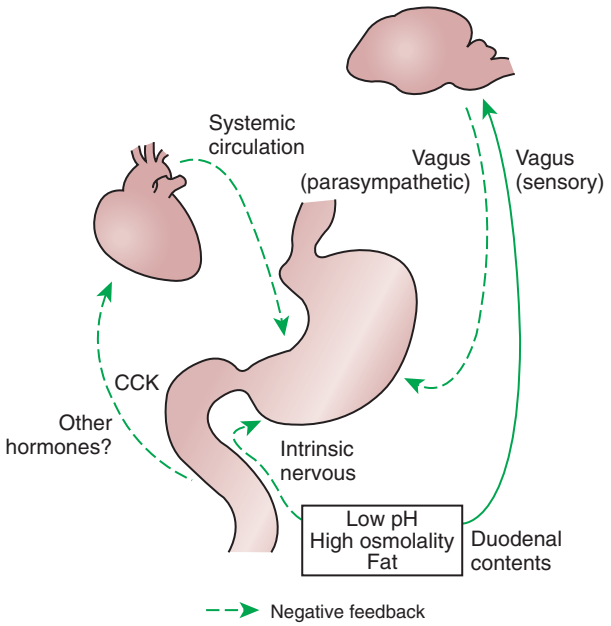
The Rate of Gastric Emptying Must Match the Small Intestine's Rate of Digestion and Absorption

The rate at which food leaves the stomach must match the rate at which it can be digested and absorbed by the small intestine. Because some types of foods can be digested and absorbed more rapidly than others, the rate at which the stomach empties must be regulated by the contents of the small intestine. Thus there are reflexes originating in the small intestine that may slow gastric emptying when the digestive capacity of the small intestine has been reached. These reflexes slow gastric emptying, allowing the stomach to serve as a storage site for ingesta while the small intestine "catches up" with its digestive and absorptive functions. The afferent receptors of these reflexes are in the duodenum and are activated by low pH, high osmolality, and the presence of fat. Separate sensory receptors apparently exist for each of these stimuli, but these receptors have not been identified anatomically.

Many reflexes occur within the GI system. Their names usually reflect the site of origin of the afferent stimulus and the site of the efferent response. Thus reflex control of gastric emptying by the duodenum is referred to as the *enterogastric reflex* ("entero" referring to the intestine).

The arc of the enterogastric reflex probably involves both the CNS and the ENS and the endocrine/paracrine system (Fig. 28.8). The extrinsic reflex pathway appears to involve afferent fibers of the vagus nerve, which receive stimuli in the duodenum. These stimuli are integrated in the brainstem, and the response is mediated by vagal efferent fibers to the stomach. The enteric reflex arc involves receptors in the duodenum and nerve fiber connections in the ENS that directly affect gastric emptying.

A contribution of the GI endocrine system to the enterogastric reflex has long been suspected, but the exact hormones responsible for the reflex are not known. CCK and secretin may be important. Both hormones are secreted by cells in the duodenum; CCK is secreted in response to fat and secretin in response to low pH; both appear to have suppression of gastric emptying as secondary



• **Fig. 28.8** Inhibitory arcs of the enterogastric reflex. Low pH, high osmolality, and the presence of fat in the duodenum stimulate vagal, enteric neuronal, and hormonal reflexes that inhibit stomach emptying. After the duodenal pH and osmolality have moderated and some of the fat has been absorbed, the inhibitory influences on the stomach are removed. *CCK,* Cholecystokinin.

effects. Gastric inhibitory peptide is a hormone produced in the duodenum in response to the presence of carbohydrate. In the dog, gastric inhibitory peptide may function as an inhibitor of gastric emptying, although stimulation of insulin secretion is probably its major action.

Enterogastric reflexes control gastric emptying by regulating stomach motility. The manner in which motility affects gastric emptying of solids is different from that for liquids. The rate at which solids are expelled from the stomach is regulated by the rate at which they are broken down into particles small enough to pass through the pylorus. This in turn is controlled by the motility of the antrum, or distal stomach; the greater the motility of the antrum, the faster material is broken down. Thus antrum motility regulates the rate of release of solid material from the stomach. Liquid material leaves the stomach more quickly than solid matter, and the release of liquid may be less dependent on antral motility than on the motility of the proximal stomach.

Minimal mixing activity occurs in the proximal stomach. Therefore liquids and solids tend to separate, the liquids moving to the outside and the solids to the center of the mass of food in the proximal stomach. Increased tension in the wall of the stomach body forces liquid into the antrum. Liquid may leave the antrum quickly, dependent on the activity of the pylorus. On the other hand, increased tension in the stomach body has little effect on the transport of solid material, because such material cannot leave the gastric body until sufficient space has been made available in the antrum. Thus motility of the stomach body appears primarily responsible for the liquid-emptying rate, whereas motility of the antrum is most responsible for the solid-emptying rate. The effect of the pylorus itself on gastric emptying is not as great as might be expected; removal of the pylorus results in a slight increase in the liquid-emptying rate and little increase in the rate of emptying of solid material. It appears that the distal portion of the antrum can account for much of the sieving action usually attributed to the pylorus. The rate of emptying of an isotonic liquid from the stomach is exponential and dependent on the initial volume of the liquid meal. Under usual circumstances, a liquid meal in the canine stomach has a half-life of about 18 minutes and is essentially gone by 1 hour after ingestion. Solid material is emptied more slowly, and its rate depends on its fat content. Low-fat meat meals are usually gone from the stomach in 3 to 4 hours after ingestion.

Between Meals, the Stomach Is Cleared of Indigestible Material

Some types of ingested materials, such as bone and indigestible foreign objects, cannot be reduced to particles less than 2 mm in diameter. During the digestive phase of gastric motility, such material does not leave the stomach. To clear the stomach of indigestible debris, a particular type of motility occurs between meals. This motility pattern is called the *interdigestive motility complex.* This motility pattern in the stomach is similar to, and probably continuous with, the *migrating motility complex* (MMC) of the small intestine. The latter will be discussed in the next section. In association with the interdigestive motility complex, the pylorus relaxes as strong waves of peristalsis sweep over the antrum, forcing less digestible material into the duodenum. This type of motility appears to have a "housekeeping" function in clearing the stomach of indigestible material.

The peristaltic waves of the interdigestive motility complex occur at approximately 1-hour intervals during the periods when the stomach is relatively empty of digestible material. Eating disrupts the complex and causes the resumption of the digestive motility pattern. Herbivores, which eat almost constantly, have a slightly different pattern; the interdigestive motility complex occurs at approximately hourly intervals, even with digestible food present in the stomach.

Vomiting Is a Complex Reflex Coordinated From the Brainstem

Vomiting is a complex reflex activity, and its integration, or coordination, is centered in the brainstem. The act of vomiting involves many striated muscle groups and other structures outside the GI tract. Vomiting is associated with the following actions:

1. Relaxation of the muscles of the stomach and lower esophageal sphincter, and closing of the pylorus.
2. Contraction of the abdominal musculature, creating an increase in intraabdominal pressure.
3. Expansion of the chest cavity while the glottis remains closed; this action lowers intrathoracic pressure and thus the pressure in the body of the esophagus.
4. Opening of the upper esophageal sphincter.
5. Antiperistaltic motility (peristaltic motility propelling ingesta toward the mouth) in the duodenum, which may precede the previous actions; thus vomiting may include ingesta of intestinal origin.

The efferent limb of this reflex arc involves motor fibers in many different peripheral nerves.

Afferent stimulation of the vomiting reflex comes from a large number of receptors. Of particular importance are mechanoreceptors in the pharynx and tension receptors and chemoreceptors in the

gastric and duodenal mucosa. Stimulation of these receptors sends signals to the *vomit center* in the brainstem. Thus noxious tactile or chemical stimulation of the GI mucosa can result in vomiting, clearing, or attempting to clear the offending stimulus from the GI tract. Direct irritation of GI structures, however, is not the only stimulus for vomiting. The vomit center receives afferent input from a variety of organs, thus vomiting is not always an indication of a primary GI problem.

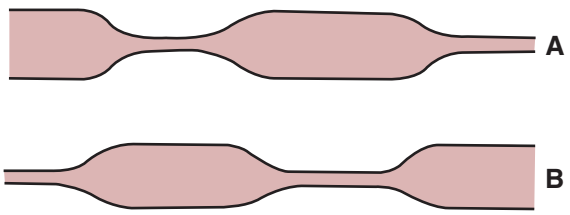
An important structure outside the GI tract that supplies afferent input in the vomit center is the *chemoreceptor trigger zone.* This is an area of the brainstem that lies in contact with the third ventricle. The chemoreceptor trigger zone is sensitive to the presence of some drugs and toxins in the blood. When stimulated, this zone sends signals to the vomit center and induces vomiting. Some of the products of inflammation stimulate the chemoreceptor trigger zone. Thus inflammatory disease, even outside the GI tract, can sometimes lead to vomiting.

The semicircular canals of the inner ear are other important structures that supply afferent input to the vomit center. Constant stimulation of the semicircular canals may induce vomiting, as occurs in motion sickness. Other sites in the body also may stimulate the vomit center; thus vomiting is a rather nonspecific sign of disease.

Motility of the Small Intestine Has Digestive and Interdigestive Phases

Motility of the small intestine occurs in two distinct phases: (1) during the digestive period after food intake and (2) during the interdigestive period when little food is present in the gut. In the digestive phase there are two primary motility patterns: propulsive and nonpropulsive. The *nonpropulsive* pattern is referred to as *segmentation.* Segmentation results from localized contractions of circular muscle. Portions of small intestine, usually 3 to 4 cm long, contract tightly, dividing the gut into segments of constricted and dilated lumen. Within a few seconds, the constricted portions relax and new areas constrict (Fig. 28.9). This action tends to "milk" gut contents back and forth within the small intestine, mixing them with digestive juices and circulating them over the absorptive mucosal surfaces. This type of motility does not contribute much to the net aboral propulsion of ingesta. In fact, segmentation tends to slow down the aboral movement of material because of closure of the intestinal lumen in the constricted segments.

Propulsive activity during the digestive phase consists of peristaltic contractions that migrate down the gut in phase with the slow waves. Digestive-phase peristaltic contractions pass over short segments of intestine and then die out. Thus ingesta are pushed down the gut for a short distance and then subjected to additional segmentation contractions and mixing activity. The interaction of segmentation and peristaltic motility has caused some authors to describe the movement of ingesta during digestive-phase motility as "two steps forward, one step back."



• **Fig. 28.9** Segmentation in the small intestine. (A) Areas of circular muscle constriction close the lumen and divide the gut into dilated segments containing ingesta. (B) At periodic intervals the areas of constriction and dilation alternate, exerting a mixing and circulating action on the ingesta.

The interdigestive phase of small-intestinal motility is characterized by waves of powerful peristaltic contractions that sweep over a large length of small intestine, sometimes traversing the entire organ. These waves are referred to as the *migrating motility complex* (MMC) or, alternatively, the *migrating myoelectric complex.* The MMC begins in the duodenum as groups of slow waves that stimulate intense action potential and muscular contraction activity. The complex migrates down the intestine at the rate of the slow waves. Some of the MMCs die out before reaching the ileum, but some travel the entire length of the small intestine.

The MMC is the basic motor activity in the gut during a fasting or the period between meals, which may be referred to as interdigestive states. Meal consumption interrupts the MMC. The MMC typically lasts 80 to 120 minutes and consists of three successive phases: phase I (60–70 min), which has no contractions; phase II (20–30 min), which has intermittent and irregular contractions; and phase III, which has strong peristaltic contractions that last 3 to 10 minutes and that start from the stomach and lower esophagus and migrate distally to reach the colon.

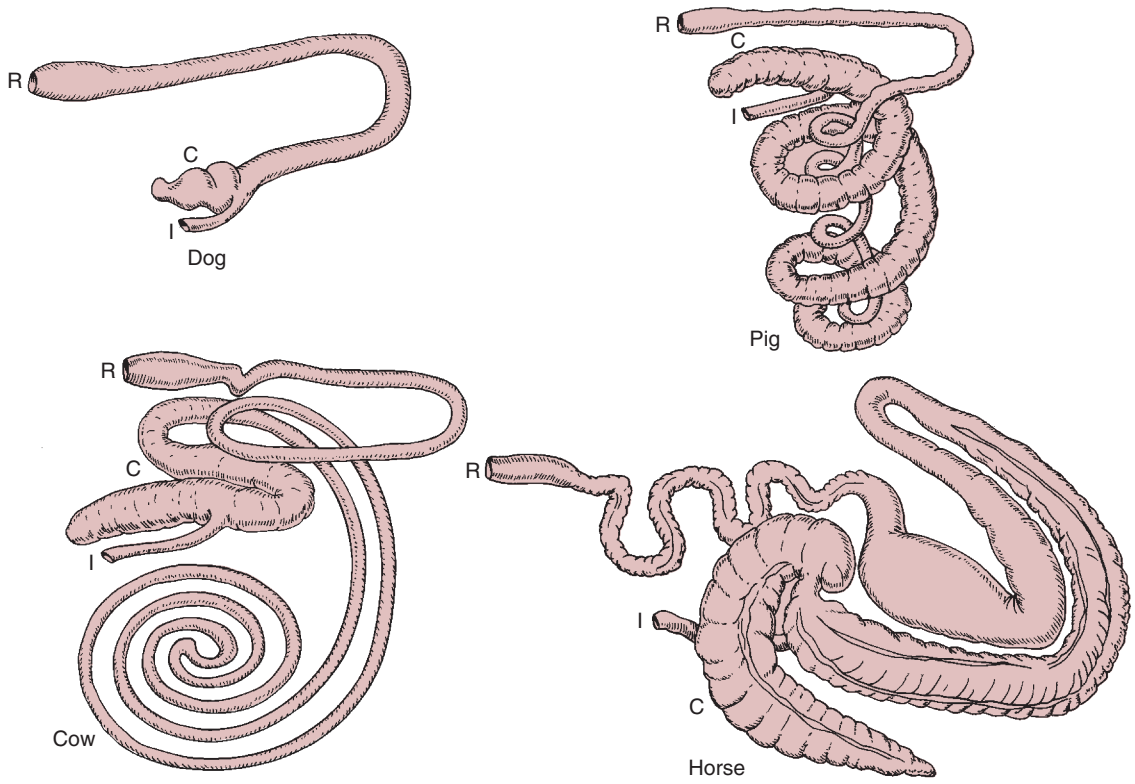
The MMC probably has a "housekeeping" function and serves to push undigested material out of the small intestine. The MMC is also important in controlling the bacterial population in the upper gut. Normally the duodenum harbors a relatively small population of bacteria, and the population increases distally into the ileum, which has a moderately large number of bacterial organisms. The colon is heavily colonized by numerous species of bacteria. It is important for digestive function that this relative distribution of bacteria be maintained within the gut. The MMC may help to impede the migration of bacteria from the ileum to the duodenum.

The Ileocecal Sphincter Prevents Movement of Colon Contents Back Into the Ileum

The *ileocecal sphincter* is at the junction of the small and large bowel and prevents the retrograde movement of colon contents into the ileum. It consists of a well-developed ring of circular muscle that remains constricted at most times. In addition to the muscular sphincter, in many species there is a flap of mucosa that acts as a one-way valve, further blocking movement of colon contents into the ileum. During periods of peristaltic activity in the ileum, the sphincter relaxes, allowing movement of material into the colon. When colonic pressure increases, the ileocecal sphincter constricts more tightly.

Motility of the Colon Causes Mixing, Retropulsion, and Propulsion of Ingesta

The colon has multiple functions, including (1) absorption of water and electrolytes, (2) storage of feces, and (3) fermentation of organic matter that escapes digestion and absorption in the small intestine. The relative importance of these functions varies with the species, and tremendous differences in colon size and shape exist among animals. The major determinant of colon size is the importance of colonic fermentation to the energy needs of



• **Fig. 28.10** Variations of colon anatomy of four mammals. Animals with simple colons, such as the dog, are not dependent on colonic fermentation to supply energy needs. Horses, which have tremendous colonic development, rely on colonic fermentation for a large portion of their energy needs. In animals such as pigs and cattle, the importance of colonic fermentation to digestive needs is intermediate between that of the horse and dog, and this intermediate position is noted in their colon development. *C,* Cecum; *I,* ileum; *R,* rectum.

the animal. Some species, such as the horse and rabbit, make extensive use of fermentation products for nutritional needs and have large and complex colons. (The ruminant's fermentation chamber is in the stomach.) Other species, such as the dog and cat, do not rely on fermentation products and have relatively simple colons. Fig. 28.10 illustrates differences in colonic anatomy among four species with different needs for fermentative digestion.

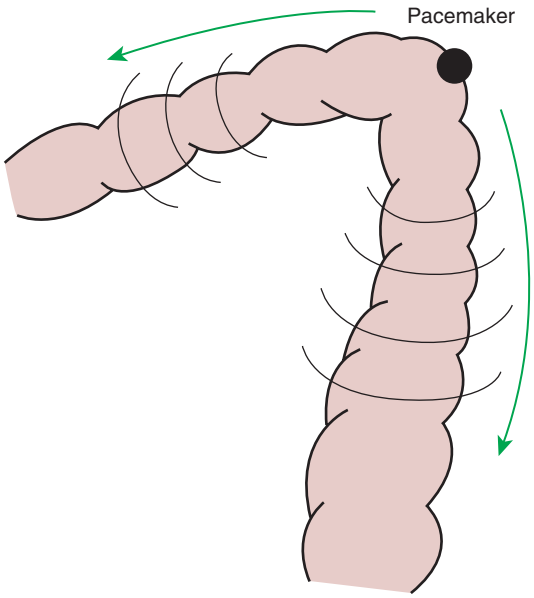
Considerable similarity appears to exist in colon motility patterns among animals, despite the anatomical diversity. *Mixing* activity is prominent in the colons of all species because mixing and circulating are important to both absorptive and fermentative functions. Mixing is achieved by segmentation contractions along with other types of motility. In many species, such as the horse and pig, colonic segmentation is pronounced and in some areas results in the formation of sacculations, known as *haustra,* which are visible even after death.

A particular characteristic of colonic motility is *retropulsion,* or *antiperistalsis.* This type of peristaltic contraction migrates orally, the opposite of normal peristaltic movement. Such motility results from colonic slow-wave activity that is somewhat more complex than that of the small intestine. In the colon, as in the small intestine, the slow waves originate in the ICC. The colonic ENS, however, can influence the ICC in such a manner as to shift the site of origin of slow waves and the direction of their propagation. Under resting conditions in the colon, slow waves originate from *pacemakers* in one or more central sites. The pacemakers are not anatomical structures; rather, they are areas defined by activities of the ENS. Thus the pacemakers are not always the same areas; they can disappear and form in different locations in response to the need for different motility patterns. Antiperistaltic contractions occur in the segments in which slow waves migrate in an oral direction. Antiperistaltic contractions are retropulsive and impede the movement of ingesta, causing intense mixing activity and forcing material to accumulate in the proximal portions of the colon. Retropulsion appears to be particularly strong near the pacemakers, and the pacemakers therefore represent sites of high resistance to the flow of colonic ingesta.

Because of continued inflow of material from the ileum into the colon, some ingesta escape the retropulsive antiperistaltic motility, move into areas of propulsive peristaltic activity, and proceed along the colon. In addition, there are periods of intense propulsive activity that involve the entire colon. These are called *mass movements* and frequently involve the distal translocation of the entire colonic content.

The Colon Is an Important Site of Storage and Absorption in All Animals

The colon of the dog and cat is a relatively simple organ consisting of a short cecum, an ascending part, a transverse part, and a descending part. During the resting phase, there is a colonic pacemaker at about the junction of the transverse and descending colons (Fig. 28.11). This gives rise to antiperistaltic activity in the proximal colon with resultant accumulation of ingesta in the cecum and ascending colon areas. Moderate peristaltic activity usually



• **Fig. 28.11** A pacemaker is present at the junction of the transverse and descending parts of the colon of the cat and probably in other mammals with similar colonic anatomy. Slow waves and peristaltic activity emanate in both directions from the pacemaker. Retrograde, or reverse, peristalsis in the proximal portions of the colon causes ingesta to be retained there, promoting the storage and absorptive functions of the colon.

occurs in the descending colon, whereas the distal colon and rectum are usually constricted and empty.

Material entering the carnivore colon is of a fluid consistency. It is thoroughly mixed in the ascending and transverse colons, and much of the water and many of the electrolytes are absorbed. By the time it reaches the descending colon, it is semisolid and becoming feces.

Despite Large Anatomical Differences in the Colons of Herbivores Compared With Omnivores and Carnivores, There Are Similarities in Motility

There are important similarities in motility among various species, even those with extensive anatomical differences. The purpose of this discussion is to describe the similarities in motility between species with simple colons and those with complex colons. Chapter 31 provides a more extensive discussion of the highly developed colons of herbivores.

The equine hindgut, as an example of an herbivore hindgut, is complex and highly developed (see Fig. 28.10). The cecum is large and separated into haustra. The equine cecum is unique among ceca of most species, even other herbivores, because a distinct, sphincter-like orifice joins it to the colon. The colon is divided into a large and a small portion, and the large colon is folded on itself so that three distinct flexures exist. The longitudinal muscles of the cecum and most areas of the colon are not evenly dispersed around the circumference of the gut. Instead, they form discrete bands, or teniae, that course along the longitudinal axis of the gut. The teniae divide the haustra longitudinally, giving the equine cecum and large colon a sacculated appearance.

Motility in the equine cecum consists of active segmentation and mixing, along with occasional mass movements that appear to transfer large amounts of ingesta to the colon. Motility in the colon consists of segmentation, antiperistalsis, and peristalsis. A colonic pacemaker appears to exist at the pelvic flexure and creates an area of high resistance to flow that results in prolonged retention of material in the ventral portions of the large colon. The pelvic flexure pacemaker in the equine colon is similar in function to the colonic pacemaker in the transverse colon of the dog and cat. Little is known about regulation of motility in the equine small colon. The characteristic ball-shaped form of equine feces probably represents intense segmentation-type motility in the small colon, where the feces are formed (see Chapter 31).

In ruminants and swine, the hindgut consists of a cecum of intermediate complexity, a spiral colon, and a straight colon. Compared with other species, less is known about the hindgut motility of animals with spiral colons. There seems to be an area of high flow resistance at the flexure, or central point, of the spiral colon. This site of flow resistance may represent a pacemaker that generates antiperistaltic motility in the centripetal portion of the colon.

The Anal Sphincter Has Two Layers With Separate Innervation

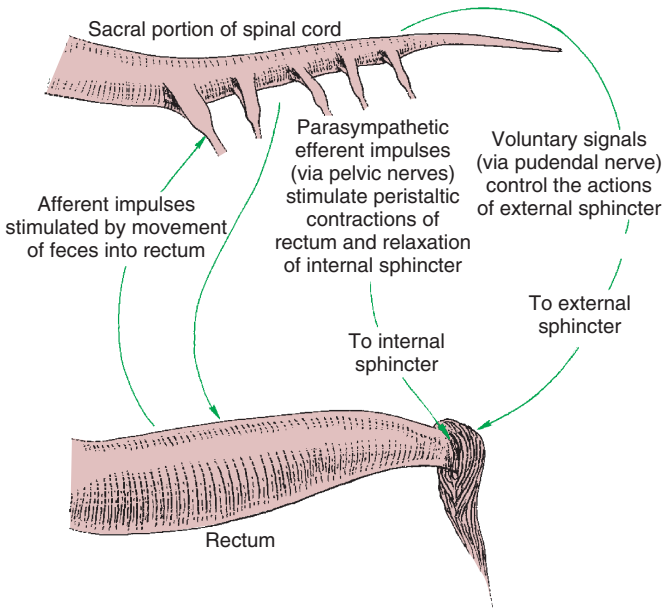
The anal opening is constricted by two sphincters: an *internal sphincter* of smooth muscle, which is a direct extension of the circular muscle layer of the rectum, and an *external sphincter* of striated muscle. The internal anal sphincter usually remains tonically contracted and is responsible for anal continence. The internal sphincter receives parasympathetic innervation from the sacral spinal segments through the pelvic nerve and sympathetic innervation from the lumbar spinal segments through the hypogastric nerve. In most species, sympathetic stimulation results in constriction of the sphincter, and parasympathetic stimulation results in relaxation.

The external sphincter maintains some degree of tonic contraction, but the consistent tone of the anus is primarily regulated by the internal sphincter. The external sphincter is innervated by general somatic efferent fibers that have cell bodies in the cranial sacral spinal segments and course in the pudendal nerve.

The Rectosphincteric Reflex Is Important in Defecation

The entry of feces into the rectum is accompanied by the reflex relaxation of the internal anal sphincter, followed by peristaltic contractions of the rectum. This is known as the *rectosphincteric reflex* and is an important part of the act of defecation (Fig. 28.12). The reflex normally results in defecation, but in trained animals its effects can be blocked by voluntary constriction of the external anal sphincter. When defecation is voluntarily suppressed, the rectum soon relaxes to accommodate the fecal bolus, and the internal anal sphincter regains its tone. In humans, and presumably in dogs and cats, relaxation of the rectum and constriction of the internal sphincter are associated with fading of the urge to defecate, until another bolus of feces enters the rectum.

Uninhibited animals respond to the presence of feces in the rectum with a number of voluntary actions associated with defecation. In carnivores, the diaphragm and abdominal muscles contract



• **Fig. 28.12** Arcs of the rectosphincteric reflex. The reflex is initiated by the movement of feces into the rectum and results in peristaltic movements of the rectal wall and relaxation of the internal anal sphincter. Fecal passage is the normal effect of the reflex, but voluntary constriction of the external anal sphincter can prevent the passage of feces and eventually override the reflex, apparently enabling trained animals to suppress the urge to defecate.

to increase intraabdominal pressure, and striated muscles of the anal canal relax as the animal assumes the defecation posture. These acts are important for complete evacuation of the rectum.

Major Differences Between Avian and Mammalian Digestive Systems Include, in Birds, Both the Lack of Teeth and the Separation of Gastric Functions Into Distinct Anatomical Regions

Important anatomical differences exist between the digestive systems of birds and mammals. These differences affect motility functions more than other aspects of digestion, such as secretion, digestion, and absorption. Therefore the digestive tract of birds is covered in this chapter as a separate topic. In other chapters of this section, aspects of avian digestion are integrated into the general discussion.

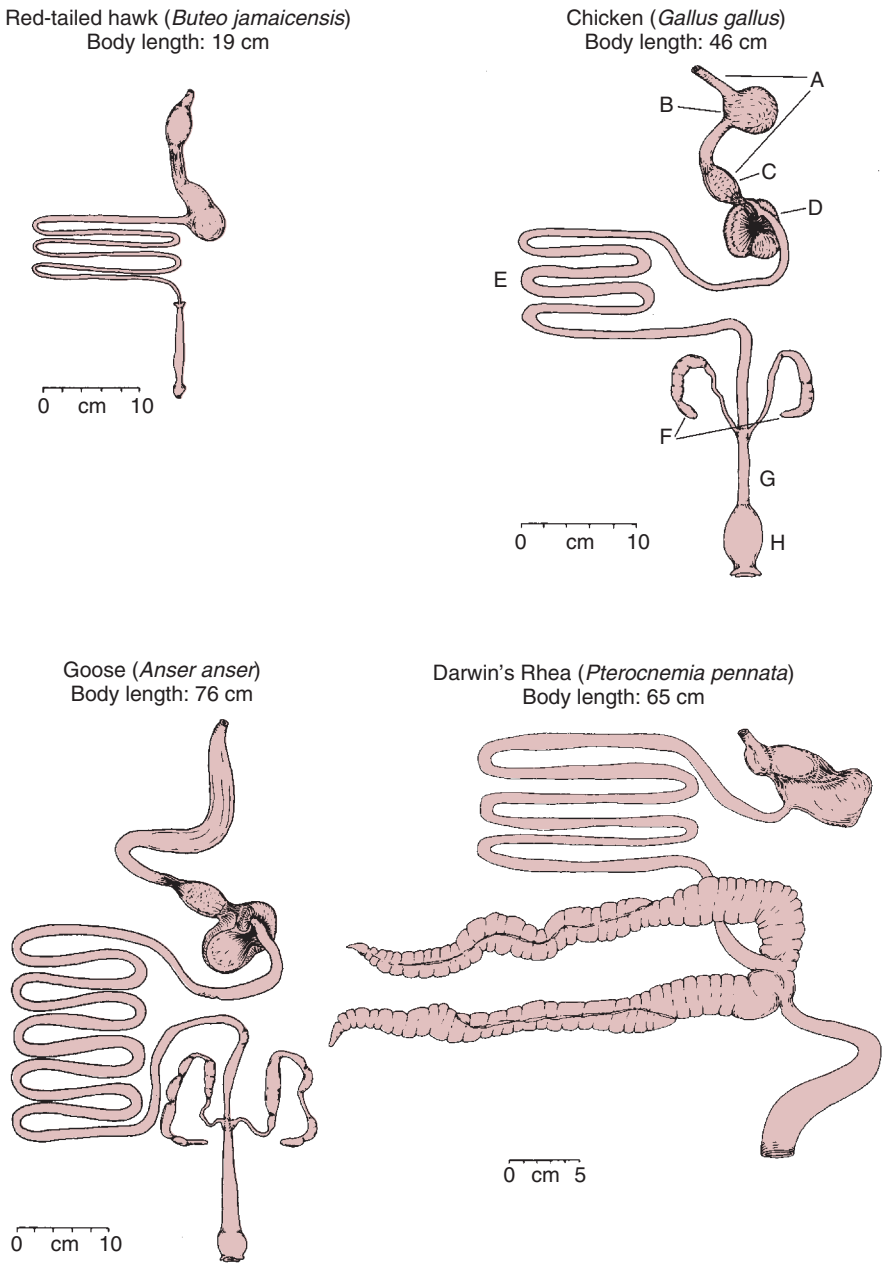
Fig. 28.13 illustrates the general anatomy of the avian digestive system. The pharynx of birds is simpler than that in mammals, with birds having no soft palate. There are no teeth, although in carnivorous species, the beak is modified for tearing food into pieces small enough to swallow. The esophagus has a large diameter so as to accommodate unmasticated food. An outpouching of the esophagus is known as the *crop.* The extent of development of the crop varies widely among avian species. The glandular portion of the stomach is the *proventriculus,* which is separated from the muscular stomach, known as the *ventriculus* or *gizzard,* by a short isthmus. The small intestine varies greatly in length among avian species but is generally rather short compared with that of mammals of similar size. The ceca are usually paired and vary tremendously in development among avian species. In some carnivorous species, such as the hawk, the ceca are rudimentary, whereas in some of the nonflying herbivores, such as the ostrich, cecal development is extensive (see Fig. 28.13). The colon and rectum are very simple; the rectum ends in the cloaca, which is a common passageway for digestive, urinary, and reproductive discharges.

The crop performs a storage function. In some species, the crop is little more than a pouch in the esophagus, whereas in others, such as the chicken, there is a distinct sphincter-like opening between the esophagus and crop. In general, ingesta do not begin to accumulate in the crop until the gizzard is full. The crop is richly populated with mucus-secreting cells, but no digestive glands are present. However, digestive glandular secretions originating from the salivary glands and proventriculus are present in the crop. In many species, it appears that ingesta and secretions pass in a retrograde manner up the esophagus from the gizzard and proventriculus to the crop. The motility of the crop is under regulation of vagal impulses. Crop motility and rate of emptying are coordinated so as to release ingesta at a rate matching the emptying rate of the proventriculus and gizzard. In some avian species, the crop also functions as a storage place for food being transported to the young. In this case, food is swallowed into the crop and later regurgitated as feed for offspring.

The proventriculus is a low-volume organ with a glandular epithelium resembling that of the stomach of mammals (see Chapter 29). Motility functions of the proventriculus are to propel ingesta and digestive secretions into the gizzard for mixing and grinding. The gizzard is a muscular organ that grinds and liquefies ingesta. In addition, particle-size discrimination occurs in the gizzard; small particles are passed into the duodenum, whereas large particles are retained for further comminution or ejected back into the proventriculus for further addition of digestive secretions. In carnivorous birds, concretions of bone, hair, feathers, and other indigestible material accumulate in the gizzard and are occasionally ejected orally in an action known as *egestion.* In grain-eating birds, small stones or gravel are swallowed and retained in the gizzard to aid in the comminution of ingesta. This inorganic material is referred to as *grit,* and its presence increases digestive efficiency, although it is not essential. The mucosa of the gizzard is covered by a tough coating known as *koilin.* This coating is composed of glandular secretions and desquamated cells. It protects the mucosa from the physical grinding actions of the gizzard.

The motility and function of the various stomach areas of birds are easily comparable with gastric motility and function in mammals. The crop and proventriculus function in much the same manner as the fundus and body of the mammalian stomach, with storage and secretory functions. The gizzard functions in much the same way as the antrum of the mammalian stomach, with grinding and particle-size discrimination functions. The major functional differences between birds and mammals include the physical separation of the stomach compartments in birds and the advanced grinding function of the gizzard.

The motility patterns of the avian small intestine appear to be generally similar to those in mammals. The motility of the hindgut also shares some characteristics of other animals. Reverse peristalsis is a dominant characteristic of the avian colon and rectum, moving ingesta into the ceca. Urinary excretions arriving at the cloaca become incorporated with ingesta and move in a retrograde manner into the ceca, thus facilitating reabsorption of the remaining water and electrolytes from urine. Cecal motility is characterized primarily by mixing and reverse peristalsis, with occasional mass movements resulting in the evacuation of the ceca. These mass movements in avian species are followed by defecation.



• **Fig. 28.13** Comparative anatomy of the digestive tracts of four avian species. Note the variation in crop and cecal development. The carnivorous red-tailed hawk has a small crop and rudimentary ceca. The chicken has a well-developed crop, and the rhea has tremendous cecal development. *A,* Esophagus; *B,* crop; *C,* proventriculus; *D,* gizzard or ventriculus; *E,* small intestine; *F,* ceca; *G,* rectum; *H,* cloaca. (From Stevens CD. *Comparative Physiology of the Vertebrate Digestive System,* Cambridge, UK: Cambridge University Press; 1988.)

CLINICAL CORRELATIONS

Equine Rabies

History

Owners report that their horse has not "been itself" for the past few days. Today, the animal is extremely lethargic and stands with forelegs wide apart and head held low. The nostrils are soiled, and the owners report that water and feed come out of the nostrils when the animal attempts to eat or drink.

Clinical Examination

From the history and presenting signs, you recognize that the horse may have paralysis of muscles of the pharynx and larynx. Because these lesions are typically associated with rabies in horses, you don a pair of plastic gloves and sleeves and proceed with your examination. To assess the function of the swallow reflex, you attempt to pass a stomach tube. You observe that the swallow reflex appears to be diminished, but with some persistence, the tube can be passed. This indicates that there is no physical obstruction in the pharynx or esophagus, and the problem is functional. These findings support, but do not confirm, a diagnosis of rabies.

Comment

Rabies in herbivorous animals may take a number of forms. One of the most common signs in cattle and horses is paralysis of the pharynx and larynx as a result of viral lesions in the brainstem nuclei supplying the appropriate cranial nerves. If rabies is suspected, no one should come into direct contact with the excretions of the animal, especially saliva.

CLINICAL CORRELATIONS—cont'd

Treatment

In this horse, treatment should consist of oral fluid and electrolyte therapy administered through an indwelling stomach tube. If there is no response to this conservative therapy, and if the animal's condition appears to deteriorate, euthanasia is necessary, and the animal's head should be submitted for evaluation for a positive diagnosis of rabies.

Horse With Colic Due to Large Colon Displacement

History

A 16-year-old thoroughbred gelding presents to you for colic. The trainer noticed that he was fine when he was checked this morning and fed. The barn staff walked by the stall a little bit later, and the horse was down, rolling, and very sweaty. They could not keep him comfortable, so they brought him straight to your clinic.

Clinical Examination

He has injected (dark red) and has tacky mucous membranes with prolonged capillary refill time (CRT), demonstrating dehydration and poor perfusion. His temperature, pulse, and respiratory rate are all increased. He is sweating and trying to lay down. After sedating him, you pass a nasogastric tube, and you get a lot of reflux back which smells fetid, suggesting that ingesta could have been sitting in the GI tract without moving much and rotting and/or there could be necrotic bowel present. On palpation per rectum, it feels as though he has displaced his large intestine. Blood work supports a severe inflammatory process with low electrolytes and albumin.

Comment

You take him to surgery and put his large intestine back in place. Unfortunately, he has a compromised small intestine that needs to be resected. When he recovers from surgery, he is put on intravenous fluids to rehydrate him and correct his electrolyte deficits, which likely include sodium, chloride, and calcium. He is also given a few liters of plasma to correct his deficit in albumin, and both help with oncotic pressure (favoring capillary fluid retention) and his immune system. He is kept on intravenous antibiotics, K-penicillin, and gentamicin. Unfortunately, he has poor GI motility (ileus). Ileus can result for several reasons. GI motility can be decreased immediately after surgery because of the effect of anesthesia. In addition, endotoxemia, which is common in animals experiencing colic, may result in ileus. Endotoxins are molecules derived from the cell walls of gram-negative bacteria. These substances, which are normally confined to the gut lumen, may enter the vascular system of animals suffering from colic or other conditions affecting mucosal integrity. Endotoxin can induce activation of α_2 adrenoreceptors and the production of certain prostanoids, which limit GI motility. Moreover, ileus may be caused by a reduction in the number of ICC, which control the electrical activity in the intestine. This appears to be a response to the surgery. In many cases, we may not know the exact cause of the ileus.

Treatment

Restoring motility is often multimodal. Horses are often given antiinflammatory medication to help limit inflammation. Hand walking is encouraged to help restore motility. Correction of electrolyte deficits is critical for normalization of muscle contractions, which influence motility. In addition, motility modifying drugs, such as lidocaine or yohimbine, may be used. Lidocaine may suppress primary afferent neurons, limiting reflex efferent inhibition of motility. Its antiinflammatory properties are also thought to help improve mucosal repair. Yohimbine is another drug that is used. It acts as an α_2 adrenoreceptor antagonist to counteract increased sympathetic outflow. Erythromycin has also been used to restore motility. It is a direct motilin agonist on smooth muscle cells to stimulate release of acetylcholine and motilin within the ENS. Clinicians usually start with just one motility modifying drug to determine whether it is beneficial. Lidocaine is not infrequently chosen. Clinicians will monitor the horse's progress in addition to the other supportive treatments. Medications are adjusted as needed to help restore motility.

Practice Questions

1. A unique feature of gastrointestinal (GI) smooth muscle cells is that:
- **a.** Their resting transcellular electrical potential has the positive pole on the outside surface of the cell membrane.
- **b.** Action potentials, or spikes of membrane depolarization, are not associated with muscle contractions.
- **c.** Muscle contractions are stimulated by partial depolarization of the membrane.
- **d.** There are spontaneous, rhythmical undulations in the electrical potential across the cell membrane.
- **e.** Contraction of the muscles is never influenced by nervous activity.

2. The interstitial cells of Cajal are:
- **a.** Modified neurons capable of generating contraction.
- **b.** Modified neurons capable of generating only action potentials.
- **c.** Modified neurons capable of generating only slow waves.
- **d.** Modified smooth muscle cells capable of generating only slow waves.
- **e.** Modified smooth muscle cells capable of generating only action potentials.

3. The term *slow waves* as applied to the gut refers to:
- **a.** Slowly moving fronts of electrical activity that are propagated down the enteric nervous system.
- **b.** Slowly moving fronts of electrical activity that result from coordinated changes in cell membrane potential occurring throughout the smooth muscle of the intestinal wall.
- **c.** Slowly moving fronts of ingesta that proceed down the intestine in response to peristaltic movement.
- **d.** Slowly moving fronts of action potentials that are constantly passing over the gut smooth muscle.
- **e.** Slowly moving fronts of peristaltic contractions that pass uniformly over the entire small intestine during the digestive period.

4. An animal is presented to you with aspiration pneumonia (the result of food material entering the lower respiratory tract). Which of the following lesions would be a likely cause?
- **a.** Loss of myenteric plexus function in the pharynx and upper esophagus
- **b.** Loss of slow-wave activity in the pharynx and upper esophagus
- **c.** A lesion in the brainstem
- **d.** A lesion in the trachea
- **e.** None of the above

5. The term *cephalic phase* is used in reference to a number of activities occurring in the GI tract. In general, the term means:
   a. The early phases of digestion, when food is nearest the head.
   b. Any actions stimulated directly by the presence of food in the stomach.
   c. Any actions stimulated directly by the presence of food in the mouth.
   d. Digestive events stimulated by the presence of food in the GI tract but requiring reflexes integrated in the central nervous system.
   e. Digestive events that occur before the ingestion of food and in response to central nervous system stimulation brought on by the anticipation of eating.

6. Conditions in the duodenum, such as low pH or high fat concentration, can reflexively inhibit gastric emptying. Which reflex arc is involved in this inhibition?
   a. Parasympathetic nervous system
   b. GI enteric nervous system
   c. GI endocrine system
   d. All the above

7. Which of the following best describes the motility of the proximal region of the monogastric stomach?
   a. Rhythmic segmentation
   b. Peristalsis
   c. Retropulsion
   d. Adaptive relaxation

8. Which of the following is characteristic of the interdigestive phase of small intestinal motility?
   a. Migrating motility complexes consisting of waves of peristaltic contractions that pass over the entire length of the small intestine
   b. Rhythmic segmentation
   c. Short waves of peristalsis that die out after a few centimeters
   d. Complete relaxation of small intestinal smooth muscle

9. Which of the following aspects of colon physiology is common to many species, irrespective of interspecies anatomical differences in colon structure?
   a. Rapid flow of ingesta
   b. Adaptive relaxation
   c. Retropulsion, or antiperistalsis
   d. Haustra formation

10. Colonic "pacemakers":
    a. Are anatomically distinct structures composed of specialized smooth muscle cells.
    b. Shift in their sites under the influence of the ENS.
    c. Are involved in segmentation but not peristalsis.
    d. Control defecation.

11. The rectosphincteric reflex is integrated in the:
    a. Brainstem
    b. ENS
    c. Lumbar spinal cord
    d. Sacral spinal cord

Bibliography

Entry-Level Supplements

Barrett KE. *Gastrointestinal Physiology (Chapters 7-9)*. McGraw-Hill Education; 2014.

Johnson LR. *Gastrointestinal Physiology (Chapters 2-6)*. 8th ed. Mosby; 2014.

Advanced Reading

Bharucha AE, Fletcher JG. Recent advances in assessing anorectal structure and functions. *Gastroenterology*. 2007;133(4):1069–1074.

Denbow DM. Gastrointestinal anatomy and physiology. In: Scanes GC, eds. *Sturkie's Avian Physiology*. 6th ed. Elsevier; 2015.

Dinning PG, Arkwright JW, Costa M, et al. Temporal relationships between wall motion, intraluminal pressure, and flow in the isolated rabbit small intestine. *Am J Physiol Gastrointest Liver Physiol*. 2011;300:G577–G585.

Hanani M, Farrugia G, Komuro T. Intercellular coupling of interstitial cells of Cajal in the digestive tract. *Int Rev Cytol*. 2005;242:249–282.

Hanani M, Freund HR. Interstitial cells of Cajal: their role in pacing and signal transmission in the digestive system. *Acta Physiol Scand*. 2000;170(3):177–190.

Horowitz B, Ward SM, Sanders KM. Cellular and molecular basis for electrical rhythmicity in gastrointestinal muscles. *Annu Rev Physiol*. 1999;61:19–43.

Johnson LR, ed. *Physiology of the Gastrointestinal Tract (Chapters 33-36)*. 5th ed. Elsevier; 2012.

Kerlin P, Zinsmeister A, Phillips S. Relationship of motility to flow of contents in the human small intestine. *Gastroenterology*. 1982;82:701.

Merritt AM. Normal equine gastroduodenal secretion and motility. *Equine Vet J Suppl*. 1999;29:7–13.

Pfannkuche H, Gaebel G. Characteristics and control of equine intestinal motility. *Pferdeheilkunde Equine Med*. 2007;23(4):333.

Plumb D. *Veterinary Drug Handbook*. 4th ed. Ames, IA: Iowa State Press; 2002.

Podolsky DK, Camilleri M, Fitz JG, eds. *Yamada's Textbook of Gastroenterology (Chapters 3-6)*. John Wiley & Sons, Ltd; 2016.

Rasmussen OO, Christiansen J. Physiology and pathophysiology of anal function. *Scand J Gastroenterol Suppl*. 1996;216:169–174.

Reed SM, Bayly WM, Sellon DC. *Equine Internal Medicine*. 4th ed. St. Louis, MO: Saunders.; 2018.

Smith BP. *Large Animal Internal Medicine*. 5th ed. St. Louis, MO: Mosby Elsevier; 2015.

Stevens CE, Hume ID. *Comparative Physiology of the Vertebrate Digestive System*. 2nd ed. Cambridge, UK: Cambridge University Press; 1996.

29

Secretions of the Gastrointestinal Tract

THOMAS H. HERDT

KEY POINTS

The Salivary Glands

1. Saliva moistens, lubricates, and partially digests food.
2. Salivary secretions originate in the gland acini and are modified in the collecting ducts.
3. Salivary glands are regulated by the parasympathetic nervous system.
4. Ruminant saliva is a bicarbonate-phosphate buffer secreted in large quantities.

Gastric Secretion

1. Depending on the species, there may be two general types of gastric mucosa: glandular and nonglandular.
2. The gastric mucosa contains many different cell types.
3. The parietal cells of the gastric glands secrete hydrochloric acid.
4. The chief cells of the gastric glands secrete pepsin. It is secreted in an inactive form and is subsequently activated in the stomach lumen.
5. The parietal cells are stimulated to secrete by the action of acetylcholine, gastrin, and histamine.

The Pancreas

1. Pancreatic exocrine secretions are indispensable for digestion of complex nutrients: proteins, starches, and triglycerides.
2. Acinar cells secrete enzymes, whereas centroacinar cells and duct cells secrete an electrolyte solution rich in sodium bicarbonate.
3. Pancreatic cells have cell surface receptors stimulated by acetylcholine, cholecystokinin, and secretin.

Bile Secretion

1. The liver is an acinar gland with small acinar lumina known as *canaliculi*.
2. Bile contains phospholipids and cholesterol maintained in aqueous solution by the detergent action of bile acids.
3. The gallbladder stores and concentrates bile during the periods between feeding.
4. Bile secretion is initiated by the presence of food in the duodenum and stimulated by the return of bile acids to the liver.

Digestion and absorption can take place only in the aqueous milieu of digestive secretions. Synthesis and secretion of these fluids represent a well-controlled process regulated by endocrine, paracrine, and neural events. The total volume of digestive secretions is large, with the daily amount substantially larger than the volume of fluid ingested over a similar period. In addition, most of the digestive secretions have a relatively large concentration of electrolytes. This large outpouring of fluid and electrolyte into the gut makes reabsorption of these secretions imperative if fluid and electrolyte homeostasis of the body is to be maintained. Indeed, one of the major life-threatening ramifications of digestive diseases is the loss of water and electrolytes from the body caused by inadequate reabsorption of digestive secretions.

The Salivary Glands

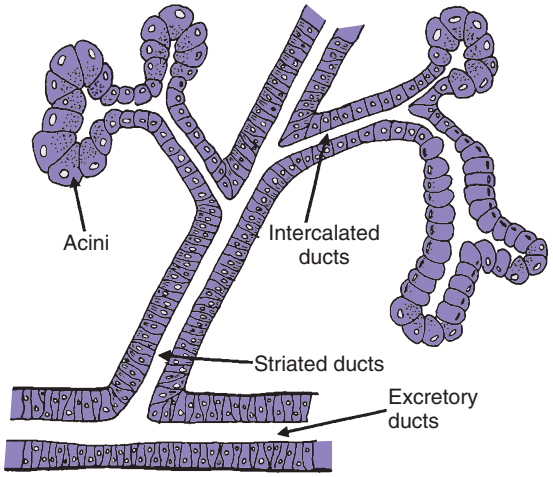
Saliva Moistens, Lubricates, and Partially Digests Food

As food is chewed, it is mixed with salivary secretions that allow it to be molded into well-lubricated boluses that facilitate swallowing. In addition, saliva may have antibacterial, digestive, and evaporative cooling functions, depending on the species.

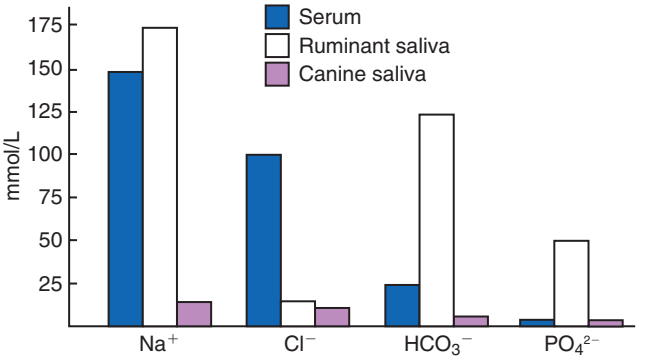
The antibacterial activity of saliva results from antibodies and antimicrobial enzymes known as *lysozymes*. Initially, you may think that the antibacterial properties of saliva are inefficient because the mouth normally contains a large, thriving population of bacteria. However, saliva aids in keeping this population in check, and animals with impaired salivary function are prone to infectious diseases of the oral cavity.

In omnivorous animals, such as rats and pigs, saliva contains a starch-digesting enzyme known as *salivary amylase*. This enzyme is usually absent from the saliva of carnivorous animals, such as cats. The saliva of some species also contains a fat-digesting enzyme known as *lingual lipase*. This enzyme is frequently present in young animals, such as calves, while they are on a milk diet; the enzyme disappears as they mature.

Salivary enzymes probably have their major digestive effect in the proximal stomach, because food is not retained in the mouth long enough to permit extensive digestion. The lack of mixing activity in the proximal stomach may be essential for the starch-digesting function of saliva. This is because the amylase enzyme is functional at neutral to slightly basic pH, which characterizes saliva. The low pH of the distal stomach probably inactivates the enzyme; therefore it may be important that food entering the stomach initially not be mixed with gastric secretions, so as to allow the salivary enzymes some time to work before being



• **Fig. 29.1** Schematic illustration of the salivary gland. Saliva is initially secreted by the acinar cells and then modified as it passes through the intercalated, or collecting, ducts. Modification of acinar secretions by duct epithelia is a physiological phenomenon common among several types of glands, including the pancreas.



• **Fig. 29.2** Electrolyte composition of blood serum and of canine and ruminant saliva. Note that the electrolyte concentration of canine saliva is much lower than that of serum, in contrast to the concentration in ruminant saliva. Also note the high concentrations of bicarbonate (HCO_3^-) and phosphate (PO_4^{2-}) in ruminant saliva; these ions give ruminant saliva its alkalizing quality.

inactivated by gastric acid. Some birds have salivary amylase that is active in the environment of the crop.

The evaporative cooling function of saliva is covered in Chapter 53.

Salivary Secretions Originate in the Gland Acini and Are Modified in the Collecting Ducts

The salivary gland is a typical *acinar* gland composed of an arborizing system of collecting ducts that end in cellular evaginations known as *acini* (Fig. 29.1). The cellular epithelium of the acini is functionally distinct from that of the collecting ducts. Saliva is initially secreted into the lumen of the acini. The glandular cells lining the acini secrete water, electrolytes, enzymes, and mucus. As the newly formed saliva progresses through the collecting ducts, its composition is modified. The duct epithelium reabsorbs electrolytes, especially sodium and chloride, in a manner similar to that in the proximal tubules of the kidneys. The final product, saliva, is hypotonic and has a sodium concentration substantially less than that of extracellular fluid. The extent to which the acinar secretion is modified in the collecting ducts depends on the rate of saliva production. At high rates of salivary flow, there is little modification, which results in higher tonicity and electrolyte concentration, compared with low rates of flow.

Most mammals have at least three pairs of salivary glands: the *parotid* glands, which lie just under the ear and behind the vertical ramus of the mandible; the *mandibular* glands, which are in the intramandibular space; and the *lingual* glands, which lie in the base of the tongue. Each of these glands drains into a main duct that has a single opening into the mouth. In addition to these major glands, there are minor glands in the tongue and buccal mucosa. These small, indistinct glands often have numerous secretory ducts emptying into the mouth. The concentration of mucus is different in the secretions of the various salivary glands. The parotid gland secretes watery, or *serous,* saliva, whereas many of the minor glands secrete highly mucous saliva. Other glands secrete a mixed type of saliva containing both mucous and serous material. Avian salivary glands secrete a copious amount of mucus to lubricate unmasticated food for swallowing.

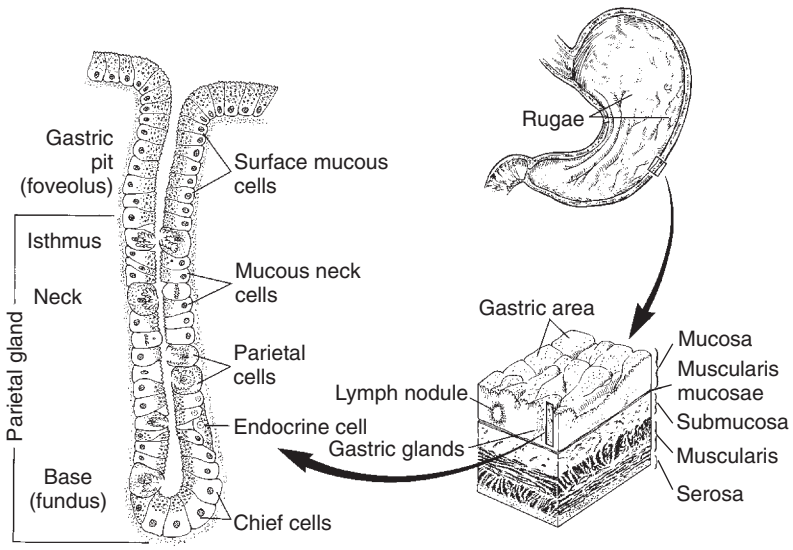
Salivary Glands Are Regulated by the Parasympathetic Nervous System

Autonomic, parasympathetic nerve fibers of the facial and glossopharyngeal nerves end on the secretory cells of the salivary gland acini and stimulate the cells through cholinergic receptors. All phases of salivary activity are stimulated by this mechanism, including electrolyte, water, and enzyme secretion. The anticipation of eating can initiate a parasympathetic response that results in salivary secretion. In Pavlov's famous experiment, parasympathetic stimulation of the salivary gland was evoked in dogs by the sound of a ringing bell. The dogs had been trained to anticipate eating after hearing the bell. This well-known experiment was one of the first demonstrations that the central nervous system (CNS) could regulate digestive functions. Chewing and stimulation of taste buds, in addition to the anticipation of eating, are afferent stimuli for salivation.

Salivary secretory cells also contain β-adrenergic receptors activated by sympathetic nerve stimulation or circulating catecholamines. This form of stimulation probably has little association with normal digestive activity but is related to the salivation and drooling seen in carnivores preparing to attack. Among digestive glands, the salivary glands are unique because there is no endocrine regulatory component.

Ruminant Saliva Is a Bicarbonate-Phosphate Buffer Secreted in Large Quantities

The normal composition of ruminant parotid saliva is quite different from the saliva of monogastric animals. Bovine and canine saliva are compared in Fig. 29.2. Ruminant saliva is isotonic and, compared with blood serum, has high concentrations of bicarbonate and phosphate and a high pH. This well-buffered solution is necessary for neutralizing acids formed by fermentation in the rumen, and ruminants secrete it in enormous quantities. An adult cow may secrete 100 to 200 L of saliva per day. This volume is approximately equivalent to the extracellular fluid volume of most adult cattle. It is obvious that much of the water and electrolytes secreted in saliva must be reabsorbed rapidly and recirculated through the total body water, or the animal would die of dehydration in a short period of time. In abnormal circumstances, such as blockage



• **Fig. 29.3** Anatomical illustration of glands of the stomach body. Other portions of the glandular stomach mucosa have similar structures but may differ somewhat in the cell types present in the glands. The gland openings are large enough to be seen with a handheld magnifying glass.

of the esophagus, in which the flow of saliva is diverted from the gastrointestinal (GI) tract, cattle quickly become dehydrated and acidotic.

In general, the salivary glands of domestic animals are seldom involved in disease processes and infrequently require veterinary attention.

Gastric Secretion

Depending on the Species, There May Be Two General Types of Gastric Mucosa: Glandular and Nonglandular

Most domestic, monogastric animals have only glandular mucosa in the stomach, but horses and rats have an area in the proximal portion of their stomachs covered by nonglandular, stratified squamous epithelium. This area is visibly different from the glandular area, to which it adjoins with a sharp line of demarcation. The function of the nonglandular area of gastric mucosa is unclear. The nonglandular area may serve as a place where a small amount of fermentative (rumen-like) digestion could occur. Because there is little mixing activity in the proximal stomach, food in the nonglandular area would be protected from the secretions of the gastric glands. These acid secretions kill bacteria, and thus their presence would prevent fermentation. Fermentative digestion is discussed in detail in Chapter 31.

The glandular area of the stomach is divided into three regions: *cardiac mucosa, parietal mucosa,* and *pyloric mucosa.* These areas contain glands of similar structure but with different types of secretions, as described later. In most species, the cardiac mucosa forms a narrow band around the gastric opening of the esophagus. In the pig, however, the cardiac mucosa covers a substantial portion of the proximal stomach.

The Gastric Mucosa Contains Many Different Cell Types

The glandular mucosa of the stomach has frequent invaginations, or pores, known as *gastric pits.* The size of the pits is such that the pores leading into them can be seen with a handheld magnifying glass. At the base of each pit is a narrowing, or isthmus, that continues into the opening of one or more gastric glands (Fig. 29.3).

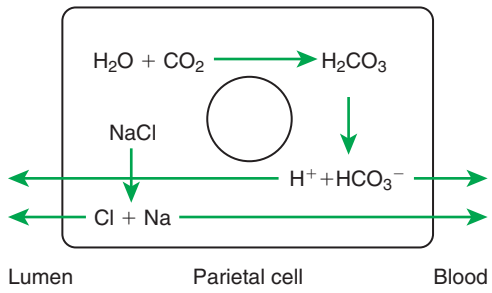
The major surface areas of the stomach and the lining of the pits are covered with *surface mucous cells.* These cells produce thick, tenacious mucus that is a special characteristic of the stomach lining. The mucous cells and their associated secretion are important for protecting the stomach epithelium from the acid conditions and grinding activity present in the lumen. When the mucous cells are injured, stomach ulcers result.

Each region of the mucosa contains glands with characteristic cell types. Within the parietal area, the glands contain *parietal* cells. These cells are clustered in the neck, or proximal area, of the gland. Their function is to secrete hydrochloric acid (HCl). Distributed among the parietal cells in the neck of the gland is another type of cell, the *mucous neck cells.* These mucous cells secrete thin mucus, less viscous than that of the surface mucous cells. The mucous neck cells, in addition to their secretory function, appear to be the progenitor cells for the gastric mucosa. They are the only cells of the stomach lining capable of division. As they divide, they migrate either down into the glands or up into the pits and onto the surface epithelium. As they migrate, the mucous neck cells differentiate into any of the several types of mature cells of the gastric surface and glands. In the base of the gastric glands is yet a third type of cell, the *chief cells.* These cells secrete *pepsinogen,* the precursor to the digestive enzyme *pepsin.*

The glands of the cardiac and pyloric mucosal regions resemble those of the parietal area in structure but contain different cell types. The cardiac glands secrete only mucus. Their mucus is alkaline and probably serves to protect the adjacent esophageal mucosa from the acid secretions of the stomach. The pyloric glands have no parietal cells but contain the gastrin-producing G cells. According to most reports, pyloric glands do secrete pepsinogen.

The Parietal Cells of the Gastric Glands Secrete Hydrochloric Acid

When the gastric glands are stimulated maximally, the HCl solution secreted into the lumen is isotonic and has a pH of less than 1. Both the hydrogen (H^+) and the chloride (Cl^-) ions are secreted by the parietal cells but apparently by different cellular



• **Fig. 29.4** Electrolyte movements during gastric acid secretion. The production of hydrogen and bicarbonate ions from water and carbon dioxide is stimulated by the action of the enzyme carbonic anhydrase, the activity of which is high in the gastric mucosa.

mechanisms. H^+ is secreted through an H^+,K^+-ATPase (adenosine triphosphatase) enzyme located on the luminal surface of the cell. This enzyme, sometimes referred to as a *proton pump,* exchanges H^+ for potassium ions (K^+), pumping one K^+ into the cell for each H^+ secreted into the lumen. In the exchange process, one molecule of adenosine triphosphate (ATP) is hydrolyzed to adenosine diphosphate (ADP), representing an expenditure of energy. The K^+ cations that accumulate within the cells are released back into the lumen in combination with Cl^- anions. This allows the recycling of K^+ ions as they are pumped back into the cells in exchange for H^+, resulting in the net secretion of H^+ and Cl^-, with little net movement of K^+.

Hydrogen ions for secretion come from the dissociation of intracellular carbonic acid (H_2CO_3), leaving a bicarbonate ion (HCO_3^-) in the cell for each H^+ secreted into the lumen (Fig. 29.4). Carbonic acid originates from water and carbon dioxide through the action of *carbonic anhydrase,* an enzyme found in high concentration in the gastric mucosa.

As hydrogen cations are secreted, bicarbonate anions accumulate in the cell. To counterbalance this accumulation, bicarbonate anions are exchanged for chloride anions at the cell's nonluminal surface. In this manner, additional chloride is made available to the cell for secretion into the glandular lumen, and bicarbonate is secreted into the blood. During periods of intense secretion by the gastric glands, large amounts of bicarbonate are released into the blood. This transient and mild alkalization of the blood during digestion is known as the *alkaline tide.* Normally, the alkaline tide is reversed when bicarbonate in the blood is consumed indirectly during the neutralization of gastric secretions as they enter the intestine (see the section on pancreatic secretions later in this chapter). Thus on a total-body basis, gastric acid production results in only small and transient changes in blood pH. In disease states, however, in which the secretions of the stomach are prevented from entering the intestine or are lost from the body because of vomiting, the pH of the blood can rise to dangerously high values.

Pepsin Is Secreted by the Chief Cells of the Gastric Glands; It Is Secreted in an Inactive Form and Is Subsequently Activated in the Stomach Lumen

Pepsin is usually described as a single compound, but it is actually a family of protein-digesting enzymes secreted from the gastric glands. They are formed in the chief cells as inactive proenzymes called *pepsinogens.* Pepsinogens are stored in the chief cells as granules until secreted into the lumen of the gastric glands. After secretion, pepsinogens are exposed to the acid contents of the stomach, resulting in cleavage of a small portion of the protein molecule, which leads to activation of the enzymes.

Digestive enzymes that are synthesized and stored as inactive proenzymes and activated in the lumen of the gut are known by the general name of *zymogens.* The general pattern of zymogen formation and activation is necessary because, if synthesized in the active state, the enzymes could digest and destroy the cells that synthesize them.

The Parietal Cells Are Stimulated to Secrete by the Action of Acetylcholine, Gastrin, and Histamine

Gastric acid secretion is stimulated by the anticipation of eating and the presence of undigested food in the stomach. When an animal anticipates eating, parasympathetic vagal impulses stimulate cells of the enteric nervous system (ENS), which in turn release acetylcholine (ACh) into the vicinity of G cells and parietal cells. These secretory cells have ACh receptors on their surfaces and respond by secreting gastrin and HCl, respectively. Gastrin circulates in the bloodstream and eventually finds its way to the parietal cells, which have gastrin receptors, in addition to ACh receptors, on their surfaces. The combined actions of gastrin and ACh on the parietal cells result in high rates of HCl flow. The response of the stomach to anticipatory stimuli originating in the brain is referred to as the *cephalic phase* of gastric secretion.

Food entering the stomach initiates the second phase, or *gastric phase,* of gastric secretion. Distention of the stomach by food stimulates stretch receptors, providing afferent stimulation of the ENS; the system responds by direct nervous (ACh) stimulation of the G and parietal cells. In addition, food acts as a buffer, raising the stomach pH. This removes the inhibiting effect of acid on G-cell secretion, further stimulating the production of gastrin, which leads to even greater enhancement of acid production by the parietal cells.

Histamine plays a role as an amplifying substance in gastric acid secretion. Parietal cells have surface receptors for gastrin, ACh, and histamine. They are stimulated maximally when all three receptors are occupied. Histamine is secreted by *mast cells* and *enterochromaffin-like cells* in the parietal mucosa. The histamine-secreting cells are stimulated to secrete by gastrin and ACh. Thus the effects of gastrin and ACh on gastric acid secretion are amplified through their stimulation of histamine secretion.

As gastric secretion and digestion proceed, the pH of the stomach decreases. When the stomach pH falls to about 2, gastrin secretion is suppressed, and at a pH of 1, gastrin secretion is completely abolished. Thus the gastrin stimulus to the parietal cells is removed, and acid secretion is reduced.

The intestinal environment also influences the secretion of gastric acid. As acid contents from the stomach flow into the duodenum and the duodenal pH is reduced, gastric acid production is suppressed. The exact mechanism by which duodenal acidification exerts negative feedback on the parietal cells is unclear. The hormone *secretin,* produced in the duodenum, and neuronal reflexes acting through the ENS may be involved.

The secretion of pepsinogen appears to be under the same regulatory influences as the secretion of HCl. However, regulation of pepsin secretion has been researched much less than regulation of HCl secretion.

The Pancreas

Pancreatic Exocrine Secretions Are Indispensable for Digestion of Complex Nutrients: Proteins, Starches, and Triglycerides

The pancreas is composed of two functionally separate types of glandular tissue. A small but important portion of the pancreatic tissue is arranged into discrete islets within the parenchyma of the gland. These cells are collectively called the *endocrine pancreas* because they secrete hormones into the bloodstream; the endocrine pancreas is discussed in Chapter 34. The great majority of the pancreatic tissue is involved with the elaboration of digestive secretions. This portion is known as the *exocrine pancreas* because its secretions are delivered into the intestinal lumen; the exocrine pancreas is the subject of this section.

Acinar Cells Secrete Enzymes, Whereas Centroacinar Cells and Duct Cells Secrete an Electrolyte Solution Rich in Sodium Bicarbonate

The exocrine pancreas is a typical acinar gland in which the end-pieces, or acini, are connected by an arborizing system of ducts; thus the gland conceptually resembles a bunch of grapes. Its general structure resembles the salivary gland, as illustrated in Fig. 29.1. The cells of the acini contain a generous portion of rough endoplasmic reticulum, on which large amounts of secretory proteins, the digestive enzymes, are synthesized. Each pancreatic acinar cell can produce all of the more than 10 different enzymes secreted by the pancreas. Chapter 30 discusses the functions of the major digestive enzymes of the pancreas (see Table 30.1). Protein-digesting enzymes, which are potentially harmful to the pancreatic cells, are synthesized as zymogens in the same manner as pepsinogen synthesis in the gastric glands. After synthesis, the enzymes and proenzymes are stored in vesicles, or *zymogen granules,* near the cellular apex. When the cells are stimulated, the zymogen granules fuse with the plasma membrane and release their contents into the lumen of the gland and eventually into the duodenal lumen, where they are converted to the activated form of the enzyme.

Specialized cells near the junction of the acini and ducts are called *centroacinar cells.* The function of these cells, and to a lesser degree the duct epithelial cells, is to modify the electrolyte composition of the fluid secreted by the acinar cells. The electrolyte makeup of the acinar secretion initially resembles extracellular fluid, having a relatively high concentration of sodium and chloride. The centroacinar cells have on their luminal surface a chloride-bicarbonate exchange protein that transports bicarbonate out of the cell in exchange for chloride, thus greatly enriching the bicarbonate concentration of pancreatic fluid. This exchange protein does not require an energy input, and its action is driven by a high intracellular concentration of bicarbonate. This system is facilitated by electrolyte transport proteins on the *basolateral surface* of the cell (see Chapter 30). These transport proteins consist of Na^+,K^+-ATPase, an Na^+-HCO_3^- cotransporter, an H^+-Na^+ exchanger, and an H^+-ATPase. The Na^+-HCO_3^- cotransporter, in combination with carbonic anhydrase, generates bicarbonate within the cell, thus driving the chloride-bicarbonate exchange at the luminal membrane. The H^+ remaining from the dissociation of carbonic acid is removed from the cell at the basolateral membrane by Na^+-H^+ exchange and by the H^+-ATPase pump. The net result is that pancreatic fluid is a bicarbonate-rich alkaline fluid that neutralizes the acid ingesta arriving in the duodenum from the stomach. In addition, the H^+ ions transported into the basal interstitial fluid of the pancreas are absorbed into the blood, balancing the "alkaline tide" that was created by gastric acid secretion.

In broad generality, the ion transport activities of the pancreatic duct cells are similar, but directionally opposite, to those of the parietal cells, as illustrated in Fig. 29.4. The overall effect of the two secretory cell types is to mix hydrochloric acid with ingesta in the stomach and to neutralize the acid with sodium bicarbonate in the duodenum.

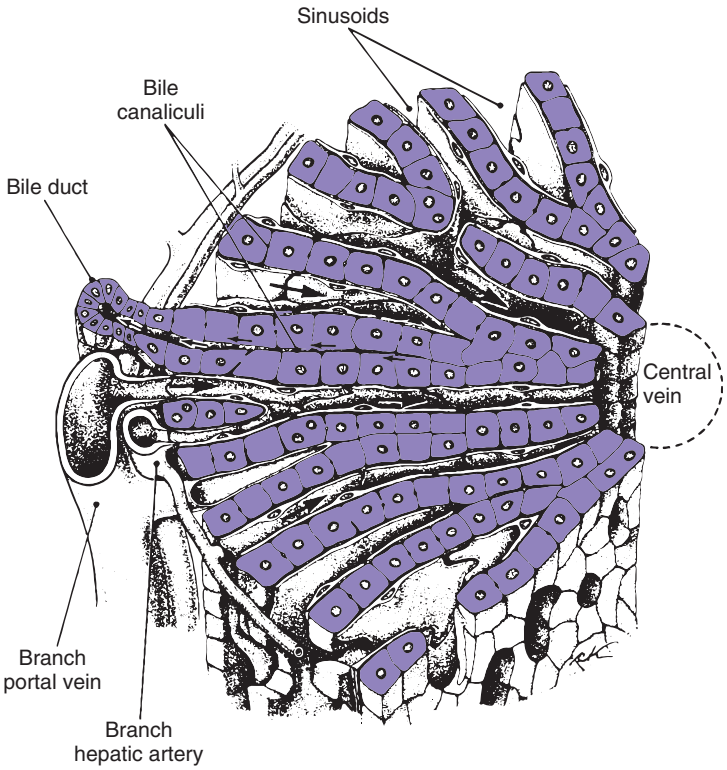
The ducts of the pancreatic lobules coalesce in an arborizing pattern to form either one or two main pancreatic ducts, depending on the species. The pancreatic duct or ducts may empty directly into the duodenum or, as in sheep, into the common bile duct. In the latter case, the pancreatic secretions enter the intestinal lumen along with bile.

Pancreatic Cells Have Cell Surface Receptors Stimulated by Acetylcholine, Cholecystokinin, and Secretin

When binding sites on the surfaces of pancreatic acinar, centroacinar, or duct cells are occupied, the cells are stimulated to secrete. Each type of cell appears to have receptors for ACh and for cholecystokinin (CCK) and secretin. ACh, released from nerve endings near the cells, stimulates secretion, as do CCK and secretin arriving in the blood. CCK is the primary hormonal stimulus for acinar cells, whereas secretin is the primary hormonal stimulus for centroacinar and duct cells. It appears, however, that maximal stimulation of the cells occurs when all receptors are occupied. Thus acinar cells secrete most actively in the presence of all three ligands: ACh, CCK, and secretin. In this manner, secretin is said to *potentiate,* or increase, the action of CCK on acinar cells, and CCK potentiates the action of secretin on centroacinar and duct cells.

Nerve fibers ending in the vicinity of pancreatic acinar glands originate from cell bodies in the ENS, traveling outside the gut wall and into the pancreas. These neurons are stimulated to release ACh by impulses arriving from other neurons of the ENS or by parasympathetic fibers arriving through the vagus nerve. Vagal stimulation of pancreatic secretion may arise as the result of several stimuli. The sight and smell of food induce centrally integrated vagal responses, leading to pancreatic secretion. This is known as the *cephalic phase* of pancreatic secretion, which is analogous in concept to the cephalic phase of salivary and gastric secretion. Distention of the stomach causes a vagovagal reflex stimulating pancreatic secretion, called the *gastric phase* of pancreatic secretion. The effects of the cephalic and gastric phases of pancreatic secretion are to "ready" the intestine for the imminent arrival of food by prior stimulation of pancreatic secretions.

The third phase, or *intestinal phase,* of pancreatic secretion is the most intense and involves endocrine and neuronal stimuli. This phase commences as food material from the stomach enters the duodenum. This leads to distention of the duodenum, which appears to produce enteric nerve impulses, resulting in ACh stimulation of pancreatic secretory cells. This stimulation reinforces and enhances the vagally mediated neuronal stimulation of the cephalic and gastric phases. The endocrine portion of the intestinal phase of pancreatic secretion occurs in response to the chemical stimulation that results from the presence of gastric contents in the duodenum. Peptides in the duodenal lumen, arising from the digestion of food protein, stimulate CCK production from endocrine cells in the duodenum. Fats in gastric ingesta also stimulate CCK



• **Fig. 29.5** Hepatic microanatomy is complex and can be visualized in several ways. Note the relationship of the bile canaliculi to the bile ducts; the biliary system may be viewed as an acinar gland with the bile canaliculi forming a long, narrow acinus. (Modified from Ham AW. *Textbook of Histology*. 5th ed. Philadelphia: Lippincott; 1965. In Fawcett DW. *Bloom and Fawcett: A Textbook of Histology*. 12th ed. New York: Chapman & Hall; 1994.)

secretion, whereas the low pH of material entering the duodenum from the stomach stimulates the secretion of secretin.

This stimulatory pattern is logical and results in a coordinated pattern of digestion. Proteins (peptides) and fats stimulate, through CCK, the secretion of protein-digesting and fat-digesting enzymes. These enzymes function best in an alkaline environment, and so the acid secretions of the stomach must be neutralized in order for these enzymes to be effective. Acid conditions in the duodenum stimulate pancreatic bicarbonate secretion through secretin, leading to alkalization of the ingesta. As food is digested and absorbed and acid is neutralized, the stimuli for pancreatic secretion are removed, and the amount of secretion diminishes to low, basal rates.

Bile Secretion

One function of the liver is being a secretory gland of the digestive system. Its secretion, bile, has an important role in fat digestion.

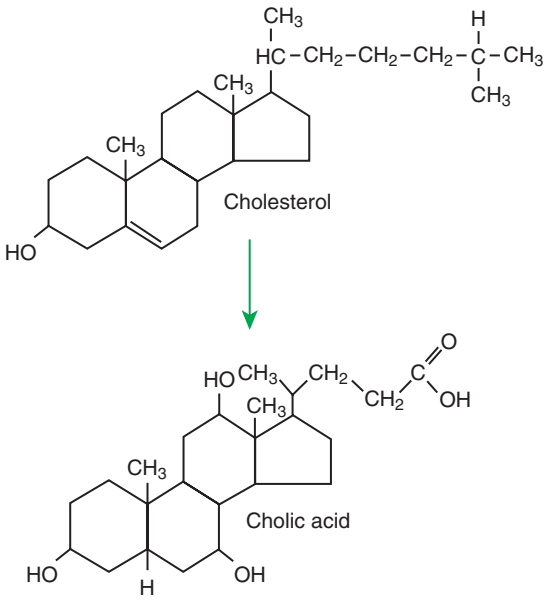
The Liver Is an Acinar Gland With Small Acinar Lumina Known as *Canaliculi*

The liver is composed of *plates,* or one-cell–thick layers of hepatocytes bathed on either side by blood from the hepatic sinusoids. Between each row of cells is a small space created by cavitations in the plasma membranes of two apposing cells. The portions of the plasma membranes lining the spaces are isolated from the remainder of the plasma membrane by tight junctions, which seal off the spaces from the surrounding extracellular environment. Within the plates of cells, these spaces join to form channels, or *canaliculi,* that connect to the *bile ductules.* Bile is secreted from the hepatocytes into the canaliculi, from which it flows into the bile duct system. From a functional standpoint, the canaliculi may be perceived as acini lined by hepatocytes and emptying into the biliary duct system, as illustrated in Fig. 29.5. The bile duct epithelium is metabolically active and capable of altering the composition of canalicular bile by adding additional water and electrolytes, especially bicarbonate. In this function, the bile duct epithelial cells function in a manner similar or identical to the centroacinar and duct cells of the pancreas. In fact, they even respond to secretin by increasing their bicarbonate secretion, just as do the pancreatic centroacinar cells.

Bile Contains Phospholipids and Cholesterol Maintained in Aqueous Solution by the Detergent Action of Bile Acids

Hepatocytes form bile acids from cholesterol. The chemical changes necessary to convert cholesterol to cholic acid, a representative bile acid, are shown in Fig. 29.6. Cholesterol is almost totally insoluble in water, but the chemical changes involved in the conversion of cholesterol to bile acids result in a molecule with a water-soluble (*hydrophilic,* or "water-loving") side and a lipid-soluble (*hydrophobic,* or "water-hating") side. This combination hydrophobic-hydrophilic attribute is the characteristic property of a detergent. Because of this dual solubility, detergents can render lipids soluble in water. The function of the bile acids is to emulsify dietary lipids and to solubilize the products of fat digestion.

Bile acids are produced in the smooth endoplasmic reticulum of the hepatocytes. As they are secreted from the cells into the lumen of the canaliculi, bile acids "dissolve away" some of the cell membrane components: phospholipids and cholesterol. These



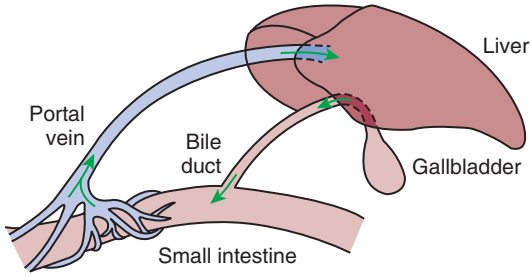
• **Fig. 29.6** Conversion of cholesterol to cholic acid, a representative bile acid. Note the presence of two additional hydroxyl groups on the ring structure of cholic acid compared with cholesterol. These hydroxyl groups enhance the water solubility and detergent action of the bile acid molecule. Other bile acids differ from cholic acid in the number and position of hydroxyl groups.

constituents—phospholipids, cholesterol, and bile acids—are the major functional components of bile and are important for the digestion and absorption of fats. The mechanism by which bile aids in fat digestion is discussed in Chapter 30.

Bile acids are secreted into the canaliculi as their sodium salts. The presence of bile acid salts and sodium in the canaliculi draws water, by osmosis, from the cells into bile. The electrolyte composition of canalicular bile usually resembles that of plasma but may be somewhat lower in chloride. As bile flows through the bile ducts, water and electrolytes are added. Bicarbonate may be secreted by the duct cells, so the bicarbonate concentration of bile is often higher than that in blood serum.

In addition to bile acids, phospholipids, and cholesterol, bile contains other lipid-soluble organic substances. Of these, the *bile pigments* are present in the highest concentration. Bile pigments are breakdown products of heme porphyrin, a portion of the hemoglobin molecule. The principal bile pigment is *bilirubin,* which is produced during the normal process of red blood cell turnover. Bilirubin gives bile its characteristic green color. In the lumen of the gut, bilirubin is converted by bacterial action to other compounds. These secondary compounds are responsible for the characteristic brown color of the feces of nonherbivorous animals. Bile pigments serve no useful digestive function: the body simply uses bile, and ultimately feces, as a route for the excretion of these waste products.

The liver serves as an excretory organ for many lipid-soluble substances in addition to bilirubin. The detergent action of the bile acids makes the liver an ideal excretory organ, in comparison with the kidney, for these types of compounds. Substances metabolized and secreted by the liver include many important drugs and toxins. This is important clinically because the actions of these agents can be potentiated by impaired liver function.



• **Fig. 29.7** Bile acids and other molecules circulate in an enterohepatic cycle. Phases of the cycle include the portal vein, biliary system, and intestinal lumen.

The Gallbladder Stores and Concentrates Bile During the Periods Between Feeding

When there is little or no food in the intestinal lumen, the *sphincter of Oddi,* at the union of the common bile duct and duodenum, is closed. With this sphincter closed, bile cannot enter the intestine and is diverted instead into the gallbladder. The gallbladder epithelium absorbs sodium, chloride, and bicarbonate from bile; water is absorbed passively. Thus in the gallbladder, the organic constituents of bile are concentrated, and the volume of bile is reduced. In species that have no gallbladder, such as horses and rats, the sphincter of Oddi is apparently nonfunctional, and bile is secreted into the intestine during all phases of the digestive cycle.

Bile Secretion Is Initiated by the Presence of Food in the Duodenum and Stimulated by the Return of Bile Acids to the Liver

When food, especially fat-containing food, reaches the duodenum, the GI endocrine cells are stimulated to secrete CCK, which in turn causes relaxation of the sphincter of Oddi and contraction of the gallbladder. These actions force stored bile into the intestine. Bile acids aid in the digestion and absorption of fats in the jejunum (see *Chapter 30*) but are not absorbed themselves until they reach the ileum. After absorption in the ileum, the bile acids travel via the hepatic portal vein to the liver. In the liver, bile acids are almost completely absorbed from the portal blood. As a result, almost no bile acids reach the posterior *venae cavae,* and they are consequently found only in low concentrations in the systemic circulation. The flow of bile acids from liver to intestine to portal blood to liver and back to intestine is known as *enterohepatic circulation* (Fig. 29.7).

Bile acids arriving at the liver, by way of the portal circulation, stimulate further bile synthesis. Thus a positive feedback system is initiated when the gallbladder contracts: the absorption of gallbladder bile from the intestine stimulates additional bile synthesis by the hepatocytes. Rapid bile synthesis and secretion continue as long as the sphincter of Oddi is open and the gallbladder is contracted. When fats have been digested and absorbed, the stimulus for CCK secretion is removed, resulting in closure of the sphincter of Oddi and diversion of bile to the gallbladder. Because bile acids are no longer reaching the intestine, they are no longer being absorbed, and thus the stimulus for bile secretion is reduced, and bile flow slows down.

In addition to the effect of CCK on bile secretion, secretin influences secretion from the bile duct epithelium. Secretin stimulates water and bicarbonate secretion from the bile ducts in a manner similar to that of its effects on the duct cells of the pancreas. Thus bile may participate in the neutralization of stomach acids.

CLINICAL CORRELATIONS

Horse in Pain With Weight Loss

History

A 4-year-old thoroughbred mare presents for weight loss, inappetence, grinding of teeth, and low-grade colic. She is off the track now but was winning until 1 month ago, and thus the trainer is concerned.

Clinical Examination

The mare seems to be quieter than expected. Her temperature, pulse, and respiration are normal. She appears thin for a horse from the track, and the trainer thinks the mare has lost 100 to 150 pounds in the past month. Her hair coat is poor. Your examination reveals no other abnormal findings.

With this history, gastric ulcers are a likely differential diagnosis. You discuss this with the trainer and decide to perform gastroscopy before doing any other diagnostic tests.

Comment

On endoscopy, the mare has several ulcers along the junction of the squamous and glandular sections of the gastric epithelium. In addition, she has two large and a few small ulcers on the nonglandular squamous compartment of the stomach. The squamous epithelium of the equine stomach has no mucus-secreting glands, in contrast to the glandular mucosa. The thick alkaline mucus that coats the glandular surface is an important component of the stomach epithelium's natural defense against acid damage. The lack of surface mucus makes the squamous portion of the equine stomach particularly prone to ulcer development. Ulcers are likely the cause of this horse's weight loss, colic, and poor performance. She will be treated, and her management will be modified to encourage healing of the ulcers.

Horses continuously secrete hydrochloric acid (HCl) in the stomach, in contrast to many other species that can modulate acid secretion based on food intake. Horses are thus suited to be constant grazers with access to food 24 h/day. This mare is kept in the stall 24 h/day unless she is being worked, and she is given a high-grain low-hay diet. Thus she eats two high-grain meals per day, and she only has a small amount of grass hay, which she typically eats quickly. When horses do not eat, the pH of the stomach decreases rapidly. Furthermore, being inside all the time adds to her stress, making this mare more prone to gastric ulcers. Histamine and gastrin also stimulate HCl secretion, whereas somatostatin inhibits it. Therefore treatment is multifactorial, aimed at changing the management to increase the pH of the stomach and administering drugs to help decrease acid secretion.

The parietal cells secrete HCl through H^+,K^+-ATPase (proton pump). Omeprazole, a once-daily medication, inhibits the proton pump. Other antiulcer medications include histamine type 2 receptor antagonists, such as cimetidine and ranitidine, which block histamine attachment to stimulatory receptors on the surface of parietal cells, decreasing HCl release. Another way to protect the stomach from acid damage is to coat the gastric lining with medications such as sucralfate, which forms a protective barrier between the mucosa and luminal contents.

Treatment

A common treatment choice in this case currently is omeprazole, which specifically decreases HCl secretion. Treatment depends on the severity of the ulcers and the inciting cause. In many cases, treatment may be recommended for up to 28 days. Additional management changes to enhance healing would include increasing the time per day the mare spends eating. Pasture would be ideal, particularly alfalfa forage, which has inherent buffering capacity. Gradually decreasing the amount of grain fed daily and increasing the amount of forage would also be beneficial. Although these management changes are ideal, they can be difficult to maintain because of the typical management of these horses.

Pancreatitis in a Dog

History

A 10-year-old female spayed overweight beagle escaped from the backyard. The owners found her in the neighbor's trashcan, eating discarded food. This was 2 days after a festive holiday, during which she had been given additional treats, including more fat than normal. About 12 hours after the dog had gotten into the garbage, she was depressed and not eating as well. Twenty-four hours after eating the trash, she was vomiting and had diarrhea and a fever.

Clinical Examination

The clients take the dog to their veterinarian. The veterinarian determines that the dog has a fever of 103.9°F, and she has a painful abdomen on palpation, particularly on the right side. She is still vomiting and has continued to have diarrhea. She is dehydrated. Blood work demonstrates an increased white cell count, demonstrating inflammation. Additionally, her amylase, lipase, and serum pancreatic lipase immunoreactivity (cPLI) are markedly increased. Abdominal radiographs demonstrate decreased contrast and a ground glass appearance in the region of the pancreas. Ultrasound of the pancreas demonstrates enlargement and some localized peritoneal fluid.

Comment

This dog has pancreatitis, which was likely triggered by the increased ingestion of fatty food. With a normal functioning pancreas, many of the enzymes are made in the inactive form (zymogens) and stored in intracellular zymogen granules. They are activated by cleavage of the aminoterminal polypeptide chain. Zymogens are normally not activated until they reach the small intestine. Cells in the duodenum have enteropeptidase, which cleaves the peptide from trypsinogen to activate trypsin. The latter cleaves the activation peptide from other digestive zymogens, thus providing considerable control over enzyme activity. Lysosomes are intracellular organelles that represent the cell's digestive system and contain enzymes capable of activating trypsinogen. Normally, the zymogens (in zymogen granules) and lysosomes are kept separately from each other. During pancreatitis, the zymogen granules and lysosomes can fuse, and their contents can mix together in intracellular vacuoles, leading to premature, increased, intracellular activation of zygomens. Such abnormal activation leads to local damage of pancreatic cells. Other regulation of pancreatic function is associated with inhibitors, such as pancreatic secretory trypsin inhibitor (PSTI). PSTI protectively inhibits trypsin if premature trypsinogen activation occurs within the acinar cell or duct system. However, if the pH is low, which occurs in the abnormal fusion vacuoles, this mechanism of regulation does not work. Additionally, if there is premature intracellular activation of trypsinogen, trypsin, and zymogens, this causes additional activation of other zymogens, thus leading to additional pancreatic damage. With pancreatitis, whether it is mild or more severe, many inflammatory mediators are also released, including TNF-alpha, IL-1, IL-2, IL-6, IL-8, IL-10, IFN-alpha, IFN-gamma, and platelet activating factor. Plasma protease inhibitors, including alpha macroglobulins and alpha proteinase inhibitors, can control inflammation. If these inhibitors are severely depleted, free proteases activate the coagulation, fibrinolytic, and complement cascades. This can lead to shock, disseminated intravascular coagulation (DIC), and death. The exact causes of pancreatitis are not known. It is more prevalent in obese dogs.

Treatment

Stabilization of patients with fluid/electrolyte therapy is recommended. Additionally, analgesics and antibiotics are often recommended. Both parenteral and enteral feeding, after vomiting ceases, have been recommended. Prognosis varies, depending on cause, severity, and chronicity of disease.

Practice Questions

1. In monogastric animals, saliva produced during periods of rapid secretion has a higher electrolyte concentration than saliva produced during periods of slow salivary secretion. From your understanding of salivary gland physiology, which appears to be the most likely explanation?
- **a.** During periods of slow salivary secretion, the acinar cells are inactive, and the duct cells secrete low-electrolyte saliva.
- **b.** Parasympathetic stimulation of the acinar cells results in the elaboration of a more electrolyte-rich saliva.
- **c.** Gastrin stimulation increases the electrolyte concentration of saliva.
- **d.** During rapid secretion, fluid produced by the acinar cells is exposed to the actions of the duct cells for a shorter time than during slow rates of secretion.
- **e.** Different cell types within the acinus are responsible for salivary production, depending on the type of stimulus.

2. Some nutritionists are experimenting with a drug that increases salivary secretion in cattle. What effect do you think this would have on rumen pH?
- **a.** Increase rumen pH
- **b.** Decrease rumen pH
- **c.** Have no effect on rumen pH

3. Inhibition of the enzyme carbonic anhydrase is likely to have what effect on gastric pH?
- **a.** Decrease gastric pH
- **b.** Increase gastric pH
- **c.** Have no effect on gastric pH

4. Which of the following is *not* a potential stimulus for gastric acid secretion?
- **a.** Norepinephrine secretion resulting from stimulation of sympathetic nerves.
- **b.** Vagal nerve activity resulting from the sight of food.
- **c.** The presence of undigested protein in the pyloric antrum.
- **d.** Acetylcholine release stimulated by gastric stretch receptors acting on nerves of the ENS.
- **e.** Histamine release from cells in the gastric mucosa.

5. Which of the following is *not* a natural ligand for receptors in the pancreas?
- **a.** Cholecystokinin
- **b.** Acetylcholine
- **c.** Gastrin
- **d.** Secretin

Bibliography

Entry-Level Supplements

Barrett K. *Gastrointestinal Physiology*. Columbus, Ohio: McGraw-Hill; 2006.

Barrett KE. *Gastrointestinal Physiology*. 2nd ed. McGraw-Hill; 2014:[chapters 3, 4, 11, 12].

Johnson LR. *Gastrointestinal Physiology*. 8th ed. Mosby; 2014:[chapters 7-10].

Advanced Reading

Chandra R, Liddle RA. Neural and hormonal regulation of pancreatic secretion. *Curr Opin Gastroenterol*. 2009;25(5):441–446.

Keating N, Keely SJ. Bile acids in regulation of intestinal physiology. *Curr Gastroenterol Rep*. 2009;11(5):375–382.

Kopic S, Geibel JP. Update on the mechanisms of gastric acid secretion. *Curr Gastroenterol Rep*. 2010;12(6):458–464.

Johnson LR, ed. *Physiology of the Gastrointestinal Tract*. 5th ed. London, UK: Elsevier; 2012:[chapters 45-53].

Monte MJ, Marin JJ, Antelo A, et al. Bile acids: chemistry, physiology, and pathophysiology. *World J Gastroenterol*. 2009;15(7):804–816.

Schubert ML. Gastric secretion. *Curr Opin Gastroenterol*. 2010;26(6): 598–603.

Stevens CE, Hume ID. *Comparative Physiology of the Vertebrate Digestive System*. 2nd ed. Cambridge, UK: Cambridge University Press; 1995.

Williams DA, Steiner JM. Canine exocrine pancreatic disease. In: Ettinger SJ, Feldman EC, eds. *Textbook of Veterinary Internal Medicine: Diseases of the Dog and Cat*. 6th ed. St Louis: Elsevier; 2005.

Williams JA. Regulation of acinar cell function in the pancreas. *Curr Opin Gastroenterol*. 2010;26(5):478–483.

30

Digestion and Absorption: The Nonfermentative Processes

THOMAS H. HERDT

KEY POINTS

1. Digestion and absorption are separate, but related, processes.
2. The small intestinal mucosa has a large surface area and epithelial cells with "leaky" junctions between them.
3. The intestinal surface microenvironment consists of glycocalyx, mucus, and an unstirred water layer.

Digestion

1. Breaking down food particle size by physical action is an important part of the digestive process.
2. Chemical digestion results in the reduction of complex nutrients into simpler molecules.
3. Luminal-phase carbohydrate digestion results in the production of short-chain polysaccharides.
4. Luminal-phase digestion of carbohydrates applies only to starches, because sugars are digested in the membranous phase.
5. Proteins are digested by a variety of luminal-phase enzymes.
6. Membranous-phase digestive enzymes are a structural part of the intestinal surface membrane.
7. Membranous-phase digestion occurs within the microenvironment of the unstirred water layer, intestinal mucus, and glycocalyx.
8. A specific membranous-phase enzyme exists for the digestion of each type of polysaccharide.
9. Complete digestion of peptides to free amino acids takes place both on the enterocyte surface and within the cells.

Intestinal Absorption

1. Specialized nutrient transport systems exist in the apical and basolateral membranes.
2. Secondary and tertiary active transport mechanisms use the transcellular sodium ion electrochemical gradient as their source of energy.
3. Passive transport occurs either through specialized channels in cell membranes or directly through the tight junctions.
4. The products of membranous-phase digestion are absorbed by sodium cotransport.

Absorption of Water and Electrolytes

1. There are at least three distinct mechanisms of sodium absorption.
2. There are three major mechanisms of chloride absorption.
3. Bicarbonate ion is secreted by several digestive glands and must be recovered from the gut if body acid-base balance is to be maintained.
4. Potassium is absorbed primarily by passive diffusion through the paracellular route.
5. The major mechanisms of electrolyte absorption are selectively distributed along the gut.
6. All intestinal water absorption is passive, occurring because of the absorption of osmotically active solutes.

Intestinal Secretion of Water and Electrolytes

1. Passive increases in luminal osmotic pressure occur during hydrolytic digestion and result in water secretion.
2. Active secretion of electrolytes from the crypt epithelium leads to intestinal water secretion.

Gastrointestinal Blood Flow

1. Water and solute movement between the lateral spaces and villous capillaries is subject to the same forces that govern water and solute movement between the extracellular and vascular fluids in other tissues.
2. Absorbed nutrients enter the capillaries by diffusion from the lateral spaces.
3. A countercurrent, osmotic-multiplier system may increase the osmolality of blood at the tips of the villi, further promoting absorption of water into the blood.
4. Disturbances in the venous drainage from the intestine can greatly affect the mechanisms of capillary absorption in the villi.

Digestion and Absorption of Fats

1. Detergent action and enzymatic action is necessary for the digestion and absorption of lipids.
2. Lipids are absorbed through the apical membrane by carrier proteins and simple diffusion.
3. Bile acids are reabsorbed from the ileum by a sodium cotransport system.
4. Absorbed lipids are packaged into chylomicrons before leaving the enterocytes.

Growth and Development of the Intestinal Epithelium

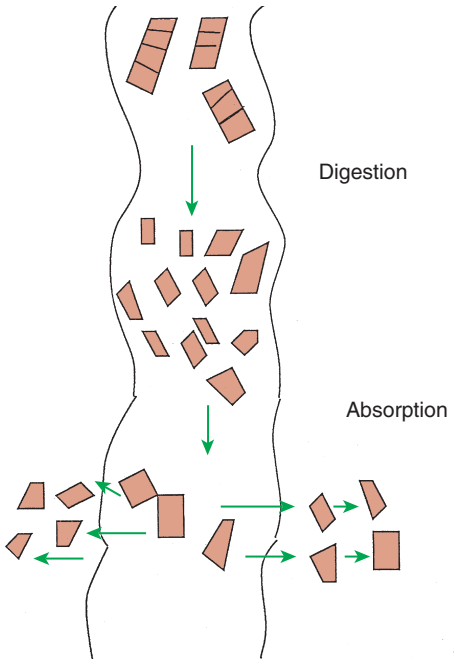
1. The length of intestinal villi is determined by the relative rates of cell loss at the tips and cell replenishment at the base.

Digestion in the Neonate

1. During the first few hours of life, proteins are not digested but are absorbed intact.
2. The major intestinal disaccharidase switches from lactase to maltase with maturity.

Pathophysiology of Diarrhea

1. Diarrhea occurs when there is a mismatch between secretion and absorption.



• **Fig. 30.1** Digestion is the process of reducing macromolecules to their constituent monomers. Absorption is the transport of the resultant monomers across the intestinal epithelium into the bloodstream.

Digestion and Absorption Are Separate, but Related, Processes

Digestion is the process of breaking down complex nutrients into simple molecules. In contrast, *absorption* is the process of transporting those simple molecules across the intestinal epithelium (Fig. 30.1). The two processes are the result of different biochemical events occurring within the gut. Both processes are necessary for the assimilation of nutrients into the body; absorption cannot occur if food is not digested, and the process of digestion is fruitless if the digested nutrients cannot be absorbed.

Disturbances of nutrient assimilation are common in veterinary medicine and may be caused by a variety of diseases, with some affecting digestion and others affecting absorption. The overt signs of failure of nutrient assimilation are often similar, but the biochemical lesions and specific therapies associated with *maldigestive* disease can be quite different from those associated with *malabsorptive* disease. Therefore diagnosing the cause of failure of nutrient assimilation is a frequent challenge faced by veterinary clinicians, a challenge that requires a thorough understanding of the physiology of nutrient digestion and absorption. This chapter first reviews the structural characteristics of the small intestinal epithelium that are of particular importance to the digestive and absorptive processes.

The Small Intestinal Mucosa Has a Large Surface Area and Epithelial Cells With "Leaky" Junctions Between Them

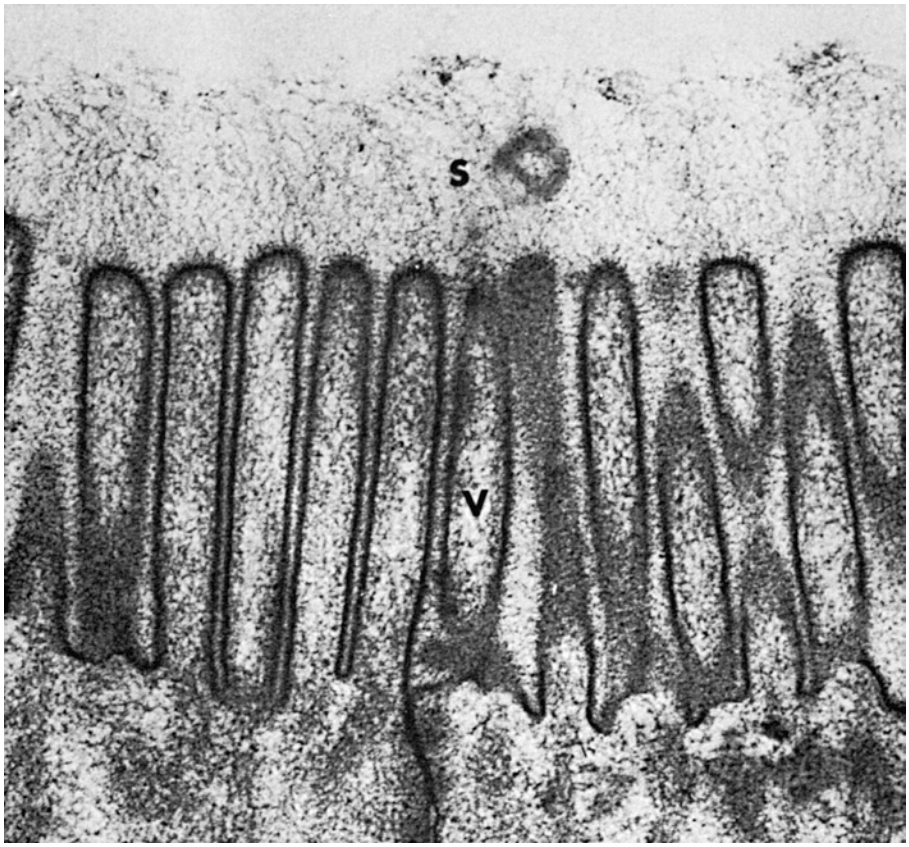
Contact between the small intestinal mucosa and the luminal contents is facilitated by an extensive intestinal surface area. Three levels of surface convolutions serve to expand the surface area of the small intestine (see Fig. 27.2). First, large folds of mucosa known as *plicae circulares* add to the intestinal surface area of some animals, but these are not present in all species. Second, the mucosal surface is covered with finger-like epithelial projections known as *villi.* These structures are present in all species and increase the intestinal surface area by 10-fold to 14-fold compared with a flat surface of equal size. Third, the villi themselves are covered with a brushlike surface membrane known as the *brush border.* The brush border is composed of submicroscopic *microvilli* that further enlarge the surface area (Fig. 30.2). At the base of the villi are glandlike structures known as *crypts of Lieberkühn* (Fig. 30.3). The villi and crypts are covered with a continuous layer of cellular epithelium.

The epithelial cells covering the villi and crypts are called *enterocytes.* Each enterocyte has two distinct types of cell membranes (Fig. 30.4). The cell surface facing the lumen is called the *apex* and is covered by the *apical membrane.* The apical membrane contains the microvilli. Under the light microscope, the microvilli give the cell surface its brushlike appearance. This appearance has led to the term *brush border,* which is synonymous with *apical membrane.* Attached to the apical membrane are many *glycoproteins.* These specialized proteins are synthesized within the enterocytes and transferred to the apical membrane. They are the enzymes and transport molecules responsible for the digestive and absorptive functions of the intestinal epithelium. Under the intense magnification of the electron microscope these proteins, which extend into the intestinal lumen, give the microvilli a fuzzy appearance (see Fig. 30.2). This rich area of glycoproteins on the surface of the apical membranes is given the name *glycocalyx.* The apical membrane is a complex cellular membrane with an unusually high protein content.

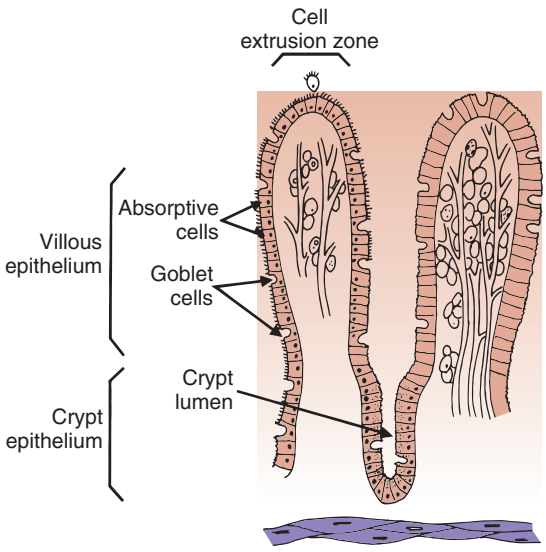
The remaining portion of the enterocyte plasma membrane, that part not facing the gut lumen, is called the *basolateral membrane,* referring to the base and sides of the cell. This membrane is not especially unusual and indeed has many similarities to cell membranes of other tissues. Although the basolateral membrane is not in direct contact with ingesta in the gut lumen, it serves an important role in intestinal absorption; nutrients absorbed into the enterocytes through the apical membrane must exit the cell through the basolateral membrane before gaining access to the bloodstream.

The attachments between adjacent enterocytes are called *tight junctions.* These connections serve a special function in the process of digestion and absorption. The tight junctions form a narrow band of attachment between adjacent enterocytes. The band is near the apical end of the cells and divides the apical membrane from the basolateral membrane. The junctions may be called "tight," but from a molecular standpoint, they are rather loose. This is especially true in the duodenum and jejunum, where the tight junctions are loose enough to allow the free passage of water and small electrolytes. Recent evidence indicates that the relative impermeability, or "tightness," of the tight junctions is not constant and can be altered by neurohumoral regulatory substances in the gut. These selective changes in permeability can affect the rates of water and ion movement across the gastrointestinal (GI) epithelium, depending on the physiological needs for secretion or absorption. However, the tight junctions are never permeable enough to permit the passage of organic molecules.

The narrow band of tight junctions leaves the majority of the basolateral membrane unattached to its neighboring membrane on the adjacent enterocyte. This arrangement creates a potential space between enterocytes. This area between the lateral surfaces of the enterocytes is called the *lateral space.* The lateral spaces are normally distended and filled with extracellular fluid (ECF). At the end of the lateral spaces nearest the apical membrane, the ECF



• **Fig. 30.2** Electron micrograph of the microvilli of the intestinal brush border. The brush border is composed of the apical membranes of enterocytes. Note the indistinct array of molecular material *(S)* that radiates away from the microvilli *(V);* membranous-phase digestion occurs within this array of molecular material, which includes the membrane-bound digestive enzymes. (From Johnson LR, Christensen J, Jacobsen ED, et al, editors. *Physiology of the Gastrointestinal Tract.* 2nd ed. New York, NY: Raven Press; 1987.)



• **Fig. 30.3** The one-cell-thick layer of intestinal epithelium is continuous over the villi and the crypts of Lieberkühn.

is separated from the fluid of the intestinal lumen only by the tight junctions. At the opposite end of the lateral spaces, the ECF is separated from the blood only by the basement membrane of the intestinal capillaries. Both the tight junctions and the capillary endothelium are permeable barriers that allow the free passage of water and small molecules. Thus there is relatively free flow of water and most electrolytes among the fluid in the lumen of the intestine, the ECF in the lateral spaces, and the blood.

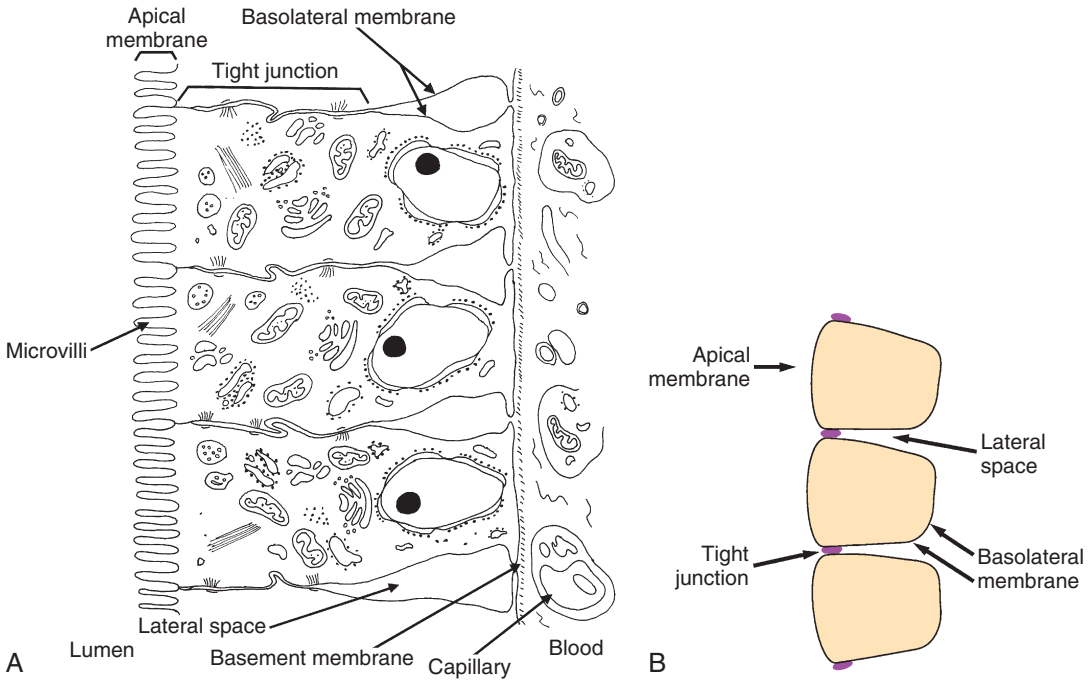
The Intestinal Surface Microenvironment Consists of Glycocalyx, Mucus, and an Unstirred Water Layer

Liberally interspersed among the enterocytes are *goblet cells,* which secrete a rich layer of mucus that covers the mucosa. At the brush border surface, the mucus blends into the glycocalyx, with the two layers forming a viscous coating that tends to trap molecules near the apical membrane. In addition to the mucus and glycocalyx layers, there is an area near the intestinal surface known as the *unstirred water layer.* With respect to the unstirred water layer, the intestine can be likened to a large stream or river, that is, the water in the center flows relatively rapidly, whereas the water near the edge or bank is quiet and flows slowly. Because of the same fluid-friction phenomenon that causes the water on the banks of rivers to be less turbulent and to flow at a slower rate than water in center stream, the water very near the intestinal surface is quiet and flows at a much slower rate than water in the central part of the lumen. The unstirred water layer, mucus, and glycocalyx form an important diffusion barrier through which nutrients must pass before entering the enterocytes.

Digestion

Breaking Down Food Particle Size by Physical Action Is an Important Part of the Digestive Process

The overall process of digestion is the physical and chemical breaking down of food particles and molecules into subunits suitable for absorption. Physical reduction of food particle size is important



• **Fig. 30.4** Understanding the anatomical relationships of the enterocytes, tight junctions, apical membrane, basolateral membrane, and lateral spaces is critical to an understanding of the physiology of intestinal absorption. (A) Anatomical illustration of the intestinal epithelium. (B) Stylized sketch of the epithelium. It is important to understand the relationship between part A and part B of this diagram.

not only because it allows food to flow through the relatively narrow digestive tube, but also because it enlarges the surface area of the food particles, thus increasing the area exposed to the actions of digestive enzymes. Physical reduction of food particle size begins with mastication (chewing) but is completed by the grinding action of the distal stomach. In the distal stomach, the physical action of grinding is aided by the chemical actions of pepsin and hydrochloric acid. The chemical actions of these stomach secretions break down connective tissue and thus aid in breaking apart food particles, especially foods of animal origin. The reduction of food particle size by physical means is essentially complete when food leaves the stomach, as described in Chapter 28 in the discussion of motility of the distal stomach.

Chemical Digestion Results in the Reduction of Complex Nutrients Into Simpler Molecules

Chemical digestion of each major nutrient is accomplished by the process of *hydrolysis,* the splitting of a chemical bond by the insertion of a water molecule. Glycosidic linkages in carbohydrates, peptide bonds in proteins, ester bonds in fats, and phosphodiester bonds in nucleic acids are all cleaved by hydrolysis during digestion. Fig. 30.5 illustrates hydrolytic splitting of these various chemical bonds.

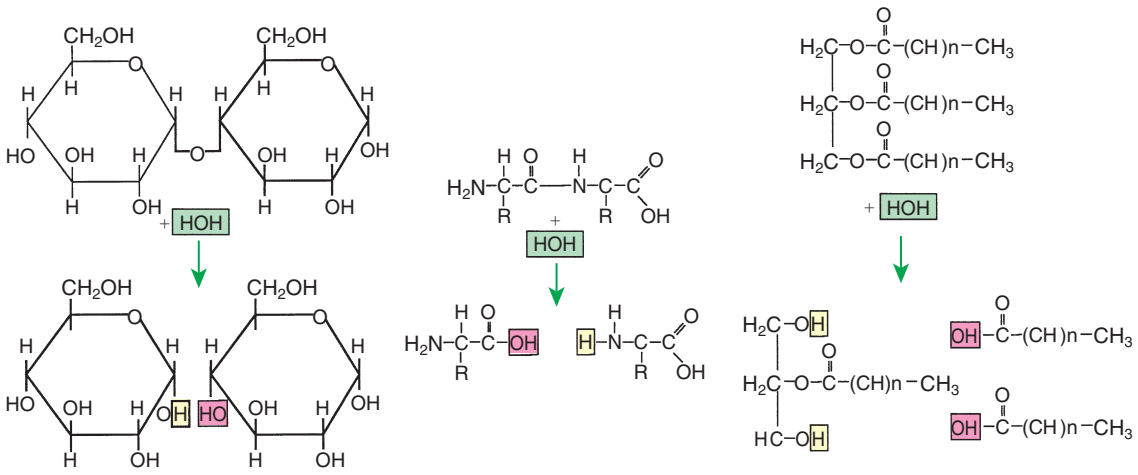
Hydrolysis in the digestive tract is catalyzed by the action of enzymes. There are two general classes of digestive enzymes: those that act within the lumen of the gut and those that act at the membrane surface of the epithelium. Enzymes acting within the lumen originate from the major GI glands, including the salivary gland, gastric glands, and especially the pancreas. The secretions of these glands become thoroughly mixed with ingesta and exert their actions throughout the lumen of their associated gut segments, thus the actions they catalyze are referred to as the *luminal phase* of digestion. In general, luminal-phase digestion results in the incomplete hydrolysis of nutrients, resulting in the formation of short-chain polymers from the original macromolecules (Fig. 30.6).

Enzymes that are chemically bound to the surface epithelium of the small intestine complete the hydrolytic process. These enzymes break the short-chain polymers resulting from luminal-phase digestion into monomers that can be absorbed across the epithelium. This final phase, which occurs at the epithelial membrane surface, is referred to as the *membranous phase* of digestion. Membranous-phase digestion is followed closely by absorption.

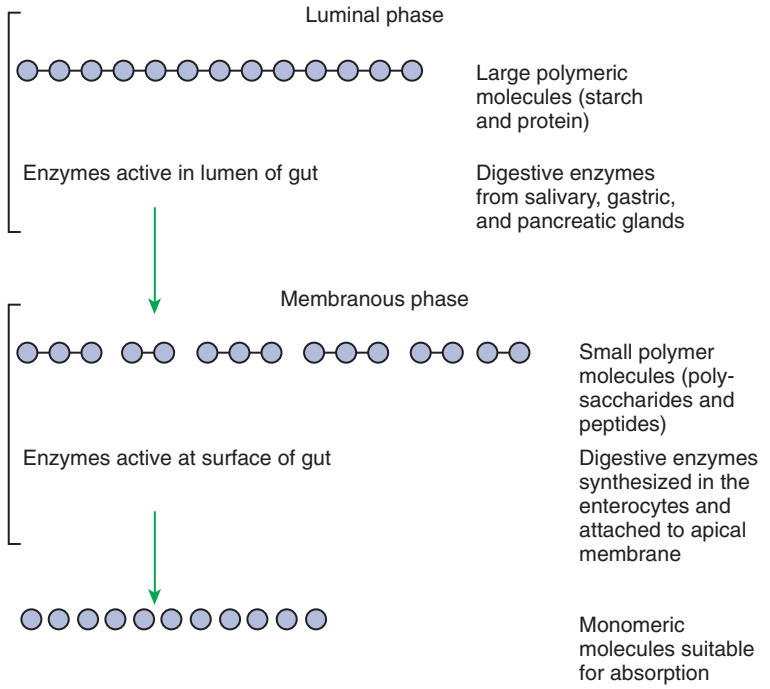
Luminal-Phase Carbohydrate Digestion Results in the Production of Short-Chain Polysaccharides

Carbohydrates are nutrients containing carbon, hydrogen, and oxygen atoms arranged as long chains of repeating simple-sugar molecules. Dietary carbohydrates originate primarily from plants. There are three general types of plant carbohydrates: fibers, sugars, and starches. *Fibers,* the structural parts of plants, form an important energy source for herbivorous animals; however, plant fibers are not subject to hydrolytic digestion by mammalian enzymes and therefore cannot be digested directly by animals (see Chapter 31).

Sugars are energy-transport molecules in plants. Sugars, or *saccharides,* may be *simple* (made up of a single molecular unit, monosaccharides) or *complex* (made up of two or more repeating saccharide subunits, polysaccharides). *Glucose, galactose,* and *fructose* are the most important simple sugars in animal diets. These monosaccharides are present, preformed in small quantities, in normal diets; however, most monosaccharides absorbed from the gut arise from the enzymatic hydrolysis of more complex carbohydrates. Complex sugars are referred to as *disaccharides, trisaccharides,* and *oligosaccharides,* depending on the number of repeating simple-sugar subunits. Oligosaccharides contain several monomer



• **Fig. 30.5** Major polymeric molecules forming food nutrients can be split into their constituent monomers by the insertion of a water molecule. This process, referred to as *hydrolysis,* is the major action of the digestive enzymes.



• **Fig. 30.6** Luminal-phase and membranous-phase oligosaccharides.

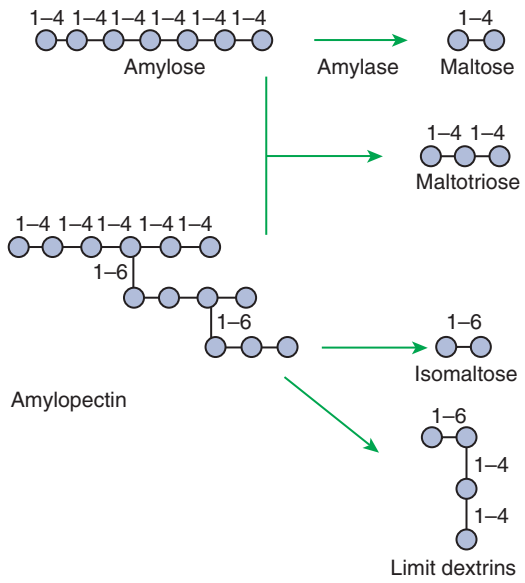
units, usually between 3 and 10. Important complex sugars in animal diets are *lactose,* or milk sugar, and *sucrose,* or table sugar. Lactose is a disaccharide composed of glucose and galactose, whereas sucrose is a disaccharide composed of glucose and fructose. Other important complex sugars are *maltose, isomaltose,* and *maltotriose;* these three sugars are composed of two or three repeating glucose units (Fig. 30.7). They are seldom present preformed in the diet but rather are formed in the gut as intermediate products of starch digestion.

Starch is an energy-storage carbohydrate of plants that forms the major energy-yielding nutrient in the diets of many omnivorous animals, such as pigs, rats, and primates. There are two chemical forms of starch, *amylose* and *amylopectin.* Both are long-chain glucose polymers, but amylose is a straight-chain molecule containing glucose monomers linked by α[1-4] glycosidic linkages. Amylopectin also contains glucose chains joined by α[1-4] glycosidic linkages, but the amylopectin chains are branched, having an α[1-6] linkage at each branch point (see Fig. 30.7). Although the chemical structure of starch is limited to these two molecular types, the physical structure and encapsulation of starches vary among plant sources. This variation results in the unique characteristics of starches from different sources, such as wheat, corn, and barley.

Luminal-Phase Digestion of Carbohydrates Applies Only to Starches, Because Sugars Are Digested in the Membranous Phase

The enzyme involved in luminal starch digestion is α-amylase, which is actually a mixture of several similar molecules. This enzyme arises from the pancreas in all species and from the salivary glands of some species (see Chapter 29). The α[1-4] linkages of either amylose or amylopectin are attacked by α-amylase. Characteristic of luminal-phase digestion, α-amylase does not break off, or *cleave,* single glucose units from the ends of the chain. Rather, the starch chains are broken in their midsections, resulting in the production of polysaccharides of intermediate chain length, known as *dextrins.* These chains continue to be attacked until disaccharide (maltose) and trisaccharide (maltotriose) units are formed.

This digestive process proceeds for amylopectin in the same way as it does for amylose, except that the α[1-6] linkages at the chain branch points of amylopectin are not hydrolyzed. Because these branch points are not hydrolyzed, branch-chain oligosaccharides, known as *limit dextrins,* and an α[1-6]-linked disaccharide, known as *isomaltose,* are formed (see Fig. 30.7). The end result of luminal-phase carbohydrate digestion is the creation of many disaccharides, trisaccharides, and oligosaccharides from large starch molecules. These complex sugars are not hydrolyzed further in the luminal phase.



• **Fig. 30.7** Two major forms of dietary starch are amylose and amylopectin. Amylose is composed of repeating glucose units joined by α[1–4] linkages. Amylopectin is a similar molecule, except that it has branch points formed by α[1–6] linkages. Because of the different linkage points, various polysaccharides result from luminal-phase digestion, as illustrated.

Proteins Are Digested by a Variety of Luminal-Phase Enzymes

Proteins are a source of amino acids, which are essential components of all animal diets. Dietary proteins come from both plant and animal sources. The general pattern of protein digestion is similar to that of carbohydrate digestion, in that large molecular proteins are broken down into small peptide chains by luminal digestion. Subsequent digestion of the peptide chains to individual amino acids occurs to a large extent by membranous-phase digestion, although unlike in carbohydrate digestion, a portion of free monomers, that is, amino acids, is released in the luminal phase.

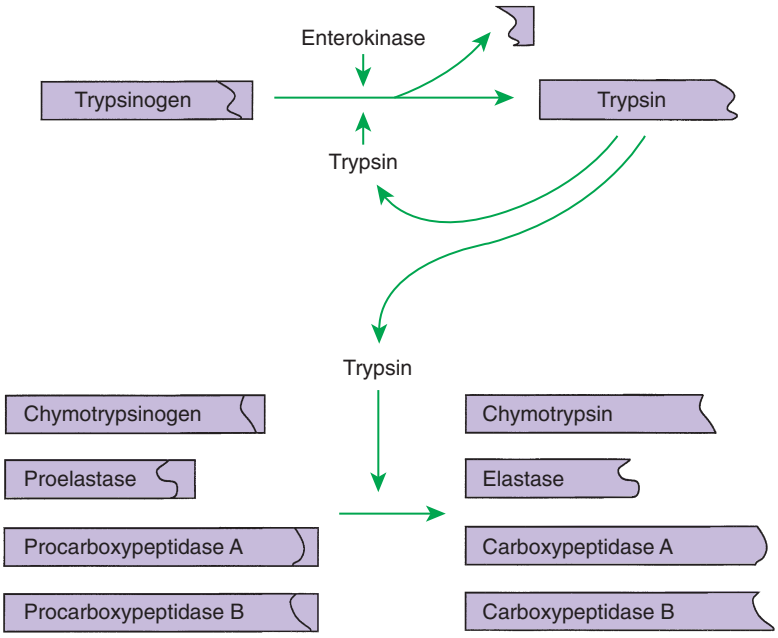
A major difference between protein digestion and carbohydrate digestion is in the number of different enzyme types involved. The relatively larger number of enzymes involved in protein digestion is expected, considering that starch molecules are made up of only one type of monomer—glucose—and protein molecules are made up of a variety of amino acids. Therefore only bonds between glucose molecules need to be broken in the case of starch. On the other hand, proteins are made up of infinite combinations of up to 20 individual types of amino acids; the various proteolytic enzymes are necessary for digestion because they differ in their efficiency in cleaving peptide bonds between specific types of amino acids.

The major luminal-phase proteolytic enzymes are listed in Table 30.1. Most proteolytic enzymes are *endopeptidases,* meaning that they break proteins at internal points along the amino acid chains, resulting in the production of short-chain peptides from complex proteins. Endopeptidases produce essentially no free amino acids. Two *exopeptidases,* which release individual amino acids from ends of peptide chains, are secreted also from the pancreas and are active in luminal-phase digestion.

The proteolytic enzymes are secreted from the stomach glands or pancreas in the form of inactive *zymogens* (see Chapter 29), which are activated in the stomach or intestinal lumen, respectively. These enzymes must be secreted in an inactive form; otherwise, the active enzymes would digest the cells in which they are synthesized. Activation of the zymogens occurs in the gut lumen. The proteolytic zymogens of the stomach, *pepsinogen* and *chymosinogen,*

TABLE 30.1 Luminal-Phase Enzymes of Protein Digestion

Enzyme	Action	Source	Precursor	Activator
Pepsin	Endopeptidase	Gastric glands	Pepsinogen	Hydrochloric acid, pepsin
Chymosin (rennin)	Endopeptidase	Gastric glands	Chymosinogen	Unknown
Trypsin	Endopeptidase	Pancreas	Trypsinogen	Enterokinase, trypsin
Chymotrypsin	Endopeptidase	Pancreas	Chymotrypsinogen	Trypsin
Elastase	Endopeptidase	Pancreas	Proelastase	Trypsin
Carboxypeptidase A	Exopeptidase	Pancreas	Procarboxypeptidase A	Trypsin
Carboxypeptidase B	Exopeptidase	Pancreas	Procarboxypeptidase B	Trypsin



• **Fig. 30.8** Activation of pancreatic zymogens. Note that trypsinogen is activated by trypsin and by the duodenal enzyme enterokinase. The autocatalytic action of trypsin on trypsinogen forms a positive feedback loop that ensures the rapid and complete activation of trypsinogen in the gut. Trypsin then activates the other zymogens.

are activated by hydrochloric acid (HCl) in the stomach lumen. Pepsinogen is also activated by pepsin in an autocatalytic feedback loop. *Trypsinogen* from the pancreas is activated by *enterokinase,* an enzyme elaborated by duodenal mucosal cells. The active enzyme, *trypsin,* then serves as an autocatalytic agent to activate additional trypsinogen and the other pancreatic protein-digesting enzymes. Fig. 30.8 illustrates the cascade of intraluminal zymogen activation.

Luminal-phase protein digestion begins in the stomach. Gastric digestion of protein is facilitated not only by the stomach enzymes but also by HCl, which has hydrolytic properties of its own. The acid environment of the stomach is suited to the action of pepsin, which has its optimal activity at pH 1 to 3. Gastric hydrolysis of protein is probably important to the physical and the chemical digestion of protein, because most connective tissue of animal origin is protein; digestion of connective tissue aids in breaking food down into particles small enough to pass the pylorus. Although stomach action is important in initiating protein digestion, it is not essential; animals without stomachs can digest proteins, provided they have a functional pancreas and are fed small, frequent meals of soft, moist food. Luminal-phase digestion of proteins is completed in the small intestine by the action of pancreatic enzymes.

Membranous-Phase Digestive Enzymes Are a Structural Part of the Intestinal Surface Membrane

Membranous-phase digestion, as with its luminal counterpart, occurs because of the hydrolytic action of enzymes. The difference between the two phases is that membranous-phase enzymes are chemically bound to the surface membrane of the intestine. They constitute a large and important portion of the glycocalyx. The substrates for these enzymes must diffuse into the glycocalyx before hydrolysis can occur. These membrane-bound digestive enzymes are synthesized within the enterocytes and subsequently transported to the luminal surface of the apical membrane. They remain attached to the surface by a short anchor segment, whereas the large, catalytic portion of the enzyme molecule projects away from the surface, toward the gut lumen.

Membranous-Phase Digestion Occurs Within the Microenvironment of the Unstirred Water Layer, Intestinal Mucus, and Glycocalyx

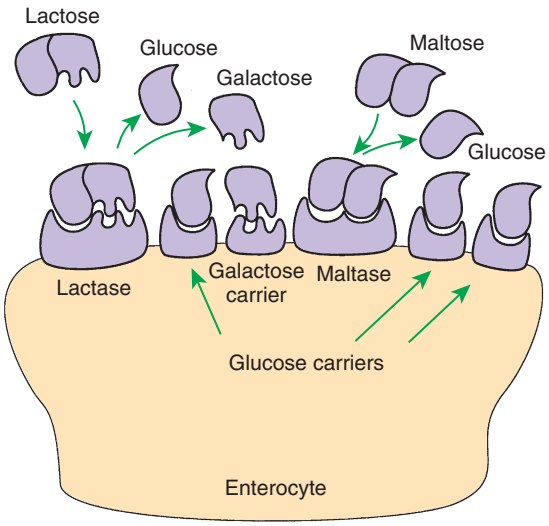
As previously described, the unstirred water layer, mucus, and glycocalyx form a diffuse zone separating the mucosal surface from the lumen of the intestine. The membranous-phase digestive enzymes project from the apical membrane into this surface layer. The quiet surface layer forms a microenvironment in which membranous-phase digestion occurs. Peptides and polysaccharides in the intestinal lumen must diffuse into the surface layer before membranous-phase digestion can take place. Furthermore, most of the products of membranous-phase digestion never diffuse away from the surface environment back into the lumen of the intestine; instead, they are absorbed soon after formation into the underlying epithelial cells. This arrangement is efficient because it ensures that the final products of carbohydrate and protein digestion are formed near their site of absorption, avoiding the need for long diffusion distances (Fig. 30.9).

A Specific Membranous-Phase Enzyme Exists for the Digestion of Each Type of Polysaccharide

The membranous-phase enzymes of carbohydrate digestion have as their substrates dietary complex carbohydrates, such as sucrose and lactose, and the polysaccharide products of luminal-phase starch digestion, including maltose and isomaltose. The specific membranous-phase enzymes are named according to their substrates and include *maltase, isomaltase, sucrase,* and *lactase.* The sole product of maltose and isomaltose digestion is glucose, whereas in addition to glucose, fructose and galactose are produced from the digestion of sucrose and lactose, respectively. All polysaccharides are digested to monosaccharides before absorption (Fig. 30.10).

Complete Digestion of Peptides to Free Amino Acids Takes Place Both on the Enterocyte Surface and Within the Cells

Membranous-phase digestion of peptides is, in some respects, similar to that of carbohydrates; peptide-digesting enzymes, or *peptidases,* are present on the enterocyte surface membrane and extend into the glycocalyx. These enzymes hydrolyze the peptide products of luminal-phase protein digestion, yielding free amino acids. Some of the longer-chain peptides are incompletely digested, yielding dipeptides and tripeptides. A large portion of dietary amino acids is absorbed directly in the form of dipeptides and tripeptides. This mode of absorption contrasts with that of carbohydrates, in which only monomeric, simple sugars may pass the apical membrane. Dipeptides and tripeptides that are absorbed intact are subsequently hydrolyzed by the action of intracellular peptidases, which results in the formation of free amino acids that are then available for passage into the blood. Thus the final digestion of peptides to free amino acids may occur at either of two sites: on the surface membrane of the enterocyte or within the cell. In either case, the final product of protein digestion is free amino acid (Fig. 30.11).



• **Fig. 30.9** Relationship of membranous-phase digestion to absorption. The enzymes responsible for digestion and the carrier molecules responsible for absorption are both part of the apical membrane. The products of digestion are thus formed in the immediate vicinity of the carrier proteins, avoiding long diffusion distances. Specific enzymes and carrier molecules are present for the various substrates, as illustrated.

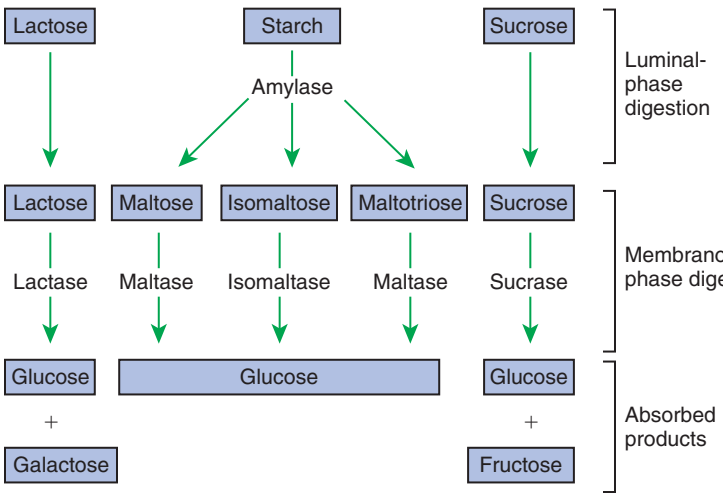
Intestinal Absorption

Absorption refers to the movement of the products of digestion across the intestinal mucosa and into the vascular system for distribution. To better understand the physiologically eloquent and clinically important processes of intestinal absorption, the reader might need to review the processes of diffusion across membranes, the difference in composition of intracellular and ECF (see Chapter 1), the electrical polarity across cell membranes, the function of the sodium-potassium (Na^+,K^+) adenosine triphosphatase (ATPase) pump, and the function of selective ion channels (see Chapters 1 and 4).

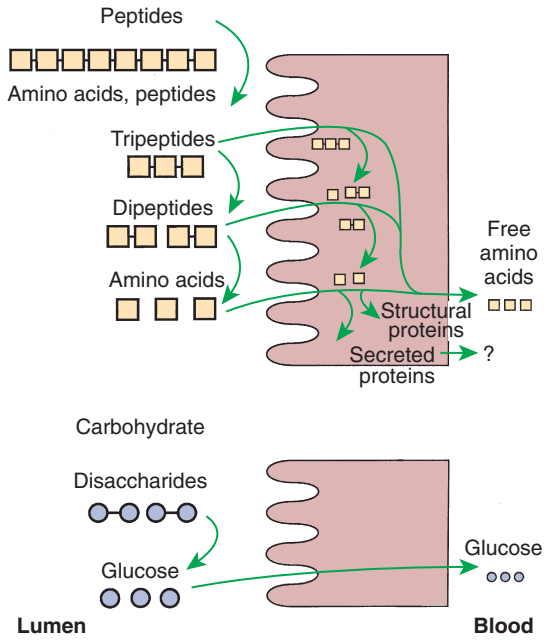
In considering intestinal absorption, keep in mind that molecules move across membrane barriers in response to chemical and electrical gradients. When molecules can freely penetrate a membrane, their movement across it is completely determined by the laws of diffusion and differences in chemical and electrical gradients: molecules flow to areas of lower concentration and charged particles move to areas of opposite charge. However, charged ions (especially cations) and most organic nutrient molecules do not freely penetrate the GI epithelium. Therefore they do not move in accordance with the laws of diffusion unless there is some mechanism to facilitate their transport across membranes.

Specialized Nutrient Transport Systems Exist in the Apical and Basolateral Membranes

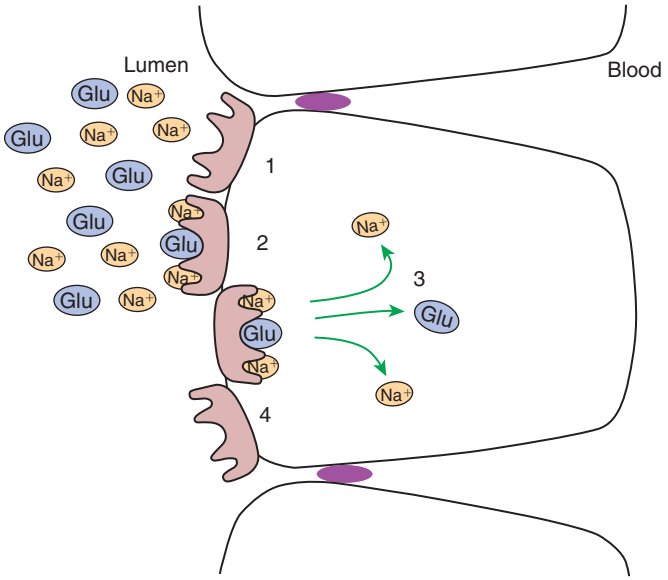
Specialized *transport mechanisms* exist for the movement of molecules across membranes in the intestinal epithelium. These mechanisms are interactions of events involving specific proteins that lie embedded in the matrix of the cell membranes of the epithelial cells. These proteins provide the *transport pathway* for the passage of ions and organic molecules across the plasma membranes of cells. As discussed here, there are many transport pathways. In general, the different pathways are polarized within the enterocytes, meaning that specific transport pathways exist on either the apical or the basolateral membrane, but not both. The transport pathway proteins chemically interact with specific organic nutrients and inorganic



• **Fig. 30.10** Luminal-phase and membranous-phase digestion of carbohydrate. Note that specific enzymes exist for each polysaccharide, and that a limited number of monomers are formed eventually from a relatively large number of starches and polysaccharides.



• **Fig. 30.11** Membranous-phase digestion of peptides and carbohydrates. Note that tripeptides and dipeptides may be hydrolyzed to their constituent amino acids either on the apical membrane or within the enterocyte. In carbohydrate digestion, however, all disaccharide hydrolysis occurs at the apical membrane. Regardless of the site at which the final hydrolysis of peptides occurs, the product absorbed into the blood is free amino acid (see Fig. 30.16).



• **Fig. 30.12** Cotransport is made possible by the allosteric transformation of transport proteins that lie in the apical membrane. The cotransport protein has two binding sites for sodium ions (Na^+) and one for glucose (*Glu*). When all three binding sites are occupied, the protein changes configuration in such a way as to transport the three ligands into the cell. The favorable gradient for sodium movement is maintained by the continuous action of the Na^+,K^+-ATPase pump (see Fig. 30.13).

ions to affect their transport across the membrane. The transport mechanisms can be classified as active transport, secondary active transport, tertiary active transport, and passive transport.

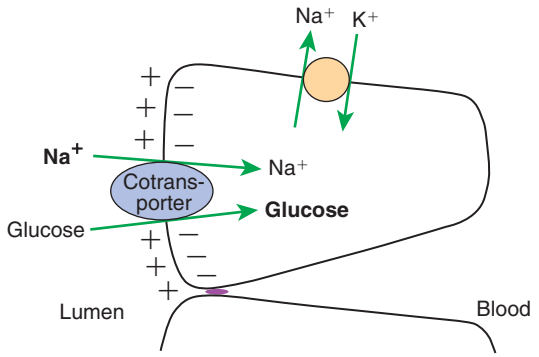
Active transport involves the direct consumption of metabolic energy. During active transport, energy stored as ATP is expended to move ions or molecules across membranes against an electrical or chemical gradient. In the large and small intestine, the active transport pathway of greatest importance is the Na^+,K^+-ATPase pump. This protein pathway lies on the basolateral membrane and uses energy from the hydrolysis of one molecule of ATP to drive three ions of sodium out of the cell, in exchange for the entry of two potassium ions into the cell. This important transport pathway exists in a wide variety of cells, in addition to enterocytes. The Na^+,K^+-ATPase pump is the mechanism by which (1) the interior of cells is kept electrically negative with respect to the ECF and (2) the concentration of sodium is kept very low in the intracellular fluid (see Chapter 1).

Secondary and Tertiary Active Transport Mechanisms Use the Transcellular Sodium Ion Electrochemical Gradient as Their Source of Energy

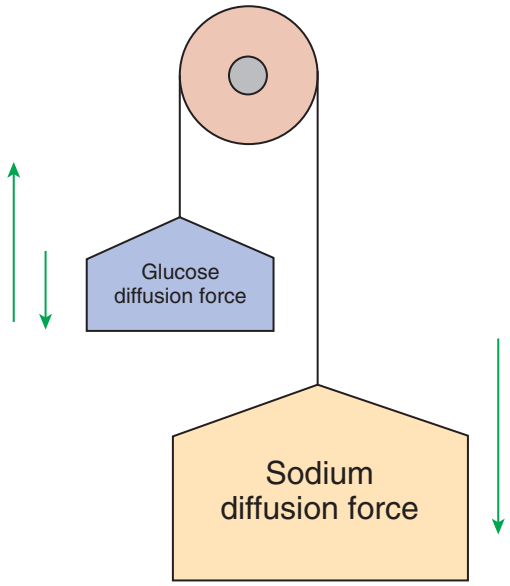
Just as a large stone resting atop a hill represents potential energy, so does the electrochemical gradient of sodium ions (Na^+) across the enterocyte membrane. Gravity imparts potential energy to the stone, whereas diffusion forces impart potential energy to Na^+ outside cells. Transport mechanisms that harness the potential energy of the sodium gradient are referred to as *secondary active transport.* Various transport pathway proteins exist for secondary active transport.

One type is referred to as a *cotransport* protein or *symport.* The characteristic of a cotransport protein is that it has binding sites for one or more Na^+ ions and an additional binding site for some other specific molecule. For example, the glucose cotransport protein has one binding site for glucose and two for Na^+. Cotransport proteins exist in the apical membrane of enterocytes. When the binding sites are unoccupied, they face the intestinal lumen. When all binding sites are occupied, a change in molecular configuration results in the translocation of binding sites, with their ligand molecules, to the interior of the cell. When this happens, the Na^+ ions along with the cotransported molecule are released into the intracellular fluid. Thus there is transport of sodium and another molecule, such as glucose, across the apical membrane. When the ligand molecules are released, the protein assumes its original configuration so that the binding sites are again on the extracellular surface of the apical membrane, ready to transport additional molecules (Fig. 30.12).

This process proceeds only as long as there is an electrochemical gradient for the Na^+. When this gradient is large, as is normally the case, it can provide the energy to "pull" the cotransported molecule, such as glucose, from an area of lower concentration to one of higher concentration, as illustrated in Figs. 30.13 and 30.14. Although the movement of a molecule against its concentration gradient represents expenditure of energy, there is no direct expenditure of metabolic energy by the sodium cotransport process. The energy expenditure is indirect and results from the direct expenditure of energy by the Na^+,K^+-ATPase pump in creating and maintaining the sodium electrochemical gradient. This is the definition of secondary active transport, with glucose transport being secondary to the active transport of sodium. Sodium cotransport processes absorb many organic nutrients, including glucose, amino acids, several vitamins, and bile acids.



• **Fig. 30.13** During cotransport, glucose is transported against an unfavorable concentration gradient. This diagram illustrates that the large sodium concentration difference across the apical membrane provides energy to transport glucose against its concentration gradient. The sodium concentration gradient, created by the action of the Na^+,K^+-ATPase pump, provides energy to drive this reaction.



• **Fig. 30.14** Secondary transport is an important concept. The tremendous sodium concentration difference between intracellular and extracellular fluid could be likened to the force of gravity; a pervasive force that affects many relationships in our environment. Movement of most ions, glucose, and many other organic molecules across the intestinal epithelium is driven by the strength of the sodium concentration difference.

In addition to sodium cotransport, there are other types of secondary active transport pathways. These pathway proteins are known as *exchangers* or *antiports.* Exchangers are usually involved with ion transport and are similar to cotransport proteins in that they have binding sites for selected ions. The difference between exchangers and cotransport proteins is that, for exchangers, the binding sites for the two different ligands are on opposite sides of the plasma membrane. For example, an important exchanger is the sodium/hydrogen (Na^+/H^+) exchanger in the apical membrane. The protein has a binding site for Na^+ and another for H^+. When the sites are unoccupied, the Na^+ site faces the intestinal lumen, and the H^+ site faces the interior of the enterocyte. When both sites are occupied, the protein flips, transporting H^+ out of and Na^+ into the cell, thus explaining the name *exchanger,* with H^+ exchanged for Na^+. As with cotransport, the force driving the exchange is the Na^+ electrochemical gradient across the cell membrane.

Another form of active transport, tertiary active transport, occurs via transport pathway proteins and is driven by electrochemical gradients established by secondary active transport. The best example of tertiary active transport is the chloride/bicarbonate (Cl^-/HCO_3^-) exchanger. This mechanism occurs in response to gradients established by the Na^+/H^+ exchanger, a secondary active transport mechanism. The Cl^-/HCO_3^- exchanger is discussed in more detail later, in the section on absorption. In essence, the term tertiary is used because the Na^+,K^+-ATPase system (primary) establishes the gradient that drives the Na^+/H^+ exchanger (secondary), which then establishes the gradient that drives the Cl^-/HCO_3^- exchanger (tertiary).

Passive Transport Occurs Either Through Specialized Channels in Cell Membranes or Directly Through the Tight Junctions

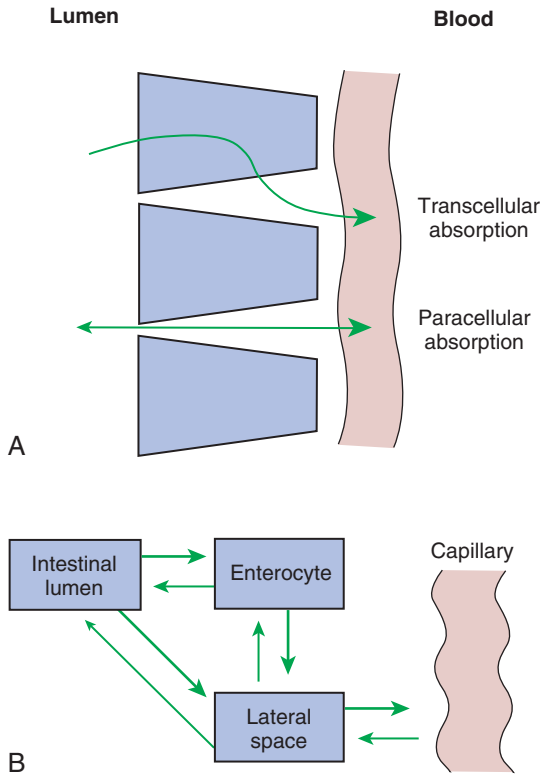
Ion channels, which are protein constituents of cell plasma membranes, are the transport pathways of passive diffusion into cells. Ions move through the channels in a completely passive manner, responding only to electrochemical gradients. No metabolic energy is directly required to effect ion movement. The only regulatory influence the cell can exert over this form of transport is in opening or closing of the channels (see Chapter 1).

A second form of passive molecular movement through the intestinal epithelium is through the tight junctions. As previously mentioned, the "tight" junctions are not so tight, especially in the duodenum and upper jejunum. In these areas, the tight junctions are freely permeable to water and small, inorganic ions. Thus water and ions move across the tight junctions in response to osmotic pressure and electrochemical gradients. Movement of materials through the tight junctions is called *paracellular* (around the cells) *absorption,* in contrast to absorption through the apical membrane, which is called *transcellular* (through the cells) *absorption.* Transcellular absorption and paracellular absorption work in a complementary manner to produce an efficient absorptive process (Fig. 30.15).

The Products of Membranous-Phase Digestion Are Absorbed by Sodium Cotransport

Sodium cotransport proteins for glucose and galactose are located in the apical membrane, in proximity to the membranous-phase digestive enzymes. Because these saccharide monomers are produced by the action of membranous-phase enzymes on polysaccharides, they move very short distances to binding sites on cotransport proteins. When both the glucose-binding (or galactose-binding) sites and the sodium-binding sites on these proteins are occupied, absorption occurs as previously described in the description of transport proteins.

In the initial phases of digestion of a starch-containing meal, the glucose concentration at the apical membrane is very high because there is ample substrate. Sodium is also readily available as a result of its presence in the various GI secretions. At this time, movement of both sodium and glucose into the enterocytes is down a concentration gradient. As digestion and absorption proceed, the glucose concentration at the apical membrane diminishes. Thus toward the end of the digestive and absorptive process, the concentration of glucose at the luminal surface of the enterocyte apical membrane becomes small. At this point, the concentration



• **Fig. 30.15** Transcellular and paracellular absorption. (A) Substances move from the intestinal lumen to the capillary by either transcellular (through the enterocyte) or paracellular (through the tight junction) absorption. (B) Intestinal lumen, enterocytes, and lateral spaces form three separate pools that may contain nutrients in different concentrations. Note that nutrients move into the capillaries from the lateral spaces and that reverse transport (from the capillary to the intestinal lumen) is possible for some substances.

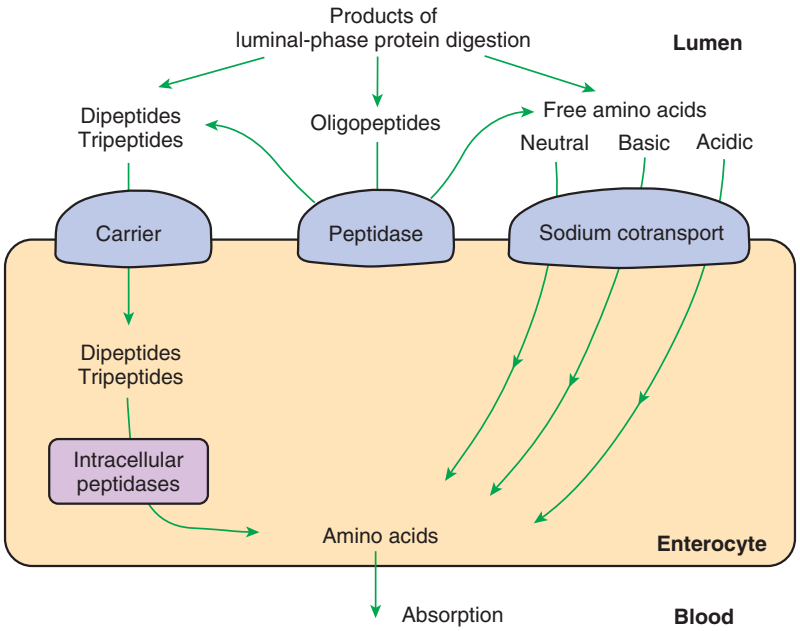
of glucose within the enterocyte can be higher than in the intestinal lumen, thus creating an unfavorable concentration gradient for glucose absorption. However, the transcellular sodium concentration gradient is maintained, driving the continued absorption of glucose (see Fig. 30.14). The process of glucose absorption by this mechanism is very efficient, and little free glucose escapes the absorptive process.

To complete the process of carbohydrate absorption, the glucose must move through the basolateral membrane into the lateral spaces, and then into the capillaries. Movement of glucose through the basolateral membrane occurs by *facilitated diffusion,* in which there is a transport pathway protein, but the direction of transport is driven only by the concentration gradient for glucose. As the intracellular glucose concentration in the enterocytes increases because of the action of sodium-glucose cotransport from the gut lumen, glucose diffuses from the cells into the lateral spaces. From the lateral spaces, it diffuses through the capillary basement membrane into the blood.

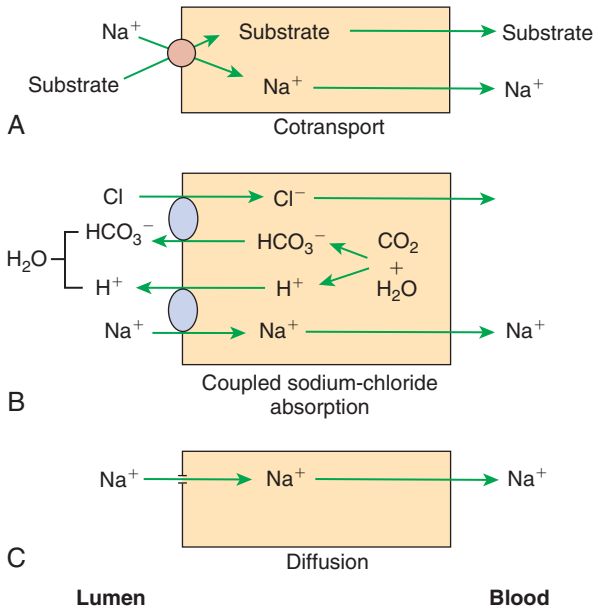
Absorption of the products of membranous-phase protein digestion occurs in a manner similar to that of carbohydrates. Sodium cotransport systems exist for free amino acids and might also exist for dipeptides and tripeptides. At least three cotransport proteins are necessary for absorption of free amino acids. The mechanism of transport for dipeptides and tripeptides might also involve sodium cotransport, but this issue is not established with certainty (Fig. 30.16).

Absorption of Water and Electrolytes

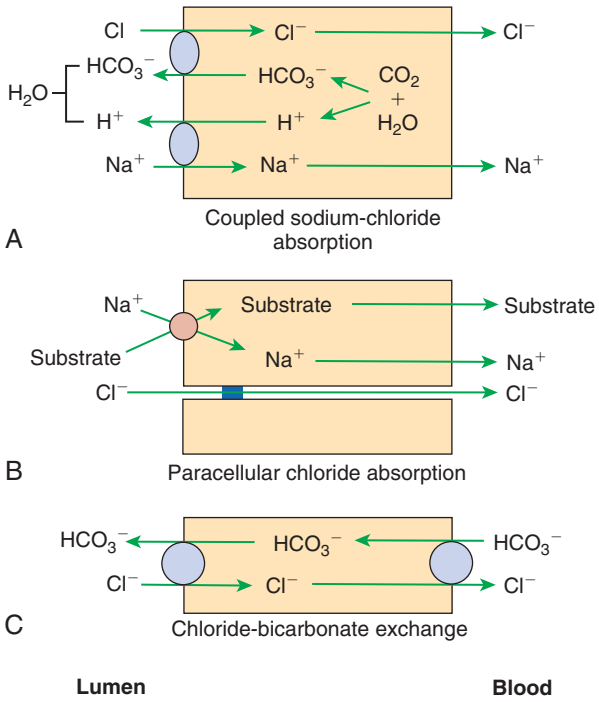
Conservation of the body's supply of water and electrolytes, primarily sodium, potassium, chloride, and bicarbonate, is a high priority for sustaining life. The gut plays a major role in this conservation, not only because it is the portal of entry for replenishment of the nutrients, but also because water and electrolytes in GI secretions must be efficiently reclaimed to maintain body composition. The most immediate clinical ramifications of GI disease usually involve



• **Fig. 30.16** At least three different sodium cotransport proteins exist for the transport of amino acids: those for neutral, basic, and acidic amino acids. A sodium cotransport process might be involved in the absorption of dipeptides and tripeptides, but this possibility is not well established.



• **Fig. 30.17** Three mechanisms of sodium (*Na⁺*) absorption. (A) Sodium cotransport with organic molecules is a major means of sodium uptake during active digestion and absorption. (B) Chloride-coupled sodium absorption is also an important means of sodium absorption and requires the action of carbonic anhydrase and the existence on the apical membrane of bicarbonate-chloride (HCO_3^-/Cl^-) and sodium-hydrogen (Na^+/H^+) exchange mechanisms. (C) Simple diffusion of sodium across the apical membrane may occur because of the large, favorable concentration gradient, but it is a relatively minor means of sodium absorption. CO_2, Carbon dioxide; H_2O, water.



• **Fig. 30.18** Three mechanisms of chloride (Cl^-) absorption. (A) Chloride-coupled sodium absorption is directly related to sodium (*Na⁺*) uptake. (B) Paracellular chloride absorption is indirectly related to sodium absorption that occurs during cotransport. (C) Chloride-bicarbonate (Cl^-/HCO_3^-) exchange occurs especially in areas where bicarbonate secretion into the intestinal lumen is important.

the loss of water and electrolytes. This section discusses the absorption of the major ions and electrolytes sequentially.

There Are at Least Three Distinct Mechanisms of Sodium Absorption

The first pathway of sodium absorption is through sodium cotransport proteins, as previously discussed. This secondary active transport pathway is not only the mechanism for glucose and amino acid absorption (Fig. 30.17A), but also a major means of sodium absorption.

The second sodium absorption mechanism is through the Na^+/H^+ exchanger (see Fig. 30.17B), mentioned previously as an example of an ion exchanger, or antiport. Through this mechanism, intracellular H^+ is exchanged for luminal Na^+ across the apical membrane. The H^+ for this exchange is formed by the action of carbonic anhydrase, which generates HCO_3^- and H^+. As H^+ is exchanged for Na^+, HCO_3^- concentrations build up in the cell. The resulting transcellular HCO_3^- gradient drives the action of the Cl^-/HCO_3^- exchanger, which results in the exchange of intracellular HCO_3^- for luminal Cl^-. Because of the close connection between Na^+ and Cl^- absorption by these pathways, this transport mechanism is often called *coupled sodium chloride transport,* as illustrated in Fig. 30.17B. One must appreciate, however, that it is only the intracellular balance of H^+ and HCO_3^- that couples the two exchange pathways. There are instances in which the intracellular pH is such that Na^+/H^+ exchange occurs without Cl^-/HCO_3^- exchange, and vice versa.

Coupled sodium chloride absorption is usually most active in the ileum and colon, where the sodium concentration in the gut is usually relatively low compared with that in the duodenum and jejunum. As usual, sodium entering the enterocytes is transported across the basolateral membrane to the lateral spaces by the action of the Na^+,K^+-ATPase pump. Chloride, however, remains in the enterocyte until its concentration is high enough to promote the diffusion of chloride through special channels, or gates, in the basolateral membranes. The rate of absorption of sodium and chloride by the coupled mechanism appears to depend on the permeability of the chloride channels; when the permeability is high, chloride passes rapidly out of the enterocyte, allowing continued chloride absorption. Conversely, when chloride channels are relatively closed, the intracellular chloride concentration rises, diminishing chloride absorption by the creation of an unfavorable concentration gradient across the apical membrane.

The third mechanism of sodium absorption is by simple diffusion through ion channels in the apical membrane (see Fig. 30.17C). The large electrochemical gradient that can exist for sodium across the enterocyte apical membrane allows direct, uncoupled movement of sodium across the membrane when the ion channels are open. Although some sodium absorption probably occurs by this mechanism, its overall importance in body sodium homeostasis is probably not great.

There Are Three Major Mechanisms of Chloride Absorption

One mechanism of chloride absorption is coupled sodium chloride absorption, as discussed previously in relation to sodium (Fig. 30.18A). Another mechanism is paracellular chloride absorption, which occurs in association with sodium cotransport of glucose

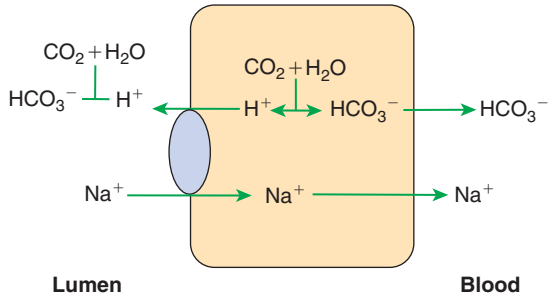
and amino acids (see Fig. 30.18B). Paracellular chloride transport occurs because of an electrical gradient. Sodium cotransport leads to the net movement of positive electrical charges (Na^+) across the apical membrane, because neither glucose nor most amino acids are charged molecules. As the sodium cations are transferred to the lateral spaces, the spaces develop a positive polarity with respect to the gut lumen. Chloride from the gut lumen passes directly into the lateral spaces through the tight junctions because these junctions are readily permeable to small anions. This provides a major mechanism for the absorption of Cl^- while maintaining electrical neutrality, although a small electrical potential is maintained across the gut surface, the lumen being negative with respect to the lateral spaces.

The last mechanism of chloride absorption is by direct exchange for bicarbonate (see Fig. 30.18C) without coupled sodium absorption. With this mechanism, there is a net movement of bicarbonate into the gut lumen, resulting in an increase in luminal pH. This can be particularly important in the colon of large herbivores where large concentrations of fermentation acids are created and require buffering.

Bicarbonate Ion Is Secreted by Several Digestive Glands and Must Be Recovered From the Gut if Body Acid-Base Balance Is to Be Maintained

Much bicarbonate is in essence "absorbed" by the neutralization of HCl from the stomach. Sodium bicarbonate entering the intestine reacts with HCl to form water, carbon dioxide, and sodium chloride, effectively resulting in the absorption of bicarbonate (HCO_3^-) and hydrogen (H^+) ions. (See Chapter 28 for an explanation of the counterbalancing effects of gastric acid secretion and pancreatic bicarbonate secretion.) However, considerable bicarbonate remains in the intestine after the neutralization of stomach acid. This remaining bicarbonate is reabsorbed, primarily in the ileum and colon via an ion-exchange mechanism.

Bicarbonate anions in the gut are electrically balanced, primarily with sodium cations, and reabsorbed essentially as sodium bicarbonate. In the absorptive process, H^+ and HCO_3^- ions are first generated within the enterocytes from water and carbon dioxide. H^+ is then exchanged for Na^+ across the apical membrane. Within the cell, Na^+ is electrically balanced by the remaining HCO_3^-, whereas the HCO_3^- remaining in the gut lumen is neutralized by the secreted H^+ (Fig. 30.19). The result is that sodium is transferred through the membrane. However, luminal bicarbonate is converted to water and carbon dioxide in the gut lumen, whereas bicarbonate anion is regenerated intracellularly. The net effect is the absorption of sodium bicarbonate.

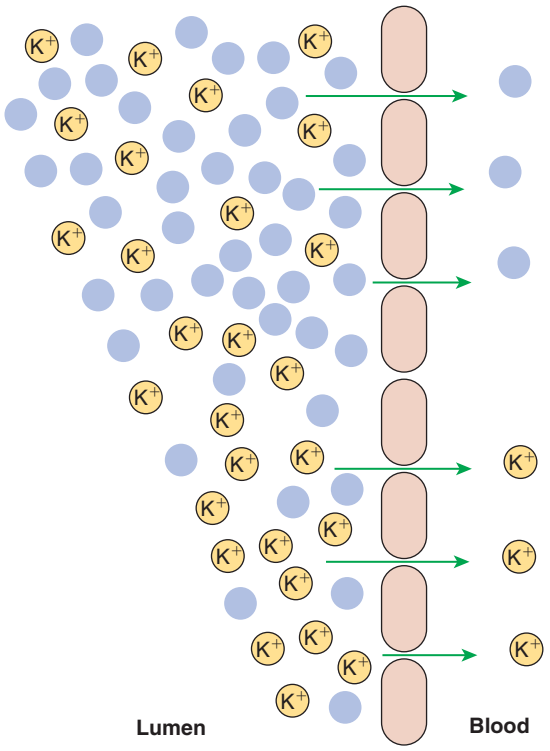


• **Fig. 30.19** Absorption of bicarbonate (HCO_3^-) is facilitated by sodium-hydrogen (Na^+/H^+) exchange at the apical membrane. The bicarbonate ion is regenerated by the action of carbonic anhydrase.

Potassium Is Absorbed Primarily by Passive Diffusion Through the Paracellular Route

Potassium (K^+), although a highly important ion in the body, is present in abundance in most animal diets. This is in contrast to sodium (Na^+), which is present in nutritionally inadequate amounts in most natural animal feeds. Therefore frequently the K^+ concentration in the material entering the intestinal lumen is relatively high compared with Na^+ concentration. In addition, dietary potassium is concentrated in the gut lumen because of the absorption of other nutrients, electrolytes, and water, unaccompanied by active potassium absorption. Thus K^+ concentration within the gut lumen increases while digestion and absorption of other osmotically active molecules proceeds.

As K^+ reaches relatively high concentrations in the intestinal lumen, a concentration gradient favorable for the diffusion of potassium across the intestinal epithelium is created. Furthermore, the concentration gradient is enhanced by the normally low K^+ concentration in the lateral spaces. The primary mechanism of potassium absorption is paracellular passive diffusion, which occurs in response to this concentration gradient (Fig. 30.20). A clinical ramification of this absorptive mechanism is that potassium absorption is directly coupled to water absorption. That is, the movement of water out of the intestinal lumen results in an increase in the luminal K^+ concentration, which in turn drives K^+ absorption. In diarrhea conditions, in which net absorption of water is impaired, K^+ absorption is impaired as well, because potassium in the gut lumen is diluted so that a concentration gradient favorable for passive diffusion of K^+ never develops. In addition to passive



• **Fig. 30.20** Potassium (K^+) is absorbed by simple diffusion through the paracellular route. Water absorption in the upper intestine increases K^+ concentration in the lower intestine, creating a favorable diffusion gradient for potassium. Note that the removal of water (*solid blue circles*) in the upper part results in a relative increase in the number of K^+ ions in the lower part.

TABLE 30.2 Distributions of Electrolyte Absorptive Mechanisms Throughout the Gut

		JEJUNUM				
Mechanism	Duodenum	Upper	Middle	Lower	Ileum	Colon
Sodium cotransport	+++++	++++	+	+	–	–
Chloride-coupled sodium absorption	+	+	+	+	++	+++
Chloride-bicarbonate exchange	–	–	–	–	++	+++
Bicarbonate absorption	–	–	–	–	++	+++
Potassium absorption	–	–	–	–	+	+++

diffusion, it appears that an H^+,K^+-ATPase pump exists in the distal colon. This transport pathway may be important for recovering the last remaining potassium from the colonic ingesta of animals with diets low in potassium.

The Major Mechanisms of Electrolyte Absorption Are Selectively Distributed Along the Gut

The activity of the assorted electrolyte absorption mechanisms discussed earlier varies along the length of the gut. The distribution of activity is listed in Table 30.2.

All Intestinal Water Absorption Is Passive, Occurring Because of the Absorption of Osmotically Active Solutes

Water moves through the intestinal mucosa by either the paracellular or the transcellular route, but always by osmosis. Those who do not have a clear understanding of the process should review the general discussion of osmosis in Chapter 1. The intestinal mucosa is freely permeable to water, allowing it to move in whatever direction is dictated by changes in osmotic pressure. As electrolytes and other soluble nutrients are actively absorbed, water is drawn along passively from lumen to intestinal capillaries. Water may move also into the intestinal lumen at times when the intraluminal osmotic pressure is high, as discussed later.

Intestinal Secretion of Water and Electrolytes

In addition to the water and electrolytes secreted into the intestine by the pancreas, liver, and other glandular organs, a considerable portion of GI water and electrolyte secretion occurs directly from the intestinal surface. All water secretion is osmotic, but the osmotic gradient driving water secretion may occur in response to either passive or active processes.

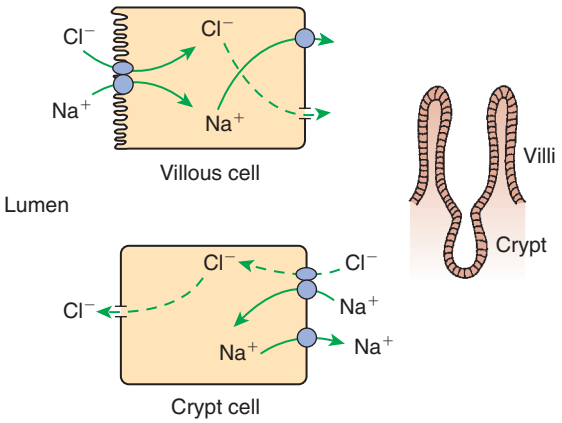
Passive Increases in Luminal Osmotic Pressure Occur During Hydrolytic Digestion and Result in Water Secretion

Food entering the intestine may be hyperosmotic because of its composition, such as salty foods and foods with high sugar content. Alternatively, food may become hyperosmotic after digestion. Digestion of foods creates many osmotically active molecules from one giant precursor molecule; thus the osmotic activity of ingesta is increased initially by digestion. When starchy meals, for example, first enter the duodenum, intraluminal digestion creates thousands of osmotically active disaccharide and trisaccharide molecules from single starch molecules. These osmotically active saccharide molecules draw water from the lateral spaces into the intestinal lumen. Water in the lateral spaces is quickly replaced by water from the intestinal capillaries, so water is essentially drawn into the intestine from the vascular system. As digestion proceeds, the saccharide molecules are absorbed, thus reducing the number of particles and lowering the osmotic pressure of the intestinal lumen. As solute molecules are absorbed, water follows them osmotically back through the epithelium and into the blood vascular system. *The cardinal rule of water movement in the intestine is that water moves in whatever direction necessary to keep ingesta iso-osmotic,* entering the gut when ingesta is hyperosmotic and leaving the gut when ingesta is hypo-osmotic. This fact has important clinical implications in the pathophysiology of diarrhea, as discussed later.

Active Secretion of Electrolytes From the Crypt Epithelium Leads to Intestinal Water Secretion

In contrast to the absorptive function of the villous cells, the crypt cells have a secretory function. This secretory function appears to use a chloride transport mechanism. The mechanism seems to be similar to coupled sodium chloride transport, as occurs in the villous enterocytes, except that the direction of transport is reversed. In the crypt cells, the coupled sodium chloride transport mechanism is on the basolateral membrane, in contrast to its position on the apical membrane of the villous cells. The effect of this arrangement is to pump Na^+ and Cl^- into the crypt enterocytes from the lateral spaces. As these ions are transported into the enterocytes, Na^+ is quickly pumped out by the Na^+,K^+-ATPase pump. In contrast, Cl^- is trapped within the cells, reaching relatively high intracellular concentrations. Under appropriate stimuli, Cl^- channels in the apical membranes of the crypt cells are opened, and the pent-up chloride from within the cells flows down its concentration gradient into the lumen of the crypt. (Ion channels and their regulation in cellular membranes are discussed in Chapter 1.) Movement of Cl^- into the lumen of the crypts creates an electrical attraction for Na^+, which move into the luminal fluid from the lateral spaces through the paracellular route. Water follows Na^+ and Cl^- osmotically; thus chloride, sodium, and water are secreted from the crypt epithelium (Fig. 30.21).

You may find the concept of ion transport processes jumping from one side of the cell to another intuitively "unappealing," especially when you consider that the cells of the intestinal crypts



• **Fig. 30.21** Water and electrolyte secretion in the crypts is affected by the secretion of chloride (Cl^-) from the apical membrane of crypt enterocytes. Sodium (Na^+) moves into the lumen by the paracellular route and electrically balances Cl^- secretion. Water follows osmotically, the net effect being secretion of a sodium chloride (NaCl) solution into the crypt lumen. In the crypts the coupled NaCl absorption mechanism appears to exist on the basolateral membrane, with Cl^- gates present on the apical membrane. The opening of Cl^- gates on the crypt cell apical membranes initiates crypt secretion. The membrane position of the coupled NaCl transport process reverses itself, moving from the basolateral membrane to the apical membrane as the cells mature and move up the villi.

will eventually mature and migrate up the villi to assume absorptive, as opposed to secretive, roles. Consider, however, that ion transport mechanisms are simply proteins inserted into the cell membranes. As with other cellular proteins, they are synthesized within the cell under the direction of genetic code. The state of maturity and cellular differentiation dictates the membrane position into which the newly synthesized proteins are directed. The differential distribution of membrane proteins to one side of a cell or another is called *polarization.* Enterocytes are said to be "polarized" with respect to membrane function.

The triggering mechanism that activates water secretion from the crypts is the opening of the chloride gates in the crypt enterocyte apical membrane. Much study has been devoted to determining the factors controlling the opening of the chloride gates in crypt cells. One important factor in regulating chloride gates appears to be the activity of the adenylate cyclase enzyme and the intracellular concentration of cyclic adenosine 3′,5′-monophosphate (cyclic AMP, or cAMP). (The role of adenyl cyclase and cAMP in cellular regulation is discussed in Chapter 1.) As cAMP concentrations rise, chloride gates open, and secretion of water and electrolytes is stimulated. Vasoactive intestinal peptide originating from effector neurons of the mucosal plexus is probably an important normal regulator of cAMP and chloride gates in crypt apical membranes. Of perhaps greater medical significance than the normal regulation of this process is the existence of pathological, or abnormal, activators of crypt cell adenyl cyclase (see later section on the pathophysiology of diarrhea).

The physiological function of water and electrolyte secretion by the crypts is to maintain an appropriate hydration and ionic environment for digestion and absorption. Ingesta must be kept sufficiently moist to allow for the mixing of nutrients with digestive enzymes and for the circulation of digested nutrients in contact with absorptive surfaces. In addition, a constant supply of sodium must be available to promote the sodium cotransport necessary for the absorption of several nutrients. The regulated process of water and electrolyte secretion from the crypts ensures the continual availability of water and sodium in the gut lumen.

Gastrointestinal Blood Flow

Water and Solute Movement Between the Lateral Spaces and Villous Capillaries Is Subject to the Same Forces That Govern Water and Solute Movement Between the Extracellular and Vascular Fluids in Other Tissues

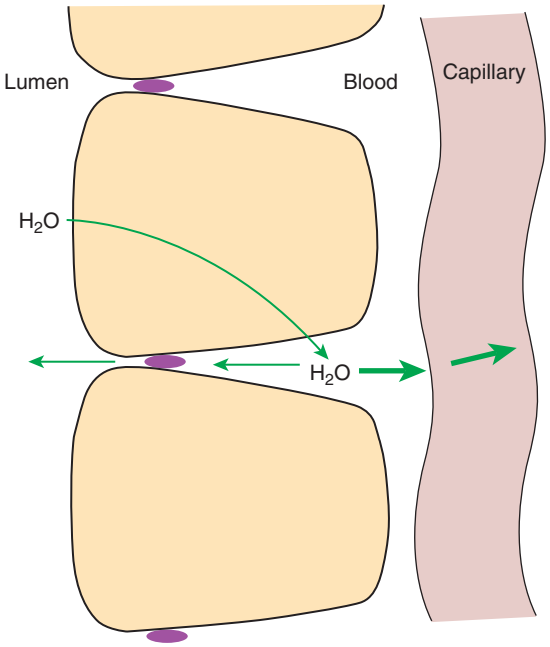
Water and all other nutrients, whether they are absorbed through the transcellular or paracellular routes, enter the ECF of the lateral spaces before entering the vascular system. Therefore the movement of ECF components into capillaries is of particular importance to intestinal absorption. The physical laws determining the distribution of water between the intravascular and extravascular fluid are the same in the villi as in other tissues. These Starling laws (which can be reviewed in Chapters 1 and 23) simply state that the movement of water is determined by the algebraic sum of osmotic and hydrostatic (created by water pressure) forces.

Absorbed Nutrients Enter the Capillaries by Diffusion From the Lateral Spaces

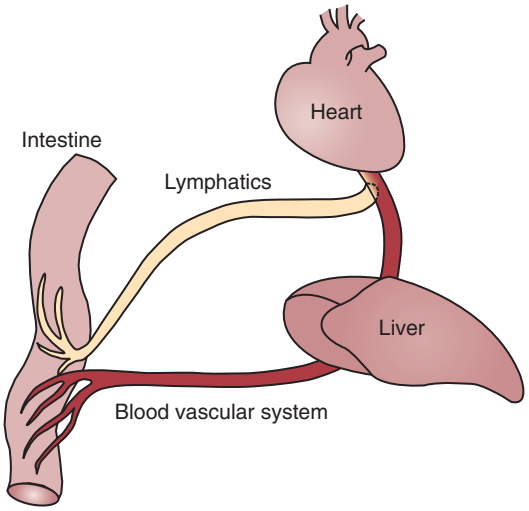
The collective action of the various intestinal absorptive mechanisms concentrates solutes (nutrients) in the lateral spaces. When the concentrations of individual solutes in the lateral spaces exceed their concentrations in blood, a gradient favoring the diffusion of nutrients from lateral spaces into the capillaries is established. The movement of solutes by diffusion into the capillaries creates an osmotic force that draws water into the capillaries (water follows solute). In addition, the oncotic force (the osmotic force exerted by plasma proteins; see Chapters 1 and 23) also tends to draw water into the capillary lumen. Moreover, hydrostatic pressure in the lateral spaces may force water directly into the capillaries. Lateral-space hydrostatic pressure can be created by the osmotic effect of absorbed solutes. As these solutes attract water from the intestinal lumen, the lateral spaces become distended, developing a small amount of hydrostatic pressure. There are two exits for the relief of this pressure: the tight junctions and the capillary endothelium, with the endothelium presenting the route of least resistance to water flow. Thus water under slight pressure within the lateral spaces tends to flow into the capillaries rather than into the intestinal lumen (Fig. 30.22).

A Countercurrent, Osmotic-Multiplier System May Increase the Osmolality of Blood at the Tips of the Villi, Further Promoting Absorption of Water Into the Blood

The villous vascular system consists of an arteriole rising up the central portion of the villi and dividing at the tip into many capillaries, which course down the outer portion of the villous stroma between the mucosa and artery. This arrangement provides for direct countercurrent flow of blood; that is, blood coming down the venules passes close to blood flowing in the opposite direction up the arteriole. Because blood in the venules contains absorbed nutrients, its osmolarity could be expected to be slightly higher than that of blood entering the villi in the arteriole. This slight



• **Fig. 30.22** Water (H_2O) enters the lateral spaces because of osmotic effects created by absorbed solutes, thus creating a hydrostatic pressure head in the lateral spaces. Under pressure, the lateral-space solution can exit through the tight junctions or through the basement membrane of the capillaries. Under normal conditions, the route of least resistance is into the capillaries, resulting in little movement of water from the lateral space into the intestinal lumen.



• **Fig. 30.23** All blood exiting the gut flows through the liver before returning to the heart. Lymphatic drainage from the gut bypasses the liver, entering the bloodstream through the thoracic duct.

difference in osmolarity can be multiplied and perpetuated by the countercurrent flow characteristics of the arterial and venous blood supplies. These conditions create a potential for the creation of an osmotic gradient along the villi; some researchers calculate osmolalities near the tips of the villi to be as high as 600 mOsm, approximately twice that of blood entering the base of the villi. (The characteristics of a countercurrent osmotic multiplier are further explained in Chapter 43 in reference to the renal loop of Henle.) The existence of the villous countercurrent osmotic multiplier is still somewhat controversial, and its presence may depend on the species in question. The effect of this osmotic-multiplier system would be to accentuate all the osmotic forces that result in the movement of water from lumen to lateral spaces and from lateral spaces to capillaries.

Disturbances in the Venous Drainage From the Intestine Can Greatly Affect the Mechanisms of Capillary Absorption in the Villi

With the exception of blood from the terminal colon and rectum, all venous blood from the GI tract is collected into the hepatic portal vein and passes through the liver before entering the vena cava and returning to the heart (Fig. 30.23). Because of this system, the liver may modify the nutrient-rich blood leaving the intestine. The liver can thus regulate the nutrient concentration of blood reaching general body tissues, keeping it relatively constant. This particular vascular arrangement of the GI system results in the passage of blood through two capillary beds, one in the gut wall and one in the liver, before its return to the heart. In most tissues, arterial hydrostatic pressure forces blood through the capillary beds. In the liver, however, this is not the case, because most of the arterial hydrostatic pressure has been dissipated during flow of blood through the intestinal capillaries. The following two circumstances tend to overcome this problem and allow hepatic blood flow to occur:

1. The capillaries (referred to as *sinusoids*) of the liver are comparatively large and thus offer little resistance to flow; therefore they can function in a low-pressure system.
2. The venous outflow of the liver goes directly into the thoracic vena cava.

The bellows-like action of the thorax transmits a negative pressure to the thoracic vena cava, which tends to aspirate blood from the hepatic veins and abdominal vena cava. Under normal circumstances, these conditions allow blood to flow readily from the intestine through the liver. However, small changes in circulatory function can have a large effect on GI blood flow. If the pumping capacity of the heart becomes reduced, it cannot quickly remove returning venous blood. Accumulation of blood and an increase in pressure in the thoracic vena cava result. This increase in pressure interferes with the flow of blood out of the liver, which in turn reduces blood flow out of the intestine. This sequence of events makes the GI system particularly susceptible to right-sided heart failure, in which the heart's pumping action is compromised.

Diffuse liver disease, in addition to right-sided heart failure, can also interfere with GI blood flow. In this condition, the resistance to blood flow through the liver is increased because of pressure on the sinusoids. Small rises in hepatic flow resistance can have large effects on intestinal blood flow because the pressure gradient across the hepatic portal vein is normally small. When the flow of blood out of the intestine is impaired, hydrostatic pressure in the capillaries of the villi is increased; the higher pressure tends to offset the osmotic and hydrostatic forces promoting water absorption, and thus water absorption is impaired.

Digestion and Absorption of Fats

Detergent Action and Enzymatic Action Is Necessary for the Digestion and Absorption of Lipids

Lipids, or fats, present a special digestive problem to the animal because they do not dissolve in water, the major medium in which

most body processes, including digestion, occur. Detergent action is necessary to emulsify or dissolve lipids so that they may be subjected to the actions of water-soluble hydrolytic enzymes in the gut. The problem of solubility makes the mechanics of digestion and absorption of lipids somewhat different from that of proteins and carbohydrates. For that reason, lipid assimilation is discussed here in a separate section.

Lipids make up a large portion of the diets of carnivores and omnivores, whereas they usually form a minor portion of the natural diets of adult herbivores. Nonetheless, it appears that herbivorous species have the capacity to digest and absorb lipids in quantities considerably higher than found in their natural diets, and frequently, supplemental lipids are added to the diets of performance horses and high-producing dairy cows. The neonates of all mammalian species have a high capacity for lipid digestion and absorption because milk has a high fat content.

The primary dietary lipid is *triglyceride,* which may originate from either plant or animal sources. Other important dietary lipids are *cholesterol* and *cholesteryl ester* from animal sources, waxes from plant sources, and *phospholipids* from both plant and animal sources. Fig. 30.24 illustrates the structures of these dietary lipids. In addition, the lipid-soluble vitamins A, D, E, and K are absorbed along with the other dietary lipids.

Lipid assimilation can be divided into four phases: (1) emulsification, (2) hydrolysis, (3) micelle formation, and (4) absorption. *Emulsification* is the process of reducing lipid droplets to a size that forms stable suspensions in water or water-based solutions. In the gut, the emulsification phase begins in the stomach as the lipids are warmed to body temperature and subjected to the intense mixing, agitating, and sieving actions of the distal stomach. This distal-stomach activity tends to break lipid globules up into droplets that pass into the small intestine. In the small intestine, emulsification is completed by the detergent action of bile acids and phospholipids. (See Chapter 29 for a discussion of bile formation and secretion.) These bile products reduce the surface tension of the lipids and allow the droplets to become even further divided and reduced in size (Fig. 30.25).

While in the bile-coated, or emulsified-droplet, stage, the lipids are subject to the actions of hydrolytic enzymes. Hydrolysis of triglyceride, the major dietary lipid component, occurs because of the combined action of the pancreatic enzymes *lipase* and *colipase.* Lipase is an enzyme secreted, in its active form, from the pancreas. However, lipase cannot directly attack the emulsified lipid droplets in the gut because it cannot penetrate the coat of bile products surrounding the droplets. The function of colipase, a relatively short peptide, is to "clear a path" through the bile products, giving lipase access to the underlying triglycerides. Lipase cleaves the fatty acids off each end of the triglyceride molecule but does not attack the central fatty acid, resulting in the formation of two *free* (or *nonesterified*) *fatty acids* and a *monoglyceride* from each molecule of triglyceride hydrolyzed (Fig. 30.26).

Other lipid-digesting pancreatic enzymes are *cholesterol esterase* and *phospholipase.* The products of these enzymes are nonesterified fatty acids, cholesterol, and lysophospholipids.

The products of hydrolytic lipid digestion (fatty acids, monoglycerides, etc.) combine with bile acids and phospholipids to form *micelles,* small water-soluble aggregations of bile acids and lipids. Micelles are considerably smaller than the emulsified fat droplets from which they are derived (see Fig. 30.25). The soluble micelles allow the lipids to diffuse through the gut lumen into the unstirred water layer and into close contact with the absorptive surface of the apical membrane (Fig. 30.27; see also Fig. 30.26).

Lipids Are Absorbed Through the Apical Membrane by Carrier Proteins and Simple Diffusion

The process of lipid absorption into the enterocytes is incompletely understood. As the micelles come close to the surface of the enterocytes, the various lipid components diffuse the short distance through the glycocalyx to the apical membrane by means of special *fatty acid–binding proteins* (not shown in Fig. 30.27). Fatty acids in the micelles appear to be taken up and transported across the apical membrane by special fatty acid–binding proteins in the apical membrane. Other micellar components appear to simply diffuse into the apical membrane; they include such lipids as monoglycerides, cholesterol, and vitamin A. The apical membrane, as with other cellular membranes, is composed primarily of phospholipids (see Chapter 1). The highly hydrophobic products of lipid digestion are soluble in the phospholipid matrix of the membrane and thus may diffuse freely through the apical membrane and into the cell. Fig. 30.27 illustrates lipid absorption from micelles.

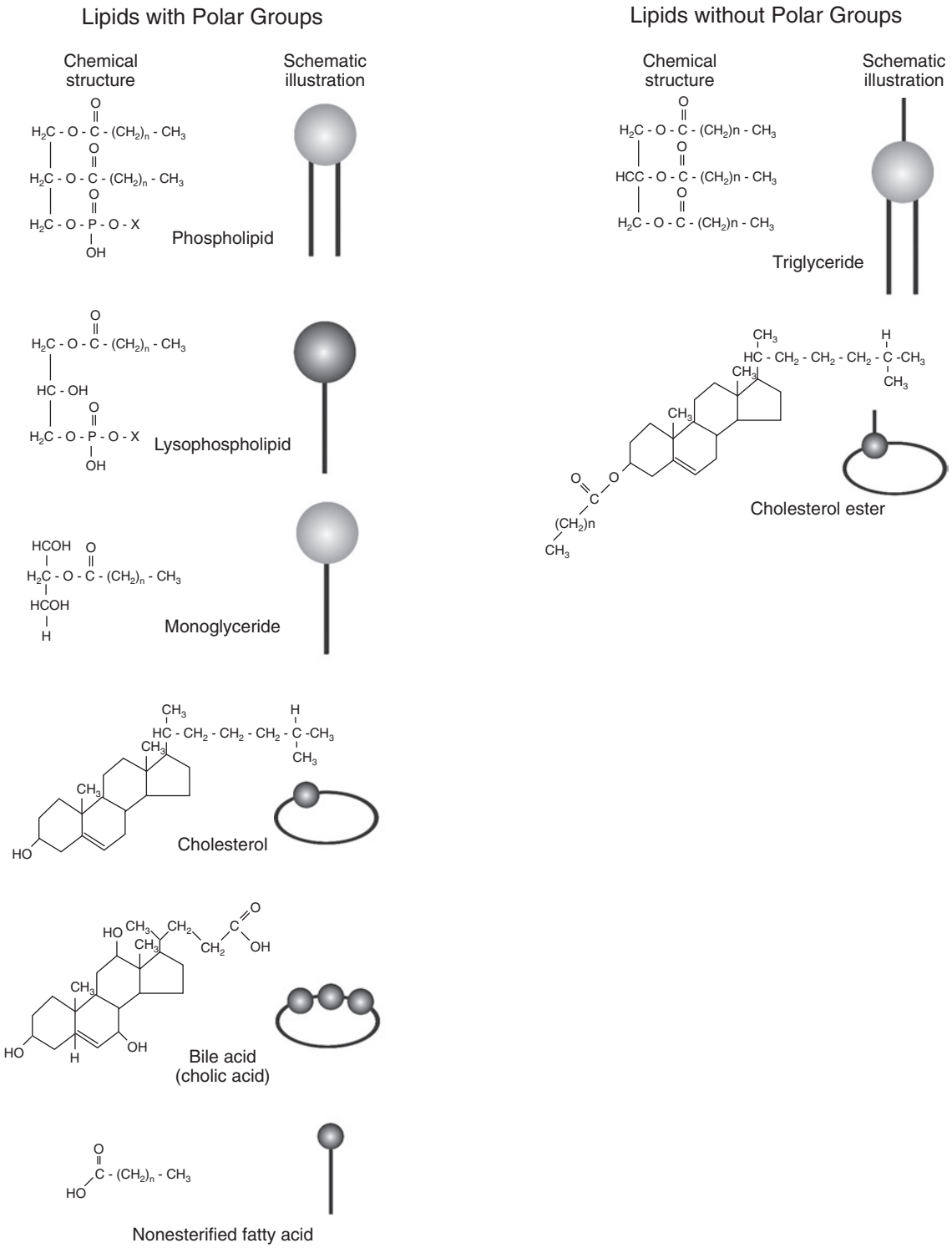
Bile Acids Are Reabsorbed From the Ileum by a Sodium Cotransport System

All components of the micelle diffuse into the enterocytes except the bile acids. Bile acids remain in the lumen of the gut, being separated from the other micellar elements as absorption proceeds. By the time bile acids reach the ileum, they are in a relatively free state, devoid of other lipids. Localized in the ileum is a specific bile acid transport system. This system operates by sodium cotransport and results in the nearly complete reabsorption of bile acids. After absorption, bile acids are transported directly back to the liver by the portal vasculature. The liver efficiently extracts bile acids from the portal blood, so normally the concentration of bile acids in the nonportal blood (systemic circulation) is small. The bile acids extracted by the liver are recycled into the bile. This recycling process occurs repeatedly, so the entire mass of bile acids within the body is circulated through the intestine several times per day.

Absorbed Lipids Are Packaged Into Chylomicrons Before Leaving the Enterocytes

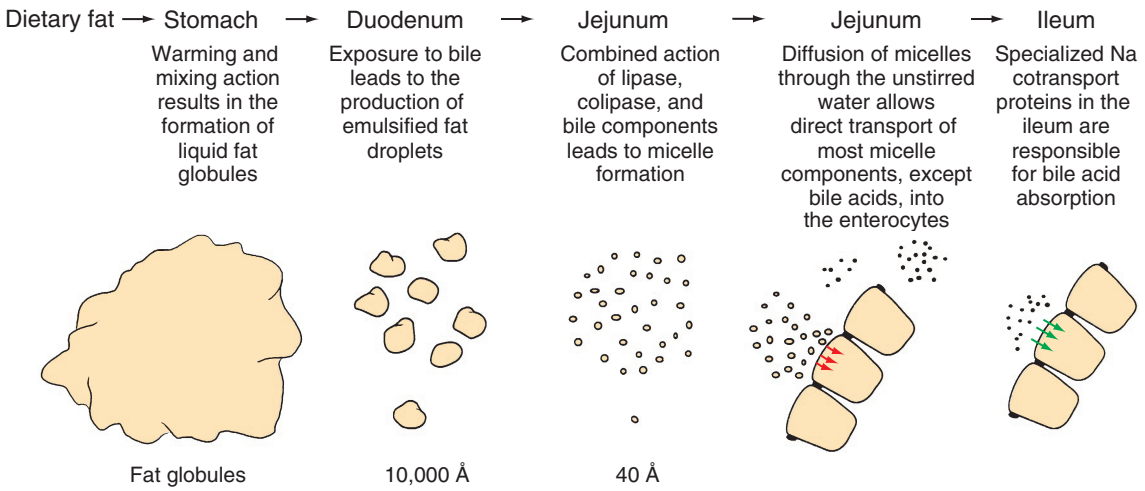
After passing the apical membrane, the absorbed lipids are quickly picked up by carrier molecules and transported within the cell to the endoplasmic reticulum. When they're on the smooth endoplasmic reticulum, the major lipids are reesterified to form triglyceride and phospholipids. The reesterified lipids are then packaged with cholesterol, minor dietary lipids, and proteins from the rough endoplasmic reticulum into structures known as *chylomicrons.* Chylomicrons are spherical structures with a core of triglyceride and cholesterol ester and a surface of phospholipid and cholesterol. The phospholipid and cholesterol are arranged with their *hydrophobic* (water-repelling) ends facing the core lipids and their *hydrophilic* (water-attracting) ends facing the surface of the chylomicron particle (Fig. 30.28). This arrangement of surface lipid makes the chylomicron water-soluble. A small number of special protein molecules are also present on the chylomicron surface. These proteins help to stabilize the surface and to direct the metabolism of the particle.

After their formation, chylomicrons are expelled from the basolateral membrane into the lateral spaces. Unlike most other nutrients entering the lateral spaces, chylomicrons are too large to

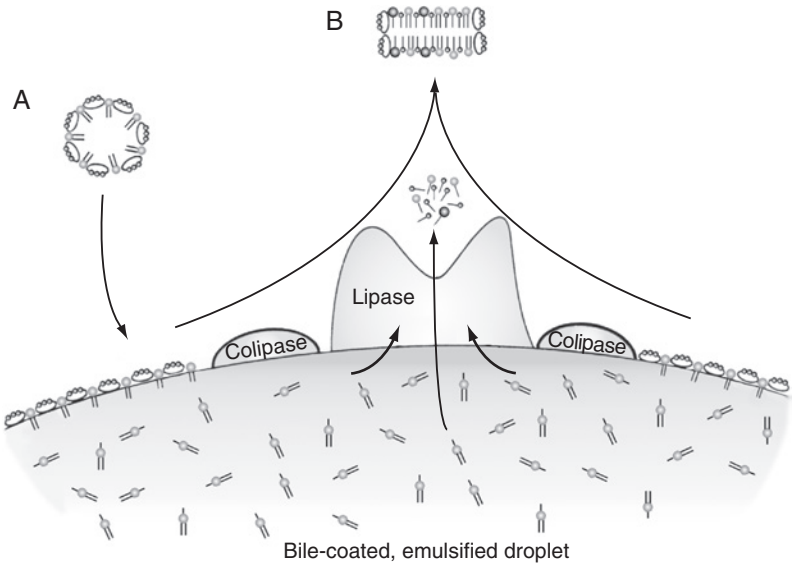


• **Fig. 30.24** Chemical structures and their schematic representations for lipid molecules involved in fat digestion and absorption. *n,* Number of carbon atoms in fatty-acid chains; *X,* phospholipid head group, most often choline.

pass through the basement membrane of the intestinal capillaries. Thus chylomicrons cannot be absorbed through the intestinal blood system. Rather, they travel through the intestinal lymphatics, which eventually form a major abdominal lymph duct that passes through the diaphragm and into the *thoracic duct.* The major lymph-collecting vessel of the body, the thoracic duct, empties into the vena cava. Through this means, chylomicrons eventually reach the blood vascular system. During absorption of a fatty meal, the character of intestinal lymph changes from water-clear to milky white because of the presence of chylomicrons. After a fatty meal, this milky white color can even be seen in blood plasma. In normal animals, this white color in blood plasma,



• **Fig. 30.25** Sites and reactions involved in fat digestion and absorption. *Å,* Angstroms.



• **Fig. 30.26** Portion of the surface of a bile-coated, emulsified fat droplet. Bile components reach the surface of the droplet through micelles (A), coming from the gallbladder. Colipase clears bile constituents from an area of the surface of the droplet, allowing the attachment of lipase. Lipase catalyzes the formation of fatty acids and monoglycerides from triglycerides. The surface components and products of lipase action combine to form micelles (B), containing fatty acids, monoglycerides, and bile constituents.

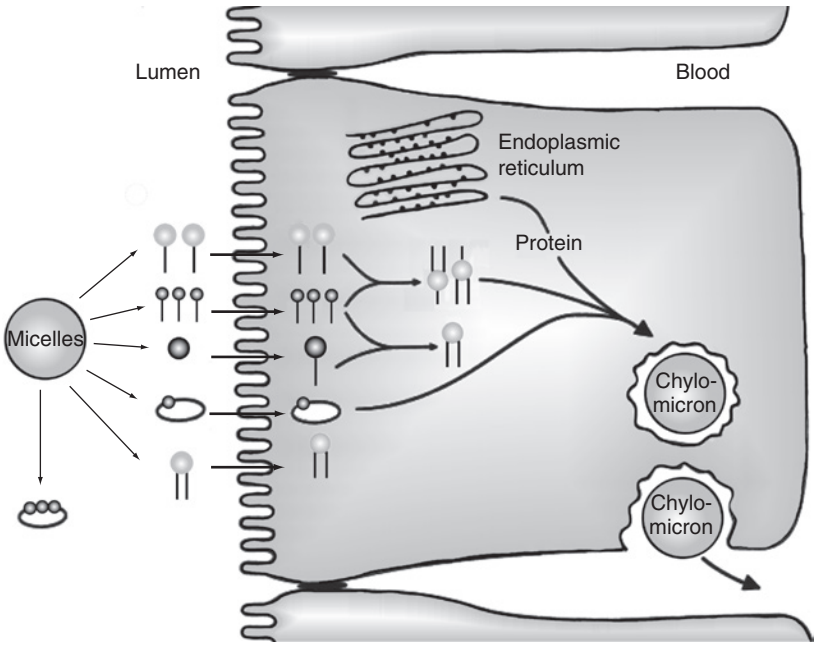
known as *lipemia,* is transient, disappearing within 1 to 2 hours after digestion of the meal. The metabolic fate of the chylomicrons is discussed in Chapter 32.

Growth and Development of the Intestinal Epithelium

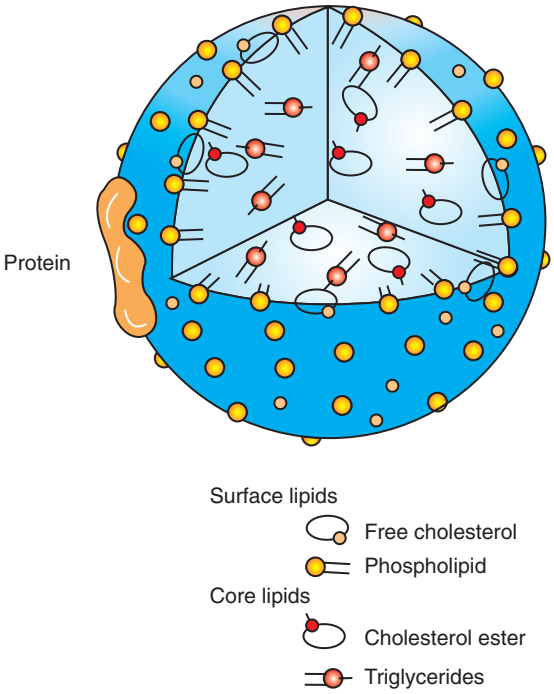
The Length of Intestinal Villi Is Determined by the Relative Rates of Cell Loss at the Tips and Cell Replenishment at the Base

Division and replication of enterocytes occur in the crypts only. Crypt enterocytes are highly mitotic and regenerate rapidly. In fact, the intestinal crypt cells are among the most rapidly regenerating cells of the body, representing the single greatest need for protein synthesis in nongrowing animals. As crypt cells multiply, they migrate upward onto the villi, pushing other villous cells ahead of them, so there is a continuous progression of cells migrating up the villi. As the cells migrate they mature, changing from relatively undifferentiated cells in the crypts to highly specialized absorptive cells on the villi. As the cells reach the tips of the villi, they are lost because of age and exposure to gut contents. The length of villi is determined by the rate at which cells are lost at the tips and the rate at which cells from the crypts replace them. An increase in cell loss at the villi tips, relative to crypt cell replication, results in shortening of the villi. Conversely, fast replication of crypt cells, relative to cell loss, leads to lengthening of the villi. The time taken for an enterocyte to migrate from its site of origin in the crypt to the tip of a villus varies with species and physiological state; on average, however, the turnover time of enterocytes is 4 to 7 days.

The rate of cell replication in the crypts appears to be stimulated by several GI hormones. When appetite and feed intake increase,



• **Fig. 30.27** Lipid absorption from micelles with subsequent formation of chylomicrons. As micelles come close to the apical membrane, lipid constituents, except bile acids, are transported through the membrane into the cell. When in the enterocyte, triglycerides are reformed from fatty acids and monoglycerides. Triglycerides are then packaged into the core of chylomicrons for transport out of the cell. The chylomicron surface is coated with phospholipids, cholesterol, and proteins.



• **Fig. 30.28** Chylomicron structure. Special proteins and lipids with polar groups form the surface coat, whereas nonpolar lipids form the core of the particle.

there is an overall increase in the secretion of GI hormones. This in turn leads to a rise in crypt cell proliferation. An increase in the rate of crypt cell replication adds cells to the villus at a rate faster than they are lost at the tips, resulting in longer villi. Appetite and feed intake may increase because of conditions of greater energy need, such as lactation, exercise, and cold environmental temperatures. The greater villi length provides greater digestive and absorptive capacity to match the need created by higher feed intake. Thus the functional capacity of the intestine is adjusted to match the nutrient needs of the animal.

Digestion in the Neonate

During the First Few Hours of Life, Proteins Are Not Digested but Are Absorbed Intact

In general, a major function of digestion is to break down proteins by hydrolysis. Under most circumstances, this process is a benefit to the animal not only from a nutritional and digestive standpoint, but also from a toxicological and allergic standpoint; potentially toxic or allergenic proteins are broken down before they are absorbed into the body. In the special case of some neonates, however, there is a need to absorb proteins intact. In most livestock species, including horses, cattle, sheep, and swine, essentially no antibodies are passed through the placenta from the dam to fetus. Thus young livestock are born without the immunological protection of their mother's antibodies. In these species, antibodies from the mother must be acquired through ingestion of colostrum, the special mammary secretion present at birth. In these animals, the digestive tract at birth is altered from the adult state, so the antibody proteins are absorbed intact rather than after digestion. There are three primary alterations, as follows:

- Acid secretion from the stomach is delayed for several days after birth.
- A similar delay appears in the development of pancreatic function, and thus acid and trypsin digestion of proteins is avoided.
- A specialized intestinal epithelium present at birth only is capable of engulfing soluble proteins in the intestinal lumen and discharging them into the lateral spaces.

The fetal intestinal epithelium has the same villous structure as the mature epithelium, but the villi are covered with special enterocytes capable of protein absorption. Immediately after birth, this special epithelium starts to disappear, and it is essentially gone after 24 hours. The loss of the protein-absorptive function in the neonate is referred to as *gut closure.*

The Major Intestinal Disaccharidase Switches From Lactase to Maltase With Maturity

Lactose from milk is the major carbohydrate in the diets of neonatal and young mammals; thus all mammals are born with high intestinal lactase activity. In contrast, maltase activity, necessary for digesting the products of luminal starch digestion, is weak or absent for several weeks after birth. As animals progress toward weaning, lactase activity wanes and maltase activity increases, allowing the animals to shift from lactose to starch as a carbohydrate source. In many species of adult animals, lactase activity is practically nonexistent.

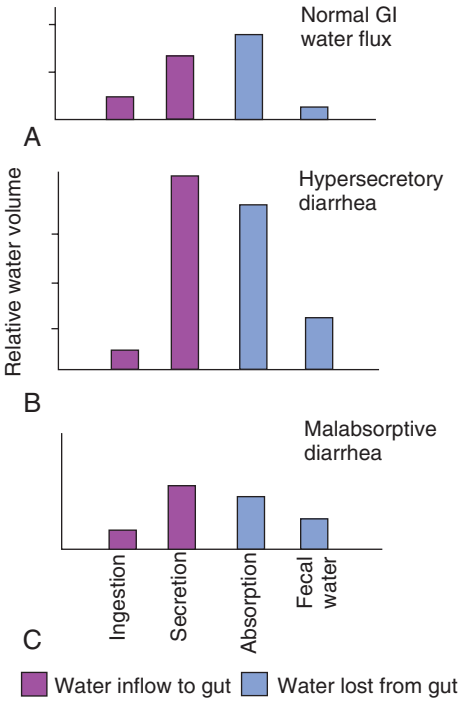
Pathophysiology of Diarrhea

Diarrhea refers to an increase in frequency of defecation or fecal volume. Volume often increases in diarrhea primarily because of increased water content. The amount of water passed in feces is the algebraic sum of GI water input and water absorption. As discussed earlier, water in the gut results from (1) ingested water, (2) water secreted by glands of the GI system, and (3) water secreted or lost directly through the mucosal epithelium. Under most circumstances, the amount of water secreted into the gut greatly exceeds the amount ingested. Normally, the amount absorbed is just slightly less than the sum of the amounts secreted and ingested, leaving a small remainder for passage in the feces (Fig. 30.29A).

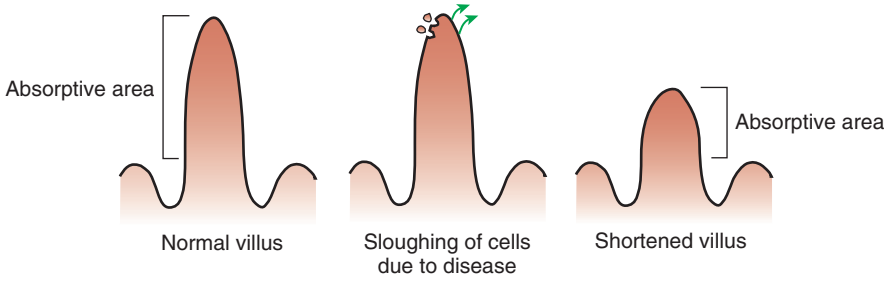
Diarrhea Occurs When There Is a Mismatch Between Secretion and Absorption

The amount of water in the feces is the result of the balance between secretion and absorption. *Malabsorptive diarrhea* occurs when absorption is inadequate to recover a sufficient portion of secreted water, as illustrated in Fig. 30.29C. Malabsorptive diarrhea usually occurs because of the loss of GI epithelium. In most instances, such losses occur because of viral, bacterial, or protozoal infections. Viral infections often cause particularly severe destruction of villous epithelium. These infections result in the loss of enterocytes from the villi. As noted previously, villous length is determined by the relative rates of cell loss and cell replacement (Fig. 30.30).

Intestinal infections result in decreased villous length because the rate of cell loss is higher than the rate of cell replacement. Short villi cause impaired absorption for two reasons: (1) there is an absolute loss of absorptive intestinal surface area and (2) the cells that are lost are the mature cells from the upper regions of the villi. It is these mature cells that possess the enzymes of membranous-phase digestion and the transport proteins for sodium cotransport; loss of these cells results in impairment of digestion and absorption of nutrients. Because nutrient absorption is necessary for the osmotic absorption of water, water absorption is diminished when nutrient absorption is impaired.



• **Fig. 30.29** Pathophysiology of diarrhea. The bars represent the relative amounts of water entering or leaving the gut. Fecal volume is the sum of the water ingested and the water secreted minus the water absorbed. Therefore fecal volume depends not on the amount of water entering the gut, but rather on the balance between water influx and efflux. See text for descriptions of hypersecretory (B) and malabsorptive (C) diarrhea.



• **Fig. 30.30** Shortening of villi caused by increased cell loss. Many infectious diseases result in an increased rate of cell sloughing from the villi. As cells are lost, the villus shrinks to fill in the gap in the epithelial coat. If villous height is to be maintained in the presence of rapid loss of enterocytes, the rate of recruitment of new cells, generated in the crypts, must be increased. Therefore when the rate of cell loss exceeds the capacity for cell replacement, shortened villi with reduced absorptive surfaces and relatively immature enterocytes appear.

Secretory diarrhea occurs when the rate of intestinal secretion increases and overwhelms the absorptive capacity. Most cases of hypersecretory diarrhea result from inappropriate secretion from the small intestinal crypts. This occurs when the normal secretory mechanism of the crypt epithelium (as discussed earlier) is abnormally stimulated. Some types of pathogenical bacteria produce toxins known as *enterotoxins*. These toxins bind to enterocytes and stimulate adenyl cyclase activity and the production of cAMP within the cells, resulting in opening of the chloride gates and the secretion of water and electrolytes from crypt epithelium. If the stimulation is mild, the gut may respond with an increase in absorption, and diarrhea does not result. However, when the secretion exceeds the capacity of the gut to increase absorption, as illustrated in Fig. 30.29B, diarrhea results. Hypersecretory diarrhea has devastating effects on the water, electrolyte, and acid-base status of animals, especially neonates. Hypersecretory diarrhea caused by enterotoxin-producing *Escherichia coli* is an extremely common disease of neonatal calves and pigs. This disease causes large economic losses in the cattle and swine industries as a result of treatment costs and death.

CLINICAL CORRELATIONS

Diarrhea With Dehydration and Acidosis in a Calf

History

You are asked to examine a 2-day-old calf. The owner reports that the calf appeared normal the night before, but this morning she is recumbent and will not rise. In addition, she shows no interest in suckling a bottle.

Clinical Examination

The calf's body temperature is subnormal. The mouth is dry, and the eyes are sunken into the orbits. The ears, tail, and distal legs are cool to the touch. The tail and perineum of the calf are wet. As you remove your thermometer, the calf passes a stream of liquid feces. The feces are almost clear and slightly yellow, with the consistency of water. Simple laboratory tests indicate that the packed cell volume is 50% (normal, 30%–35%) and the serum total-solids concentration 7.5% (normal, 5.5%–6.5%).

Comment

The calf has diarrhea, and the physical examination and laboratory findings indicate a state of advanced dehydration. Loss of body fluid volume is so severe that the calf appears to be in or near a state of hypovolemic shock. Although you cannot be certain from your examination of the calf at the farm, the severity of dehydration, the rapidity of onset, and the age of the calf all suggest a hypersecretory diarrhea caused by enterotoxigenic *E. coli* bacteria. Animals with such clinical signs are usually severely acidotic, although blood pH is seldom measured in field cases. Diarrhea, acidosis, and dehydration occur because toxins produced by the bacteria stimulate opening of chloride gates in the apical membranes of crypt cells, stimulating copious secretion of water and electrolytes, including bicarbonate. The sodium cotransport system on the villi is unaffected by the bacterial toxin, but the simultaneous presence of glucose and sodium in the lumen is necessary to promote cotransport, which can offset some of the fluid and electrolyte losses caused by hypersecretion from the intestinal crypts.

Treatment

Vascular volume expansion and correction of acidosis are primary concerns in such cases. This calf should receive 2 L of alkalizing fluids by rapid intravenous (IV) administration, with an additional 2 L or more over the next 24 hours. Frequently, the response of calves to such treatment is remarkable, and calves that appear almost dead can often be saved by vigorous fluid therapy. After the initial replacement of lost fluids by IV therapy, further dehydration caused by ongoing fluid losses can be prevented by oral administration of fluids containing glucose and sodium.

Juvenile Pancreatic Atrophy in a Dog

History

You are presented with a thin, 3-year-old German shepherd. The owners report that the dog appeared normal until 6 months ago. At that time they noticed that he started losing weight and began to eat his own feces. Recently the weight loss has become more severe, even though he has had a good appetite and seems normal otherwise. Lately, the owners have noted that the dog seems to pass a large amount of feces, which is soft and gray with a claylike consistency.

Clinical Examination

Physical examination reveals an extremely thin dog with a dull, uneven hair coat. Other physical findings are unremarkable, and the dog seems bright and friendly. You hospitalize the animal for further testing and note that he readily eats two cans of commercial dog food per day. Laboratory analysis of feces collected over a 24-hour period reveals that the dog is passing 25 g of fat in the feces per day (normal, <5 g, assuming normal diet).

Comment

This degree of fat malabsorption is characteristic of *pancreatic exocrine insufficiency.* Because there is insufficient pancreatic lipase, fats cannot be hydrolyzed to fatty acids for absorption and thus pass unabsorbed through the gut. Other laboratory tests for assessing pancreatic exocrine function include (1) examination of blood for the presence of orally administered markers that require pancreatic enzymes for digestion and absorption and (2) direct examination of feces for the presence of pancreatic enzymes or undigested nutrients. Currently, the most definitive laboratory test for the diagnosis of pancreatic exocrine insufficiency is serum *trypsin-like immunoreactivity.* Under normal conditions, a very small proportion of the trypsinogen synthesized by the pancreatic acinar cells escapes into the peripheral bloodstream, and a similar small portion of trypsin is absorbed from the intestine. Although these concentrations are too small to have an effect on the body, they can be measured in blood. The assay is based on an antibody reaction and the quantitative result is referred to as trypsin-like immunoreactivity. A low value for *trypsin-like immunoreactivity* is indicative of pancreatic exocrine insufficiency.

Treatment

Feeding highly digestible diets mixed with commercially prepared pancreatic enzymes is usually successful in promoting adequate nutrient absorption in animals with *juvenile pancreatic atrophy.* Digestion may not be completely normal but is sufficient for the dog to attain a normal body weight. Treatment must be continued for life. You may question how orally administered pancreatic enzymes can survive intact through the proteolytic environment of the stomach. Undoubtedly, portions are destroyed, but sufficient enzyme appears to make it through the stomach to be effective.

Practice Questions

1. Finding triglycerides and starch in the feces of a thin dog with a normal feed intake would suggest:
- **a.** Malabsorption
- **b.** Maldigestion

2. Which statement about the tight junctions is *false?*
- **a.** Tight junctions encircle the enterocyte near its apical end.
- **b.** Tight junctions form the dividing line between the apical membrane and the basolateral membrane.
- **c.** Tight junctions are impermeable to water.
- **d.** Tight junctions separate the lateral space from the intestinal lumen.
- **e.** Tight junctions are the only points that attach enterocytes together.

3. Which of the following molecules is consumed during the process of hydrolytic digestion?
- **a.** Glucose
- **b.** Alanine
- **c.** Dipeptides
- **d.** Fatty acids
- **e.** Water

4. A drug that blocks the activity of the Na^+,K^+-ATPase pump could be expected to have what effect on sodium-glucose cotransport?
- **a.** Increased sodium-glucose cotransport
- **b.** Decreased sodium-glucose cotransport
- **c.** No effect on sodium-glucose cotransport

5. During sodium absorption by glucose cotransport:
- **a.** Chloride is absorbed by the paracellular route.
- **b.** Chloride absorption is not affected.
- **c.** Chloride is absorbed in exchange for bicarbonate.
- **d.** Chloride absorption is coupled with potassium absorption.
- **e.** Chloride is absorbed in exchange for hydrogen ion.

6. Before entering the intestinal capillaries, all nutrients pass through the:
- **a.** Apical membrane
- **b.** Tight junction
- **c.** Lateral space
- **d.** Basolateral membrane
- **e.** Enterocyte cytoplasm

Bibliography

Entry-Level Supplements

Barrett KE. *Chapters 15 and 16. in Gastrointestinal Physiology*. 2nd ed. McGraw-Hill Education; 2014.

Johnson LR. *Chapters 11 and 12 in Gastrointestinal Physiology*. 8th ed. Mosby; 2014.

Advanced Reading

Bornhorst GM, Paul Singh R. Gastric digestion in vivo and in vitro: how the structural aspects of food influence the digestion process. *J Food Sci*. 2008;73(5):R67–R80.

Hall EJ. Clinical laboratory evaluation of small intestinal function. *Vet Clin North Am Small Anim Pract*. 1999;29(2):441–469.

Johnson LR, ed. *Chapters 58 through 60, 65 & 66. Physiology of the Gastrointestinal Tract*. 5th ed. London, UK: Elsevier Inc.; 2012.

Kong F, Singh RP. Disintegration of solid foods in human stomach. *Annu Rev Food Sci Technol*. 2014;5:111–132.

Lundgren O. Enteric nerves and diarrhea. *Pharmacol Toxicol*. 2002;90(3):109–120.

Magallanes-Cruz PA, Flores-Silva PC, Bello-Perez LA. Starch structure influences its digestibility: a review. *Annu Rev Food Sci Technol*. 2014;5:111–132.

Mansbach CM 2nd, Siddiqi S. Control of chylomicron export from the intestine. *J Food Sci*. 2017;82(9):2016–2023.

Nagy B, Fekete PZ. Enterotoxigenic Escherichia coli (ETEC) in farm animals. *Vet Res*. 1999;30(2–3):259–284.

Naylor JM. Oral electrolyte therapy. *Vet Clin North Am Food Anim Pract*. 1999;15(3):487–504.

Rao MC. Oral rehydration therapy: new explanations for an old remedy. *Annu Rev Physiol*. 2004;66:385–417.

Shirazi-Beechey SP, Moran AW, Bravo D, Al-Rammahi M. Nonruminant nutrition symposium: intestinal glucose sensing and regulation of glucose absorption: implications for swine nutrition. *J Anim Sci*. 2011;89(6):1854–1862.

Spring KR. Routes and mechanism of fluid transport by epithelia. *Annu Rev Physiol*. 1998;60:105–119.

Stevens CE, Hume ID. *Comparative physiology of the vertebrate digestive system*. ed 2. Cambridge, UK: Cambridge University Press; 1995.

Thiagarajah JR, Donowitz M, Verkman AS. Secretory diarrhoea: mechanisms and emerging therapies. *Am J Physiol Gastrointest Liver Physiol*. 2016;310(9):G659–G668.

31

Digestion: The Fermentative Process

THOMAS H. HERDT

KEY POINTS

1. Fermentation is the metabolic action of bacteria.
2. The sites of fermentative digestion must be conducive to microbial growth.

Microbial Ecosystem of Fermentative Digestion

1. The microbes responsible for fermentative digestion include bacteria, fungi, and protozoa.
2. Cooperation and interplay among the many species of microbes give rise to a complex ecosystem in the forestomach and hindgut.

Substrates and Products of Fermentative Digestion

1. Plant cell walls are important substrates for fermentative digestion and significant nutrient sources for many species.
2. Nutrients other than cell walls are also subject to fermentative digestion.
3. Anaerobic conditions in the rumen result in metabolic activities leading to the production of volatile fatty acids.
4. Volatile fatty acids are important energy substrates for the host animal.
5. Fermentative digestion of protein results in the deamination of a large portion of amino acids.
6. When protein and energy availability in the forestomach are well matched, rapid microbial growth and efficient protein utilization result.
7. Microbial protein can be synthesized in the rumen from nonprotein nitrogen sources.

Reticulorumen Motility and Maintenance of the Rumen Environment

1. The physiological functions of the reticulorumen maintain an environment favorable to fermentation patterns beneficial to the host.
2. Rumen fermentation is maintained by selectively retaining actively fermenting material while allowing unfermentable residue to pass on to the lower digestive tract.
3. Digestibility and physical characteristics of feed have important influences on both the rate of particle passage from the rumen and the rate of feed intake.
4. Rumination, or cud chewing, has an important effect on the reduction of particle size and the movement of solid material through the rumen.
5. Water moves through the rumen at a much faster rate than particulate matter.
6. Rumen dilution rate has important influences on fermentation and microbial cell yield.

Control of Reticulorumen Motility

1. Reticulorumen motility is controlled by the central nervous system and affected by intraluminal conditions.

Omasal Function

1. Passage of material from the reticulum to the omasum occurs during reticular contraction.

Volatile Fatty Acid Absorption

1. Volatile fatty acids, representing 60% to 80% of the energy needs of the animal, are absorbed directly from the forestomach epithelium.

Rumen Development and Esophageal Groove Function

1. Significant changes in forestomach size and function occur with dietary changes in early life.
2. The esophageal groove diverts the flow of ingested milk past the forestomach and into the abomasum.

Function of the Equine Large Hindgut

1. The equine hindgut has a great capacity for fermentation.
2. The types of substrate and fermentation patterns are essentially identical for forestomach and hindgut fermentation.
3. The motility functions of the cecum and colon retain material for fermentation and separate particles by size.
4. The rate of fermentation and volatile fatty acid production in the equine colon is similar to that in the rumen.
5. Hindgut anatomy and function vary greatly among the many species of veterinary interest.

Fermentation Is the Metabolic Action of Bacteria

In fermentative digestion, molecular substrates are broken down by the action of bacteria and other microorganisms. Enzymatic hydrolysis of large molecules is an essential part of fermentative digestion, just as it is for glandular (nonfermentative) digestion. The major difference between the two processes is that the enzymes of fermentative digestion are microbial in origin, rather than coming from the host animal. Other major differences between fermentative and glandular digestion involve the rate of reactions and the extent of alteration of the substrate molecules. In general, fermentative

digestion is much slower than glandular digestion, and the substrates are altered to a much greater degree.

The Sites of Fermentative Digestion Must Be Conducive to Microbial Growth

Fermentative digestion occurs in specialized compartments positioned either before or after the stomach and small intestine. Fermentative compartments positioned before the stomach are called *forestomachs* and are most highly developed in the ruminants and camelids. The size and development of the forestomach fermentation compartments vary greatly among species; many species have distinct forestomachs that are less developed than those of ruminants. Some species, including the horse and rat, have no anatomically distinct forestomach; however, some fermentative digestion may occur in a nonglandular portion of the proximal stomach.

Fermentation compartments positioned distal to the small intestine are the cecum and colon, often collectively called the *hindgut*. As with the forestomach, great anatomical differences exist in the hindgut of various species. This variation can be so extensive that the cecum and colon may appear to be functionally different organs in different species; however, when the variations are evaluated critically, important similarities can be seen in hindgut function among species.

The forestomach and hindgut can support fermentative digestion because their pH, moisture, ionic strength, and oxidation-reduction conditions are maintained in a range compatible for the growth of suitable microbes. In addition, the flow of ingesta through these areas is comparatively slow, allowing microbes time to maintain their population size. The importance of these factors can be illustrated through comparison of the forestomach and colon to the stomach and small intestine. In the stomach, bacterial numbers are kept low by the acid pH, whereas in the small intestine, bacterial numbers are kept in check by the constant flushing action of ingesta and secretions. In contrast, the pH in the forestomach and large colon is close to neutral, and the flow rate is comparatively slow.

In general, the fermentative patterns of the hindgut appear to be similar to those of the forestomach, although forestomach fermentation, especially that of the rumen, appears to be the better studied of the two. The following discussion focuses on rumen digestion but includes comments on hindgut digestion. Digestion in the equine cecum and colon is discussed at the end of the chapter.

Microbial Ecosystem of Fermentative Digestion

The Microbes Responsible for Fermentative Digestion Include Bacteria, Fungi, and Protozoa

The bacterial population associated with fermentative digestion is vast, with at least 28 functionally important species occurring in the rumen. Box 31.1 lists some of the major species found in the

• BOX 31.1 Grouping of Rumen Bacterial Species According to Type of Substrates Fermented

Major Cellulolytic Species
Bacteroides succinogenes
Ruminococcus flavefaciens
Ruminococcus albus
Butyrivibrio fibrisolvens

Major Hemicellulolytic Species
Butyrivibrio fibrisolvens
Bacteroides ruminicola
Ruminococcus species

Major Pectinolytic Species
Butyrivibrio fibrisolvens
Bacteroides ruminicola
Lachnospira multiparus
Succinivibrio dextrinosolvens
Treponema bryantii
Streptococcus bovis

Major Amylolytic Species
Bacteroides amylophilus
Streptococcus bovis
Succinimonas amylolytica
Bacteroides ruminicola

Major Ureolytic Species
Succinivibrio dextrinosolvens
Selenomonas species
Bacteroides ruminicola
Ruminococcus bromii
Butyrivibrio species
Treponema species

Major Methane-Producing Species
Methanobrevibacter ruminantium
Methanobacterium formicicum
Methanomicrobium mobile

Major Sugar-Utilizing Species
Treponema bryantii
Lactobacillus vitulinus
Lactobacillus ruminis

Major Acid-Utilizing Species
Megasphaera elsdenii
Selenomonas ruminantium

Major Proteolytic Species
Bacteroides amylophilus
Bacteroides ruminicola
Butyrivibrio fibrisolvens
Streptococcus bovis

Major Ammonia-Producing Species
Bacteroides ruminicola
Megasphaera elsdenii
Selenomonas ruminantium

Major Lipid-Utilizing Species
Anaerovibrio lipolytica
Butyrivibrio fibrisolvens
Treponema bryantii
Eubacterium species
Fusocillus species
Micrococcus species

Yokoyama MT and Johnson KA, Microbiology of the rumen and intestine, in Church DC. The Ruminant Animal: Digestive Physiology and Nutrition, chapter 7, page 125. Englewood Cliffs, NJ: Prentice-Hall, 1988.

rumen and their preferred substrates. Total bacterial numbers in the forestomach or hindgut normally range from 10^{10} to 10^{11} cells per gram of ingesta. Most of these bacteria are strict anaerobes that cannot survive in the presence of oxygen, although facultative organisms are also present. In the rumen, fungi are present, and research suggests that fungi may play an important role in the digestion of plant cell walls.

There is also a large population of protozoa in the rumen and in the cecum and colon. Protozoal numbers average about 10^5 to 10^6 cells per gram of rumen contents. Although this number is considerably smaller than the number of bacteria, the relatively larger size of the individual protozoa compared with bacteria results in a total rumen protozoal cell mass approximately equal to the bacterial cell mass, under most dietary conditions. Most of the rumen protozoa are ciliated and belong to the genus *Isotricha* or *Entodinium,* although flagellate species are also present, especially in young ruminants. As with the other organisms of the rumen, the protozoa are anaerobic.

The digestive abilities, or capacities, of protozoa and bacteria are similar; thus either type of organism can perform most of the fermentative functions of the rumen. Protozoa ingest large numbers of bacteria and hold rumen bacterial numbers in check. However, none of the actions of protozoa appears essential to rumen function because ruminants can survive well without protozoa. Thus the role of protozoa in the total ecological picture of the rumen is uncertain. One potentially important function of protozoa involves their ability to slow down the digestion of rapidly fermentable substrates, such as starch and some proteins. Protozoa are capable of ingesting particles of starch and protein, and storing them in their bodies, protected from bacterial action. The starch and protein remain engulfed until digested by the protozoa, or until the protozoa die or are swept from the rumen into the lower digestive tract. Thus protozoa may have the effect of delaying or prolonging the digestion of these substrates. Especially in the case of starch, this protozoal effect may be beneficial to the host through modulation or delay of the digestion of rapidly fermentable substrate.

Cooperation and Interplay Among the Many Species of Microbes Give Rise to a Complex Ecosystem in the Forestomach and Hindgut

The digestive process in the rumen or colon involves the interplay among the many species of bacteria and other microbes. The ecosystem of fermentative digestion is extremely complex, with the waste products of one microbial species serving as substrate for another. For example, *Ruminococcus albus* and *Bacteroides ruminicola* appear to exist synergistically. *R. albus* digests cellulose (is *cellulolytic*) but cannot digest protein. *B. ruminicola,* on the other hand, can digest protein but cannot digest cellulose. When the microbes are grown together, cellulose digestion by *R. albus* provides hexoses for the energy needs of *B. ruminicola,* and protein digestion by *B. ruminicola* provides ammonia and branch-chain fatty acids for the growth needs of *R. albus.*

In addition to substrate needs, growth factor needs are also supplied synergistically within the rumen ecosystem. For example, B vitamins are necessary for the growth of several rumen microbes, but these nutrients are generally not necessary in ruminant diets. The synergistic effect of B vitamins results from cross-feeding between species of those microbes that produce various B vitamins and those microbes that require them.

Despite tremendous ecological complexity, however, the entire pattern of fermentation may be viewed as a holistic process, without consideration of the roles and interactions of individual microbial species. Fermentative digestion is examined here from this viewpoint, with the actions of the entire rumen biomass considered as an overall digestive process, irrespective of the specific needs and actions of individual microbial species.

Substrates and Products of Fermentative Digestion

Plant Cell Walls Are Important Substrates for Fermentative Digestion and Significant Nutrient Sources for Many Species

Forages, or the foliage of plants, are both the major feedstuff of large herbivores and an important substrate for fermentative digestion. Some appreciation of the physical and chemical nature of plants is important to an understanding of the fermentative digestion of forages. This understanding may be aided by a brief comparison of plant and animal tissue structure.

At the cellular level, a major difference between plants and animals is the existence of a *cell wall* in plants. The cell wall is a complex of various carbohydrate molecules. The structural parts of plants, the leaves and stems, contain a large portion of cell-wall material. This material gives the plants their rigid framework and protects them from weather and other elements during growth. The cell-wall structure of plants can be roughly compared with the connective tissue structure of animals. Long, fiber-like molecules of *cellulose* have a strength-giving role similar to that of collagen, whereas *hemicellulose, pectin,* and *lignin* cement the cellulose together, much as hyaluronic acid and chondroitin sulfate do in animal connective tissue. With the exception of lignin, all these cell-wall molecules are carbohydrate.

Cellulose is composed of nonbranching chains of glucose monomers joined by β[1-4] glycosidic linkages, in contrast to the α[1-4] linkages in starch. Pectin and hemicellulose are chemically more heterogeneous than cellulose, being composed of various proportions of several sugars and sugar acids. None of the cell-wall materials is subject to hydrolytic digestion by mammalian glandular digestive enzymes. However, cellulose, hemicellulose, and pectin are subject to the hydrolytic action of a complex of microbial enzymes known as *cellulase.* This enzyme system releases monosaccharides and oligosaccharides from the complex carbohydrates of cell walls, but the released saccharides are not directly available for absorption by the animal. Rather, they are further metabolized by the microbes, as discussed later.

Lignin, a heterogeneous group of phenolic chemicals, is resistant to the action of either mammalian or microbial enzymes, and only a small portion of lignin is digested by either process. Lignin is important not only because it is indigestible itself, but also because it tends to encase the cell-wall carbohydrates, reducing their digestibility by protecting them from the action of bacterial cellulase. The lignin concentration of plants increases with age and ambient temperature; thus young, cool-season plants are more digestible than mature plants grown in hot weather.

Nutrients Other Than Cell Walls Are Also Subject to Fermentative Digestion

The fermentative digestion of plant cell-wall material and its importance to herbivore digestion are well known. In addition, however, essentially all protein and carbohydrate nutrients that

can provide substrate for energy and growth in mammals can also support the similar needs of microbes. Therefore almost all dietary proteins and carbohydrates are potentially subject to fermentative digestion. This fact is especially important in ruminants, in which food is exposed to fermentative digestion in the forestomach before its arrival at sites of glandular digestion. This temporal arrangement leads to the fermentative digestion of many nutrients that would otherwise have been available to the animal through glandular digestion. Thus forestomach fermentative digestion, which provides for the efficient use of plant cell walls, can potentially lead to the inefficient use of other nutrients because of microbial alteration.

Anaerobic Conditions in the Rumen Result in Metabolic Activities Leading to the Production of Volatile Fatty Acids

When carbohydrate material enters the rumen or colon, hydrolytic microbial enzymes attack it. In the case of insoluble carbohydrates, this attack requires the physical attachment of bacteria to the surface of the plant particle, with the enzymes themselves part of the surface coating of the bacteria. Enzymatic action liberates glucose, other monosaccharides, and short-chain polysaccharides into the fluid phase, outside the microbial cell bodies. Although free in solution, these products of microbial enzyme action do not become immediately available to the host animal; rather, they are quickly subjected to further metabolism by the microbial mass. Glucose and other sugars are absorbed into the cell bodies of the microbes.

Within the microbial cells, glucose enters the glycolytic, or *Embden-Meyerhof,* pathway. This is the same glycolytic pathway that exists in mammalian cells, and as in mammalian tissues, catabolism of glucose through this pathway yields two molecules of pyruvate for each molecule of glucose metabolized. In the process, two molecules of oxidized nicotinamide adenine dinucleotide (NAD) are reduced to NAD hydrogen (NADH), and two molecules of adenosine triphosphate (ATP) are formed from adenosine diphosphate (ADP). The potential energy represented by the ATP formed in this reaction is not directly available to the host animal but is the major source of energy for maintenance and growth of microbes.

If fermentative digestion were to occur under aerobic conditions, which it does not, the pyruvate produced by the glycolytic process would enter the citric acid (Krebs) cycle and would be metabolized to carbon dioxide and water, as occurs under the aerobic conditions in mammalian cells. Furthermore, in an aerobic system, the NADH produced would be oxidized in the cytochrome oxidase system with additional production of ATP and the regeneration of NAD. However, fermentative digestion is not an aerobic system; on the contrary, it proceeds in a reductive, highly anaerobic environment. Therefore a different mechanism must be provided for the oxidation of NADH and other reduced cofactors. If such a mechanism were not available, all the oxidized cofactors present would soon be reduced, and metabolism would come to a halt. Because no atmospheric oxygen is available, some other compound must serve as an "electron sink" for the oxidation of enzyme cofactors.

In fermentative digestion, pyruvate can act as an electron sink, being further reduced to provide for regeneration of NAD and the general removal of excess electrons, with an additional yield of ATP. Also, carbon dioxide can be reduced to methane, accepting electrons for the regeneration of NAD. Fig. 31.1 illustrates the metabolic pathways of these reactions. These pathways lead to the major end products of the fermentative digestion of carbohydrate, the *volatile fatty acids* (VFAs). The primary VFAs are *acetic acid, propionic acid,* and *butyric acid;* the VFAs are often referred to as their dissociated ions: acetate, propionate, and butyrate, respectively. Other quantitatively minor but metabolically important VFAs are valeric acid, isovaleric acid, isobutyric acid, and 2-methylbutyric acid. Fig. 31.2 shows the chemical structures of the VFAs.

Production of propionic acid from pyruvate results in the efficient regeneration of NAD with no net production of NADH. In fact, production of available oxygen by the *randomizing branch* of the propionic acid pathway leads to oxidation of excess NADH originating from the acetic or butyric acid pathways (see Fig. 31.1). The production of acetic acid leads to the efficient generation of ATP but, in contrast to the production of propionic acid, does not result in the regeneration of NAD from NADH. In the acetic acid pathway, excess NADH is produced. In this case, NAD is regenerated by the formation of free hydrogen, which is subsequently used to reduce carbon dioxide to methane and water (see Fig. 31.1, *lower portion*).

Thus a direct relationship exists between acetic acid production and methane production; as the amount of pyruvate entering the acetic acid pathway increases, there must be a concomitant rise in methane production. Likewise, a reciprocal relationship exists between methane production and propionic acid production; as pyruvate is diverted to propionic acid production, there is less need for methane synthesis. These relationships are shown in the stoichiometric equations of Box 31.2. These reactions do not, however, fully describe the flow of hydrogen, or reducing substances, in rumen or colonic metabolism. The chemical reactions of fermentation are extremely complex and interdependent, and NADH can donate its electrons to reactions other than those described in Box 31.2, such as the synthesis of microbial protein and the saturation of unsaturated fatty acids.

In the rumen, methanogenic bacteria, such as *Methanobacterium ruminantium,* facilitate methane production. This fragile bacterium is

• BOX 31.2 Theoretical Stoichiometric Carbon-Hydrogen Balance Equations Describing Conversion of Glucose in Rumen

Case 1

$\text{Glucose} \rightarrow 2\ \text{acetate} + 2\ CO_2 + 8\ H$
$\text{Glucose} \rightarrow \text{Butyrate} + 2\ CO_2 + 4\ H$
$\text{Glucose} + 4\ H \rightarrow 2\ \text{propionate} + 2\ H_2O$
$CO_2 + 8\ H \rightarrow CH_4 + H_2O$

Net*

$3\ \text{glucose} \rightarrow 2\ \text{acetate} + \text{butyrate} + 2\ \text{propionate} + 3\ CO_2 + CH_4 + 2\ H_2O$

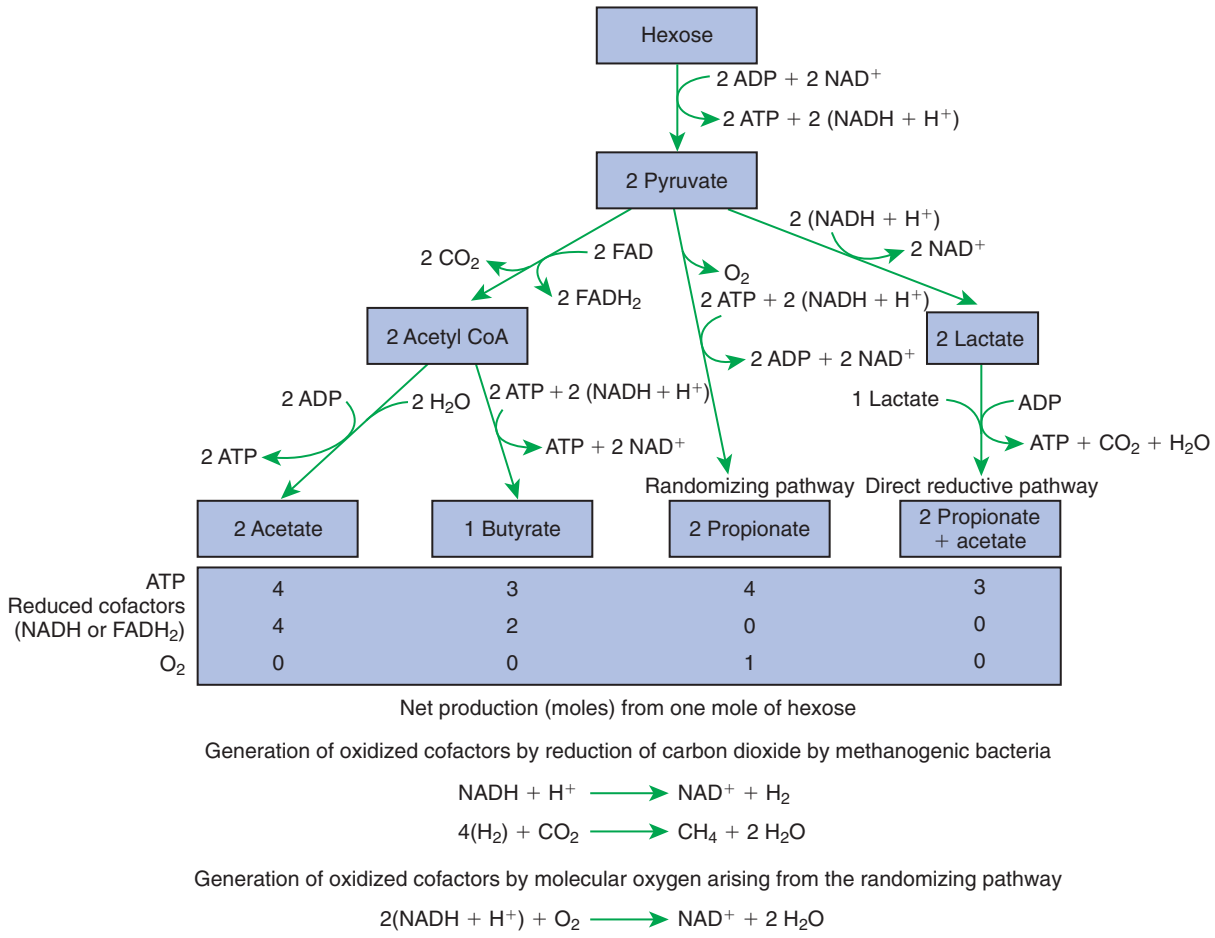
Case 2

$3\ \text{glucose} \rightarrow 6\ \text{butyrate} + 2\ \text{propionate} + CO_2 + CH_4 + 24\ H$
$\text{Glucose} \rightarrow \text{Butyrate} + 2\ CO_2 + 4\ H$
$\text{Glucose} + 4\ H \rightarrow 2\ \text{propionate} + 2\ H_2O$
$3\ CO_2 + 24\ H \rightarrow 3\ CH_4 + 6\ H_2O$

Net*

$5\ \text{glucose} \rightarrow 6\ \text{acetate} + \text{butyrate} + 2\ \text{propionate} + 5\ CO_2 + 3\ CH_4 + 6\ H_2O$

**Note that, in Case 1, the acetate/propionate ratio is 1:1 and the methane/glucose ratio 1:3 ; in Case 2, however, the acetate/propionate ratio is 3:1 and the methane/glucose ratio 3:5.*
From Van Soest PJ. Nutritional Ecology of the Ruminant. *Ithaca, NY: Cornell University Press; 1982.*



• **Fig. 31.1** Pathways of volatile fatty acid (VFA) production by the rumen or colonic biomass. The production of methane is necessary for the production of oxidized cofactors in the pathways leading to acetate and butyrate production, but not in the pathways leading to propionate production. The production of oxygen by the randomizing pathway results in the net production of oxidized cofactors. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; *CO_2,* carbon dioxide; *CoA,* coenzyme A; *FAD,* flavin adenine dinucleotide; *H,* hydrogen; *NAD,* nicotinamide adenine dinucleotide.

sensitive to changing conditions in the rumen. When conditions are unfavorable for the survival of *M. ruminantium,* methane production is reduced, shifting the metabolic pathways toward propionic acid production. Some conditions that suppress methanogenic species are high levels of feed intake, use of finely ground or pelleted feeds, and high-grain or high-starch diets. Under these circumstances, the rate of methane production is reduced, resulting in a lower rate of acetic acid production with a concomitant increase in the propionic acid production rate.

The proportional rates at which acetic acid, propionic acid, and butyric acid are produced are reflected in their relative concentrations in the rumen fluid. The relative concentrations of the VFAs have important nutritional and metabolic consequences, and although seldom measured for medical purposes, VFA concentrations are frequently reported in research literature. Typically, the ruminal acetic/propionic/butyric acid concentration ratio in ruminants ranges from 70:20:10 for animals eating high-forage diets to 60:30:10 for animals eating high-grain diets. One must remember that these values represent relative proportions and not absolute amounts. The total amount of VFA produced with a high-starch diet is usually much higher than that produced with a high-fiber diet, such that total acetic acid production may be higher with a high-starch diet than with a high-fiber diet, even though the acetic acid production relative to the other VFAs may be reduced. Fig. 31.3 illustrates this principle.

Volatile Fatty Acids Are Important Energy Substrates for the Host Animal

One can appreciate the elegance and beauty of the symbiotic relationship represented by fermentative digestion by considering the metabolism of VFAs. These molecules are the end products, indeed, the waste products, of anaerobic microbial metabolism, just as carbon dioxide is the waste product of aerobic metabolism. If the VFAs were allowed to accumulate, they would suppress or alter the fermentative process by lowering the pH of the gut or forestomach. However, the host animal maintains conditions for fermentation both by buffering pH changes and by removing VFAs from the gut by absorption. The benefit derived by the host is from the chemical energy contained in the VFAs. These bacterial "waste products" represent spent compounds within the framework of the anaerobic fermentation system, but they still contain considerable energy that can be derived from aerobic metabolism. In ruminants and other large herbivores, the VFAs are the major

energy fuels, to a large extent serving the role played by glucose in omnivorous monogastric animals. The metabolic fates of the VFAs are discussed further in Chapter 32.

Fermentative Digestion of Protein Results in the Deamination of a Large Portion of Amino Acids

To this point, the discussion of fermentative digestion has centered primarily on carbohydrates, but as previously mentioned, other energy-yielding substrates are subject to microbial attack as well. Proteins are particularly vulnerable because they are composed of carbon compounds that can be further reduced to provide energy for anaerobic microbes. As proteins enter fermentative areas of the gut, they are attacked by extracellular microbial proteases. The majority of these enzymes are "trypsin-like" endopeptidases that form short-chain peptides as end products. These peptides are formed extracellularly and are absorbed into the microbial cell bodies, much as glucose is formed from carbohydrate and then absorbed. Within the microbial cells, the peptides can be used to form microbial protein or can be further degraded for the production of energy through the VFA pathways (Fig. 31.4).

To enter the VFA pathways, the individual amino acids are first deaminated to yield *ammonia* (NH_3) and a carbon skeleton. The carbon structures of many of the amino acids can fit directly into various steps of the pathways leading to the production of the three major VFAs. The three *branch-chain amino acids* (BCAAs) are exceptions, however, and lead to the production of branch-chain VFAs by the following reactions:

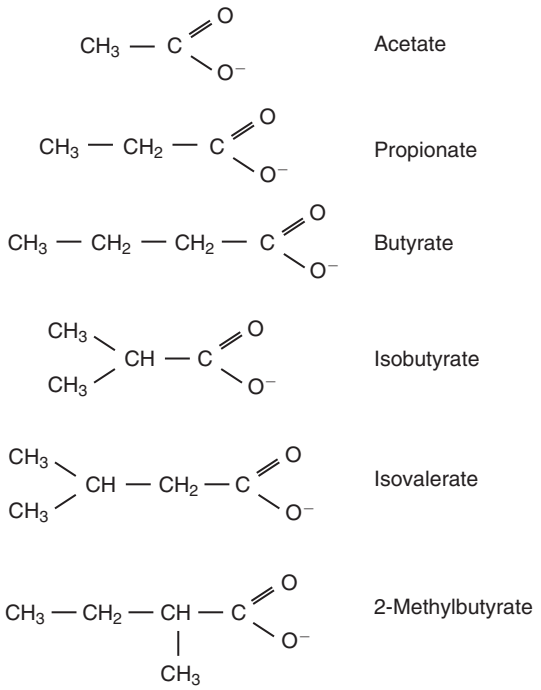
$$\text{Valine} + 2H_2O \rightarrow \text{Isobutyrate} + NH_3 + CO_2$$

$$\text{Leucine} + 2H_2O \rightarrow \text{Isovalerate} + NH_3 + CO_2$$

$$\text{Isoleucine} + 2H_2O \rightarrow \text{2-Methylbutyrate} + NH_3 + CO_2$$

These branch-chain VFAs are important growth factors for several species of bacteria, as described later.

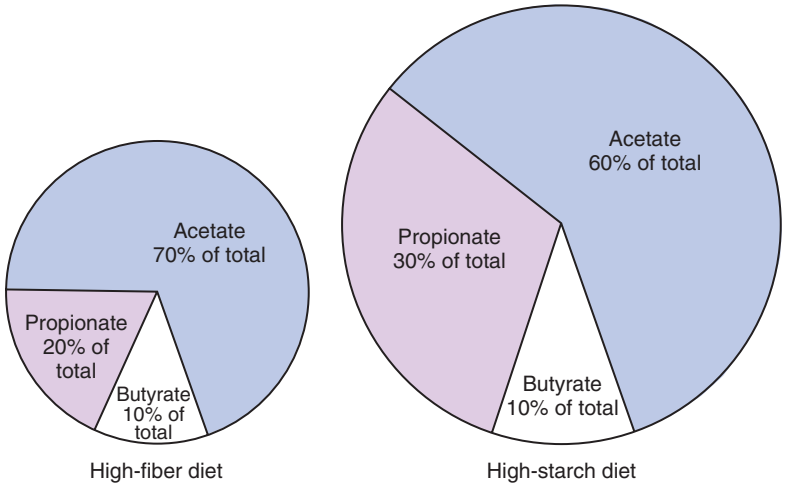
Although many species of rumen microbes appear capable of using preformed amino acids for the synthesis of protein, several species cannot do so. These species must synthesize amino acids from ammonia and the various carbon metabolites of the VFA pathways. For synthesis of the BCAAs, the branch-chain VFAs are required. Among the microbial species that require ammonia and branch-chain fatty acids are some of the important cellulose-digesting bacteria.



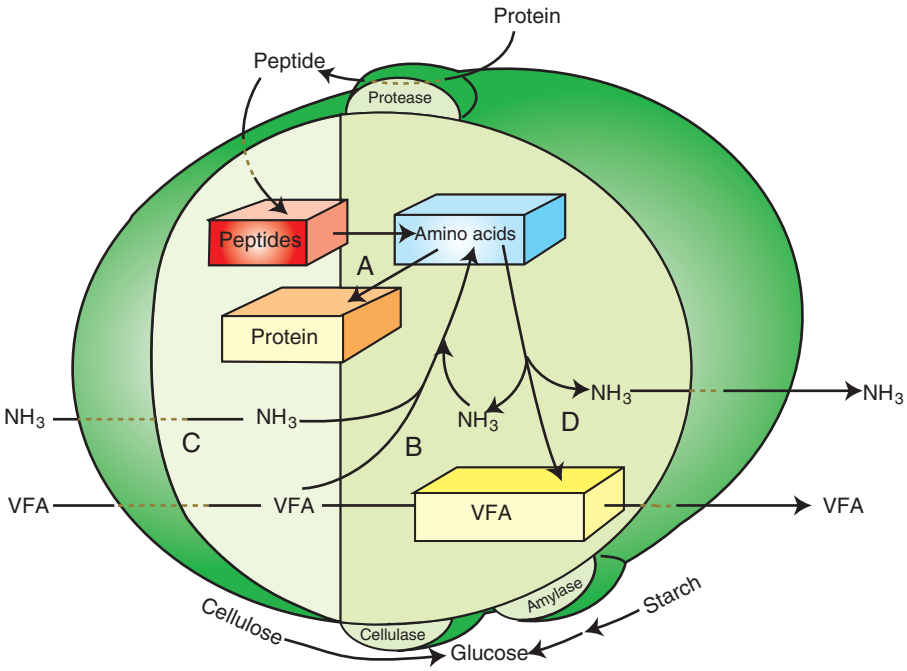
• **Fig. 31.2** Chemical structures of the major volatile fatty acids (VFAs) produced by fermentative digestion.

When Protein and Energy Availability in the Forestomach Are Well Matched, Rapid Microbial Growth and Efficient Protein Utilization Result

A large proportion of dietary amino acids are deaminated in the rumen, making them unavailable for absorption in the small intestine. Therefore ruminant animals depend, to a large extent, on microbial protein as a source of dietary amino acids. Microbial protein reaches the abomasum and small intestine when microbes are washed out of the rumen and into the lower tract. The utilization of dietary protein is optimized in ruminants when the growth rate of the rumen microorganisms is maximal, resulting in maximal delivery of microbial protein to the host animal. These conditions are best met by rapidly growing populations of microbes. The microbial growth rate depends on the supply of nutrients and the rate at which microbes are washed from the rumen. Here we



• **Fig. 31.3** Volatile fatty acid (VFA) production on high-fiber and high-starch diets. Although the percentage of acetate is lower on the high-starch diet, the total amount of acetate produced is greater on the high-starch diet. In contrast, propionate increases in both amount and proportion on the high-starch diet.



• **Fig. 31.4** Protein metabolism by rumen microbes. Protease enzymes on the microbe surfaces generate peptides that are then taken up by many types of organisms. (A) Absorbed peptides contribute to an intracellular pool of amino acids from which microbial proteins are synthesized. (B) Another source of amino acids is from intracellular synthesis using ammonia (NH_3) and volatile fatty acid (*VFA*). Many microbes appear capable of deriving their amino acids from either extracellular peptides or intracellular synthesis; (C) however, several types of bacteria seem incapable of using peptides for an amino acid source and are thus dependent on an extracellular source of ammonia for amino acid synthesis. (D) Amino acids not used for protein synthesis can be metabolized to VFA and ammonia.

consider the effect of nutrient supply on microbial growth rate; factors affecting the rate of microbial removal are discussed later.

The overall reaction in the rumen may be greatly simplified, for the purposes of this discussion, to Eq. 31.1:

(Eq. 31.1)

$$\text{glucose} + \text{peptide} = \text{microbes} + \text{VFA} + NH_3 + CH_4 + CO_2$$

The term "microbes" in this equation represents the mass of microbial cell bodies, which is a major source of dietary protein to the host animal. The faster the microbes multiply, the more protein will be available to the animal. Glucose and peptide are the major substrates that support microbial growth; as the mass of glucose and peptide increases on the left side of the equation, the potential mass of microbial cell bodies increases on the right.

Glucose and peptide represent ruminally available dietary carbohydrate and protein, respectively. In this context, *available* means available to the microbes for fermentation. Dietary carbohydrate or protein that is not susceptible, or accessible, to microbial attack is classified as *unavailable* and is not included in Eq. 1.

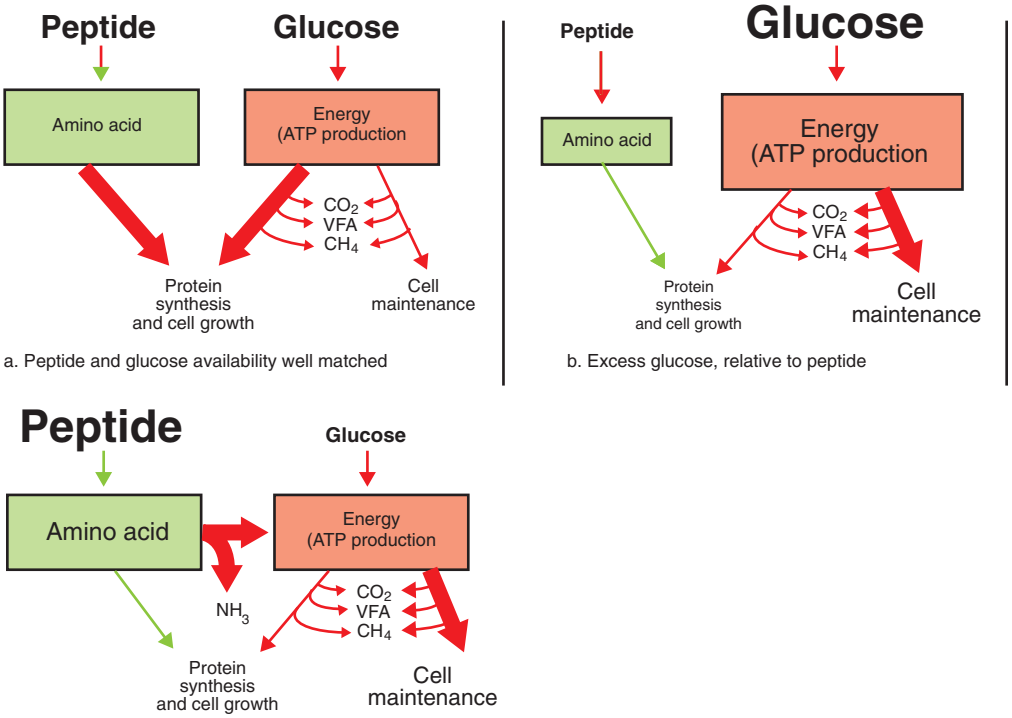
In Eq. 31.1, glucose was chosen to represent carbohydrate and peptide to represent protein because all carbohydrates must be broken down to simple sugars, and proteins to peptides, before entering microbial metabolic pathways. Peptide is the only nitrogen-containing substrate on the left in the equation, but there are two nitrogen-containing products on the right: microbes (representing microbial protein) and NH_3 (ammonia). Both substrates, glucose and peptide, contain carbon, oxygen, and hydrogen, and thus can contribute to the formation of microbial cell carbohydrate, VFA, CH_4, and CO_2.

Eq. 31.1 always balances, but the distribution of products varies according to the relative concentrations of glucose and peptide substrates, as illustrated in Fig. 31.5. For microbial cells to be produced, both energy and nitrogen are required. Energy can come from either glucose or peptide, but nitrogen must come from peptide. When glucose and peptide availability are well matched (Fig. 31.5A), energy for cellular growth comes primarily from glucose, with peptides directed toward microbial protein synthesis. Under these conditions, the products of Eq. 31.1 favor microbial protein synthesis, with little ammonia production. Glucose fermentation provides energy to support the rapid growth of the microbial mass. Peptides are used for microbial protein synthesis, not broken down for energy production, which would result in the production of ammonia. Thus microbial protein synthesis is rapid and ammonia production is low when glucose and peptide availability are well matched.

When the availability of glucose is high relative to peptide (Fig. 31.5B), it represents a mismatch, meaning the diet is not optimally balanced. There is ample energy available to the microbes, but insufficient peptide to support rapid protein synthesis; thus microbial replication is not maximal. In this case, microbial energy utilization becomes inefficient as energy is used for the maintenance of nondividing cells rather than for the energy-requiring synthetic processes of growing cells. The maintenance energy needs of the microbes still drive some fermentation of glucose with moderate VFA production, but production of microbial cells is limited because of lack of nitrogen.

When peptide availability exceeds glucose availability (Fig. 31.5C), it represents another type of mismatch. There is ample peptide potentially to support rapid microbial protein synthesis and growth, but this growth potential cannot be achieved because of insufficient energy supplies. These conditions force the microbes to use peptides to meet energy needs instead of for protein synthesis. Microbial growth rate is low, and VFA production is moderate, because fermentation is driven only by the maintenance energy needs of the microbes. Much of the VFA production comes from the carbon portions of the peptides, whereas the amine groups are shunted to ammonia production; thus the products of Eq. 31.1 favor ammonia.

The relationship between available glucose (carbohydrate) and peptide (or nitrogen) has a tremendous effect on the production of microbial cells and thus a profound effect on the nutrition of the host. This relationship, as illustrated in Fig. 31.5, is quantified by expressing microbial growth in terms of grams of microbial dry



• **Fig. 31.5** The efficiency with which dietary energy is used for protein synthesis in the rumen depends on the balance between energy and nitrogen sources. The proportion of energy used for protein synthesis and cell maintenance (as indicated by the size of the *arrows*) changes in relation to the balance of peptide (nitrogen) and glucose supplies. *ATP,* Adenosine triphosphate; *VFA,* volatile fatty acid.

matter produced per mole of energy-producing substrate used. This value, referred to as *microbial yield,* is usually designated by a capital *Y* subscripted with the abbreviation of the energy substrate to which it is referenced. A convenient but somewhat theoretical substrate with which to reference microbial cell yield is ATP. Microbial yield is then written as $Y_{ATP} = x$, where x is the number of grams of microbial dry matter produced per mole of ATP used. The value of Y_{ATP} varies between about 10 and 20 g of microbes per mole of ATP. Nitrogen availability has an important effect on the Y_{ATP} value. When microbial growth is limited by *low* nitrogen availability, a large portion of available ATP is used for maintenance rather than cell growth; thus the number of cells produced per ATP is small, and the Y_{ATP} value is low.

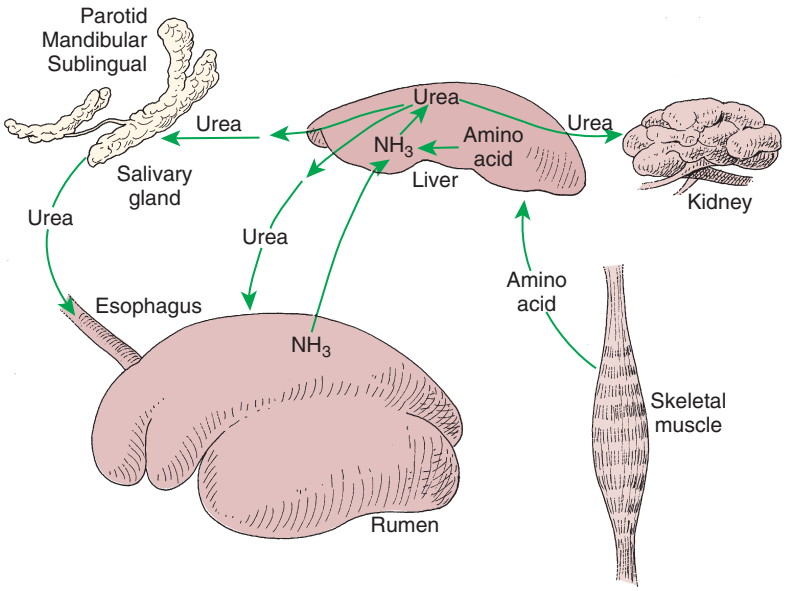
Microbial Protein Can Be Synthesized in the Rumen From Nonprotein Nitrogen Sources

If sufficient carbohydrate is available, most rumen microbes, even those capable of utilizing preformed peptides, can synthesize protein from ammonia (see Fig. 31.4). Thus protein can be produced in the rumen from such nonprotein sources as ammonia, nitrates, and urea. From a nutritional and economic point of view, this capability has been exploited by the inclusion of inexpensive nonprotein nitrogen sources in place of expensive protein in ruminant diets, allowing the microbes to synthesize protein for the amino acid needs of the host. This process also can be exploited physiologically by the recycling of endogenous urea.

Urea, the nitrogenous waste product of protein catabolism, is formed in the liver. In ruminant animals, hepatic urea production is from two sources: (1) nitrogen arising from the deamination of endogenous amino acids and (2) nitrogen absorbed as ammonia from the rumen (Fig. 31.6). Ammonia absorption from the rumen is proportional to the ruminal ammonia production rate, which is subject to the influences of ruminal carbohydrate and protein availability, as discussed earlier. Ammonia, which is toxic at moderate concentrations, is absorbed from the rumen and delivered to the liver through the hepatic-portal blood vascular system. The liver extracts ammonia from the portal blood efficiently; thus little of the potentially toxic ammonia reaches the systemic circulation.

In monogastric animals, urea is excreted from the body almost exclusively by the kidneys. In ruminants, however, urea also may be excreted into the rumen (see Fig. 31.6). Such excretion can occur by direct absorption of urea into the rumen from the blood or by excretion of urea into saliva. In either case, the urea reaches the rumen, where it is quickly transformed to ammonia, and enters the general pool of rumen nitrogen from which microbial proteins are synthesized.

The direction of nonprotein nitrogen flow, either into the rumen as urea or out of the rumen as ammonia, depends on rumen ammonia concentrations. During times of high nitrogen availability in the rumen, relative to carbohydrate availability, this system results in high blood urea concentrations and the extensive loss of precious nitrogen through urinary excretion, making ruminants nutritionally inefficient under these dietary conditions. However, during times of high carbohydrate availability relative to nitrogen availability, the major flow of urea nitrogen is from the blood into the rumen. Under these circumstances, in which ruminal ammonia concentrations are low, most of the blood urea is from endogenous protein catabolism. A portion of this urea, which in monogastric animals would be unavailable for protein synthesis, is excreted into the rumen, where it can be resynthesized into protein that will contribute eventually to the amino acid needs of the host. Thus under conditions of low dietary protein, ruminants are efficient conservers of nitrogen.



• **Fig. 31.6** Interorgan nitrogen cycling in ruminants. The diagram shows the effects of rumen ammonia (NH_3) concentration on the formation and utilization of urea. When rumen ammonia concentrations are high, the net movement of nonprotein nitrogen is toward the liver, resulting in high urea production rates and poor nitrogen conservation. When rumen ammonia concentrations are low, the net movement of nonprotein nitrogen is from liver to rumen, resulting in protein production from endogenous urea.

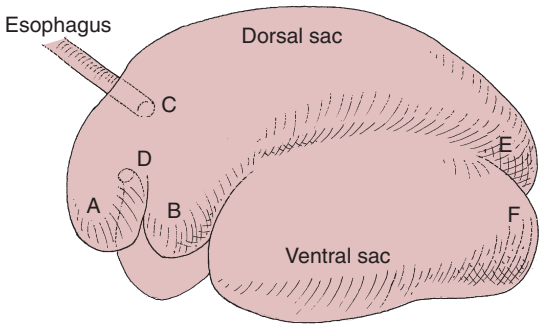
Reticulorumen Motility and Maintenance of the Rumen Environment

The Physiological Functions of the Reticulorumen Maintain an Environment Favorable to Fermentation Patterns That Are Beneficial to the Host

The host animal has no direct control over the metabolism of the microbes in its gut. However, important physiological factors influence the gastrointestinal (GI) fermentation process. For the host to ensure that the proper type of fermentation patterns occur, it must maintain within the rumen (or colon) conditions that promote the growth and favorable metabolic patterns of the most beneficial bacteria and other microbes. The following requirements must be met by the host for proper fermentation to occur:

1. Substrate for fermentation must be supplied.
2. Temperature must be maintained at or near 37°C.
3. Ionic strength (osmolality) of the rumen fluid must be kept within an optimal range (near 300 mOsm).
4. A negative oxidation-reduction potential must be maintained (−250 to −450 mV).
5. Indigestible waste (solid material) must be removed.
6. The rate of removal of microbes must be compatible with the regeneration times of the most favorable microbes.
7. Acid products of anaerobic fermentation (VFAs) must be buffered or removed.

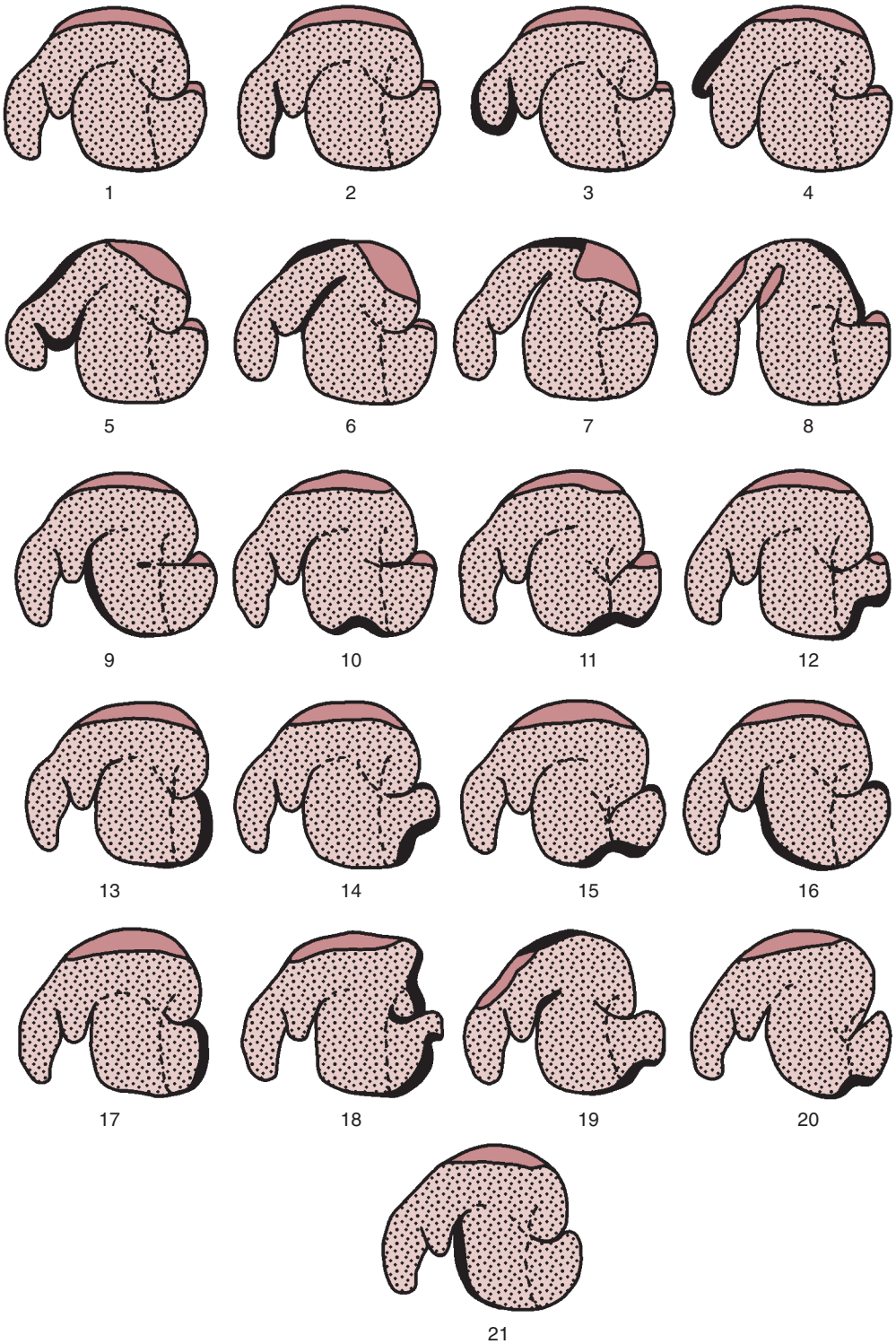
The first of these requisites, delivery of substrate, requires only eating; others (e.g., temperature, ionic strength) are met by the same homeostatic mechanisms that maintain these physiological conditions within the host body in general. Maintenance of an appropriate oxidation-reduction potential requires only that oxygen be kept away from the fermentation site. The remaining requisites for fermentation, however, have necessitated the development of special physiological functions associated with the forestomach (or hindgut). These specialized functions include the motility patterns characteristic of the reticulorumen, the direct absorption of VFA, and the production of copious amounts of saliva.



• **Fig. 31.7** Rumen anatomy. (A) Reticulum; (B) cranial sac; (C) cardia; (D) reticulo-omasal orifice; (E) caudal-dorsal blind sac; (F) caudal-ventral blind sac.

Rumen Fermentation Is Maintained by Selectively Retaining Actively Fermenting Material While Allowing Unfermentable Residue to Pass on to the Lower Digestive Tract

The walls of the reticulorumen are muscular, possess an extensive intrinsic nervous system, and are capable of highly complex and coordinated motility patterns. These motility patterns are necessary for the critical function of the rumen, which is the selective retention of actively fermenting material accompanied by the simultaneous release of unfermentable residue. An understanding of reticulorumen anatomy is necessary to appreciate the effects of reticulorumen motility patterns. Fig. 31.7 illustrates the division of the reticulorumen into compartments, or *sacs.* These divisions are created by muscular pillars that project into the lumen of the organ. The reticular fold and rumen pillars, in addition to the walls themselves, are motile. During reticulorumen contractions, the pillars alternately elevate and relax, either accentuating or reducing the divisions within the lumen of the reticulorumen. Students accustomed to studying the reticulorumens of embalmed specimens may find it difficult to visualize the extent of rumen movement. At times during contractions, the excursions of the walls and pillars are so



• **Fig. 31.8** *See legend on opposite page.*

great that the total shape of the reticulorumen is distorted; sacs and compartments are almost obliterated, and pillars elevate to the extent that compartmental divisions become nearly complete. When the magnitude of these contractions is recognized, it is not difficult to appreciate the tremendous effect that reticulorumen motility has on the flow of rumen ingesta.

Two patterns of reticulorumen motility are generally described: primary (or mixing) contractions and secondary (or *eructation*) contractions. The complexity, sequential pattern, and highly coordinated nature of these motility patterns are illustrated and described in Fig. 31.8. More important than knowing the exact sequence of contractions, however, is understanding the influence this pattern has on the flow of ingesta through the rumen. To illustrate the primary pattern of motility and describe its effect on ingesta flow, consider the path of a single bolus of feed material as it passes through the rumen. Assume this material is forage such

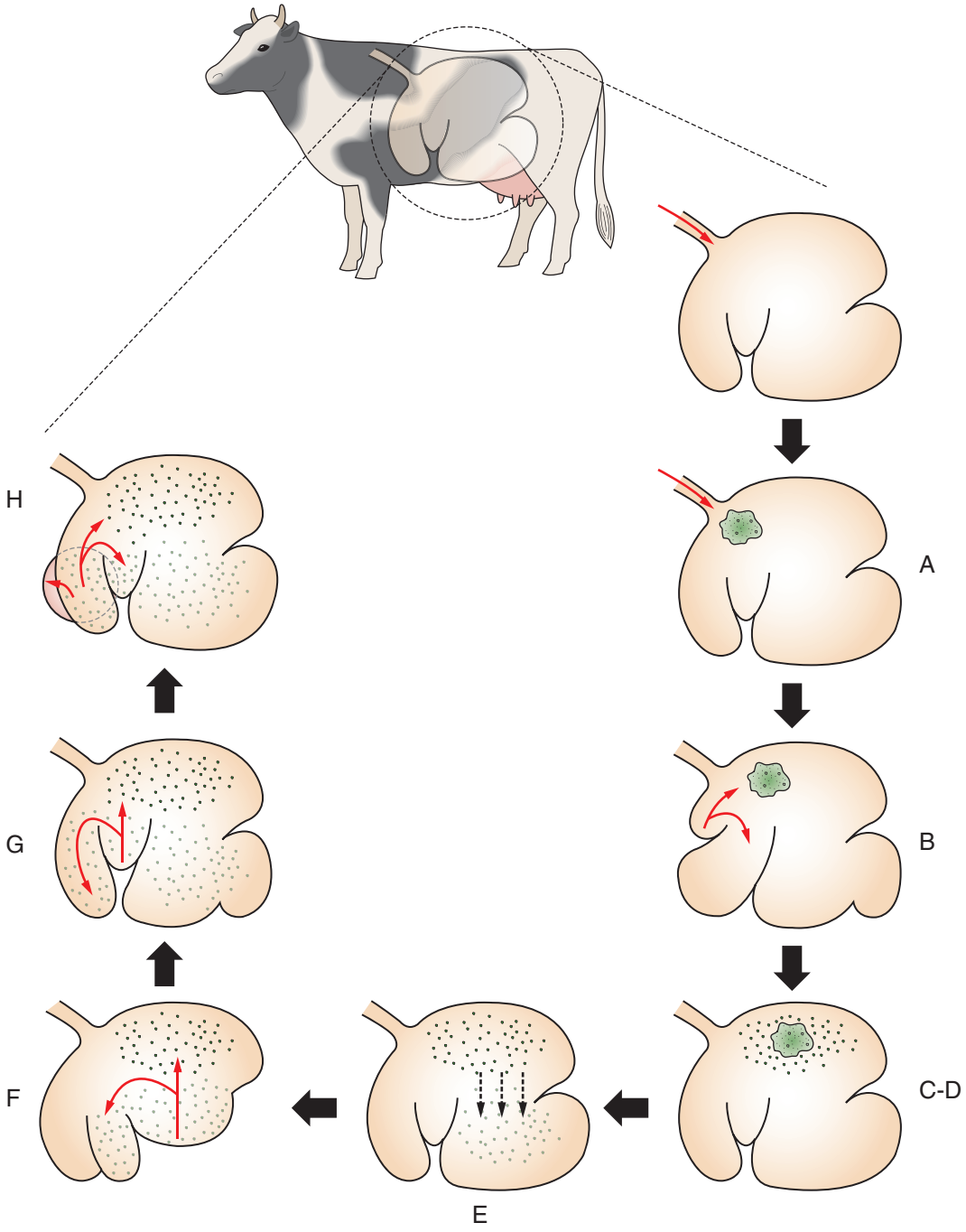
• **Fig. 31.8** Contraction sequence in the reticulorumen. These drawings were derived by taking tracings directly from radiographs. The *open regions* represent the rumen gas cap (zone), whereas the stippled region represents ingesta. The *heavy lines* indicate portions of the wall that are actively contracting. Drawings *1* through *16* represent the sequence of events in a primary contraction in a normally fed sheep. Drawings *17* through *21* represent the sequence of events in a secondary or eructation contraction. *1,* Resting stage. *2,* Initiation of sequence with elevation of reticuloruminal fold. *3,* End of first phase of reticular contraction. *4,* End of second phase of reticular contraction; note dilation of cranial sac. *5* to *7,* Contraction of cranial sac followed by contraction of cranial pillar and dorsal sac. *8,* Contraction of caudal-dorsal blind sac and caudal pillar, causing cranial displacement of gas cap toward reticulum, under cranial pillar, and into caudal-ventral blind sac. *9,* Contraction of longitudinal pillar and cranial ventral rumen; in anorectic sheep, the sequence frequently ceases at this point, and the occurrence of the remaining steps in the sequence varies according to the degree of filling of the reticulorumen. *10* to *12,* Wave of contraction migrating caudally onto the caudal-ventral blind sac, associated with a ventral displacement of the caudal pillar. *13,* Contraction of the pole of the caudal-ventral blind sac displacing gas cap around the caudal pillar. *14* to *16,* Cranial migration of contraction if no secondary contraction sequence occurs. *17,* When a secondary contraction follows a primary, the terminal contraction of the caudal-ventral blind sac may be maintained over a prolonged period or may be repeated simultaneously with a second contraction of the caudal pillar. *18,* Contraction of caudal pillar and dorsal blind sac start to push gas cap cranially; contraction starts to move cranially across caudal-ventral blind sac. *19,* Contraction has moved rapidly across dorsal rumen, and cranial pillar has moved for the second time; eructation, if it occurs, occurs at this point. *20* and *21,* Contraction migrates cranially onto ventral rumen, causing contraction of ventral coronary pillars and second ventral displacement of the caudal pillar; cycle terminates with a contraction of the cranial ventral rumen. (From Ryckebusch Y, Thivend P. *Digestive Physiology and Metabolism in Ruminants.* Westport, Conn: AVI Publishing; 1980.)

as grass, hay, or ensilage. The animal chews the feed to create an initial reduction in particle size and to form it into a bolus by mixing it with saliva. The swallowed bolus enters the rumen at the cardia, which is at the dorsal portion of the reticulum near the junction of the reticulum and the cranial ventral sac (Fig. 31.9A). The viscous nature of saliva causes the feed material to initially remain in a bolus as it enters the rumen. Air bubbles trapped in the bolus give it a relatively low specific gravity, compared with surrounding ingesta, so it remains suspended in the area near the cardia (Fig. 31.9B). As a primary rumen contraction begins, there is a biphasic, or double contraction of the reticulum. The first of these contractions is relatively weak. The second follows very quickly after the first and is extremely forceful, nearly obliterating the lumen of the reticulum and creating a strong and rapid flow of fluid ingesta. Material in the dorsal portion of the reticulum, including the recently swallowed bolus of feed, is washed back into the rumen by this flow of liquid ingesta (Fig. 31.9C). A caudal-moving contraction of the dorsal sac follows the reticular contraction, continuing to move the recently swallowed bolus and other material further back into the dorsal sac. A subsequent cranial-moving contraction of the dorsal sac serves to mix the dorsal sac ingesta, creating a mass of tangled forage fibers representing a large collection of relatively recently swallowed material (Fig. 31.9D).

This mass of material tends to remain in the dorsal rumen because of buoyancy. Microbial action creates small bubbles of gas that adhere to the plant material accounting for its buoyancy and giving it a low *functional specific gravity.* As time passes, however, microbial action tends to cause the material to break apart. The small particles that remain become less buoyant because there is less fermentable substrate from which microbes can generate gas bubbles. The functional specific gravity of these particles increases as a result of the lack of gas bubbles and the material begins to sink in the rumen as small particles (Fig. 31.9E). The material in the dorsal sac, especially in animals on a diet with a large portion of forage, is typically a relatively solid mass of tangled forage fibers whereas the material in the ventral sac is a suspension of small particles that is much more fluid and water-like. The interface between the two areas is not distinct and is composed of a "mushy" or slurry-like transition zone, a zone of intermediate-sized particles with intermediate functional specific gravity. Contractions of the ventral sac tend to push the more buoyant material back up into the dorsal sac, while allowing the smallest and least buoyant particles to spill over the cranial ruminal pillar into the cranial ventral sac (Fig. 31.9F).

Subsequent to the contractions of the ventral sac is a contraction of the cranial ventral sac, which further separates material based on functional specific gravity and results in small particles with the highest functional specific gravity flowing back into the reticulum (Fig. 31.9G). At this point, material has made a complete circuit of the rumen; entering at the cardia; passing through the dorsal, ventral, and cranial ventral sacs; and back to the reticulum. As the reticulum contracts at the beginning of a primary cycle, the reticulo-omasal orifice relaxes and dense (high functional specific gravity) material near the bottom of the reticulum is forced through the opening and into the omasum (Fig. 31.9H).

The overall effect of the transit of feed through the rumen, particularly forage feed, is a reduction in particle size associated with the loss of fermentable portions of the plant structure. Long forage material is reduced by initial mastication to particles of 1 to 2 cm or shorter. Most of the material in the dorsal rumen is of similar particle size. Particle size diminishes in the more ventral portions of the rumen. Most particles that move through the reticulo-omasal orifice are 2 to 3 mm long. The selection of small particles for passage into the omasum occurs even though the reticulo-omasal orifice, when dilated for food passage, is probably



• **Fig. 31.9** The sequential figures illustrate the pattern of flow of feed material through the reticulorumen from its arrival at the cardia (*A*) to its exit via the reticulo-omasal orifice (*H*). Details of the sequential flow pattern are given in the text.

about 2 cm in diameter, indicating that size discrimination is not based on the size of the reticulo-omasal orifice.

The previous discussion applies to the effects of the primary (or mixing) pattern of motility. *Secondary* (or *eructation*) *contractions* occur as an added sequence of events at the end of a primary sequence of contractions. In general, they occur in association with every second or third set of primary contractions. Secondary contractions consist of a cranial-moving wave that starts in the caudal-dorsal blind sac and continues over the dorsal sac (see Fig. 31.8, segments 17 through 21). The function of the secondary contraction is to force gas toward the cranial portion of the rumen. The pattern begins with a contraction of the caudoventral blind sac, expressing trapped gas in that compartment into the dorsal sac. The secondary contraction continues with a forward-moving contraction of the dorsal sac that moves gas toward the cardia, while cranial sac relaxation and cranial pillar elevation allows liquid ingesta to move away from the cardia so that gas can enter the esophagus and be eructated. Secondary contractions are important

because large amounts of gas, primarily CO_2 and CH_4, are formed during fermentation and must be removed rapidly to prevent distention of the rumen.

In general, one to three reticulorumen contractions occur per minute. Contractions occur most frequently during eating and disappear entirely during deep sleep. The rate and strength of contractions depend on the character of the diet: coarse, fibrous feeds stimulate the most frequent and strongest contractions. Secondary contractions usually occur in association with half the primary contractions, although this relationship is variable and depends on the rate of gas formation.

Digestibility and Physical Characteristics of Feed Have Important Influences on Both the Rate of Particle Passage From the Rumen and the Rate of Feed Intake

As indicated earlier, feed does not leave the rumen until it is broken down into small particles. Microbial action and remastication (as discussed later) are primarily responsible for particle size reduction in the rumen, and the rate of breakdown of fiber is chiefly a function of its digestibility. Poorly digestible fiber takes longer to be broken down sufficiently to sink into the ventral sac, compared with fiber of greater digestibility. This means that poorly digestible fiber remains in the rumen longer than fiber of greater digestibility. Because there are fixed limits to the volume of the rumen, the rate of feed intake cannot exceed the rate of ingesta outflow; therefore intake of poorly digestible feeds is always less than intake of highly digestible feeds.

Feed preparation can influence this relationship. Chopping or grinding of poorly digestible forages increases their rate of passage from the rumen because less particle size reduction is necessary before pieces can pass into the omasum. Chopping or grinding also usually increases the amount of material that an animal can eat because the rumen throughput is increased. Often, however, digestibility is decreased by chopping or grinding of forages, because the duration of exposure to microbial action is reduced as a result of rapid passage of feed through the rumen. Thus physical form (length) and digestibility each have an effect on rate of passage from the rumen and also on feed intake. In general, forage material of relatively high digestibility has a rumen half-life of approximately 30 hours, whereas poorly digestible material has a half-life of up to 50 hours.

Rumination, or Cud Chewing, Has an Important Effect on the Reduction of Particle Size and the Movement of Solid Material Through the Rumen

Rumination is the act of remasticating rumen ingesta. The initial act of rumination is regurgitation, which occurs just before the initiation of a primary rumen contraction. When regurgitation occurs, there is an extra contraction of the reticulum, which takes place just before the regular biphasic reticular contraction that initiates the primary cycle. Simultaneous with the extra reticular contraction, the cardia relaxes, and there is an inspiratory excursion of the ribs with the glottis closed. The latter action creates a negative pressure within the thorax, favoring the movement of food into the esophagus. When food enters the esophagus, a reverse peristaltic wave propels the material cranially into the mouth. As soon as the food bolus reaches the mouth, excess water is expressed by action of the tongue, the water is swallowed, and remastication of the material begins. The duration of remastication depends on the character of the diet, with coarse material requiring more time for remastication than finely ground or highly digestible feeds.

Regurgitated material comes from the dorsal portion of the reticulum, where particle size and functional specific gravity are intermediate. Thus the ingesta selected for remastication are not the coarsest material in the rumen but rather the material that has already been through the digestive actions of the dorsal sac. This appears to be an efficient system in which some of the structural material of the plant is softened by soaking and removed or weakened by microbial action during the initial phases of digestion in the dorsal sac. The partially fermented material is then subjected to remastication, causing further comminution and exposing additional fermentable substrate that may not have been directly exposed to previous microbial action.

Rumination may also aid the particle separation process: as the regurgitated bolus reaches the mouth, it is squeezed by the tongue and cheeks before mastication begins. Water and small particles are expressed from the bolus by this squeezing action and are swallowed before mastication of the remaining bolus. This action tends to separate small particles from large particles. The small particles, when reswallowed, sink into the reticulum where they are subject to passage to the omasum whereas the larger particles, when swallowed after remastication, are ejected back into the more cranial portions of the rumen.

Rumination occurs when the animal is not actively eating, usually during times of rest, but not during deep sleep. The time spent ruminating depends on the type of diet; it ranges from almost none for high-grain diets to a maximum of about 10 hours per day for high-forage diets. The feed intake level also influences the amount of rumination time, with high intakes stimulating greater rumination.

Water Moves Through the Rumen at a Much Faster Rate Than Particulate Matter

The flow of water has important effects on rumen dynamics. For small particles and soluble material to exit the rumen, liquid must constantly be washing through all sections of the rumen and moving through the reticulo-omasal orifice. This means that water must be constantly flowing through the mass of solid material. In effect, the reticulorumen functions as a giant strainer or sieve, holding the fermenting mass of particulate matter while water flows through it and washes small particles and soluble material away. Therefore the transit rate of water must be considerably greater than the transit rate of particulate matter through the rumen. The relative differences in the rates of movement of solid-phase and liquid-phase material through the rumen can be appreciated from their respective rumen half-lives: 30 to 50 hours for particulate matter and about 15 to 20 hours for liquid.

The rate of liquid flow through the rumen is often measured as the dilution rate, which is expressed as the percentage of total liquid that leaves the rumen in an hour. The term *dilution rate* comes from the way liquid turnover is measured; some soluble marker substance is mixed into the rumen, and its concentration is measured as soon as it is thoroughly dispersed into the liquid phase. Samples are then taken over time, and the rate at which the marker substance becomes diluted is measured. The rate of dilution depends on the rate at which water that contains marker leaves the rumen and is replaced with new, unmarked water; thus the dilution rate is an indirect measure of the rate of water flow

through the rumen. Normal dilution rate values vary with diet and feed intake, and are usually in the range of 5% to 30% per hour. One other point should be appreciated from the concept of dilution rate: water leaves the rumen only as it is replaced from some other source.

Almost all water that enters the rumen does so through the esophagus, from salivary flow, drinking, or succulent feeds. Thus the dilution rate depends on rates of salivation and drinking. The salivation rate is influenced by the chewing time and feed type; feeds such as long-stemmed dry roughages, which require relatively high rates of mastication, stimulate high rates of both salivary flow and dilution. Salivation occurs during rumination and during initial mastication; therefore those feeds that stimulate high rumination rates, such as forages, also stimulate high dilution rates. Conversely, feeds that do not stimulate extensive rumination (e.g., concentrates) result in relatively low dilution rates. The rate of drinking is influenced by (1) the rate of feed intake and (2) the salt, or electrolyte, content of the diet. Thus high rates of intake or diets with high electrolyte contents stimulate high dilution rates.

Little water enters the rumen by way of the mucosa. The mucosa of the forestomachs is stratified squamous epithelium and is aglandular; thus there is no direct fluid secretion. Some water can enter the rumen through osmosis, but under normal conditions, the amount appears to be minimal. Normal rumen osmolality is about 280 mOsm/kg, slightly less than the 300-mOsm/kg osmolality of blood and extracellular fluid. Thus the usual osmotic flow of water is out of the rumen. After consumption of relatively digestible feeds, rumen osmolality increases briefly because of VFA production; however, it appears that osmolalities in excess of 340 mOsm/kg are necessary for water to flow osmotically into the rumen. Under normal conditions, osmolalities that are this high are not sustained for long, and thus there is usually little osmotic flow of water into the rumen.

Rumen Dilution Rate Has Important Influences on Fermentation and Microbial Cell Yield

Small particles, including microbes, leave the rumen with the liquid phase. Therefore high dilution rates result in rapid removal of microbes and reductions in microbial cell concentrations. Because high microbe concentrations suppress microbial cell division, the growth of microbes is stimulated by high dilution rates. High growth rates are nutritionally desirable because a larger portion of the energy available to the microbes is used for growth instead of for maintenance, as occurs in older, relatively stable microbial populations. Thus high dilution rates usually increase Y_{ATP} values, provided that adequate protein is available to support cell growth.

In addition to its effect on Y_{ATP}, the dilution rate may affect the microbial makeup of the rumen biomass and also may have some influence on the fermentation pattern. The rate of microbial washout increases with the dilution rate. At high dilution rates, microbial species with slow growth rates diminish in population size because their replication rate is not great enough to match the rate at which they are removed. Thus selection pressure favors species with faster growth rates during times of high rumen dilution rates. Exceptions to this pattern occur because some microbes are able to attach themselves to the particulate matter in the solid and slurry zones. Such microbes then exit the rumen according to the kinetics of particle size reduction rather than dilution rate. In general, the changes occurring in the rumen microbial population with high dilution rates appear to favor acetic acid production and to increase the acetic/propionic acid ratio.

Control of Reticulorumen Motility

Reticulorumen Motility Is Controlled by the Central Nervous System and Affected by Intraluminal Conditions

In the dorsal vagal nucleus of the brainstem, there is a motility control center for the regulation of reticuloruminal motility. This center sends action potentials along fibers to the forestomach by way of the vagus nerve. There is an extensive enteric nervous system within the reticulorumen, but vagal innervation is necessary for coordination of normal motility patterns. When the vagal nerves are destroyed, motility of the rumen musculature ceases initially but returns within several days; however, the motility that develops after vagotomy is erratic, uncoordinated, and incapable of supporting the normal flow of ingesta through the reticulorumen. Vagotomized ruminants do not survive.

The dorsal vagal nucleus receives afferent stimuli that affect the control of forestomach motility. Important afferent signals come from the lumen of the reticulorumen and monitor distention, ingesta consistency, pH, VFA concentration, and ionic strength. Rumen volume, or distention, appears to be monitored by stretch receptors in the walls and especially in the pillars. Moderate distention increases rumen motility and rumination. Increased motility and rumination have the effect of raising the rate at which particles are broken down, leading to a higher passage rate. Thus rumen throughput is enhanced when increased intake expands rumen volume. Severe distention, as occurs pathologically in bloat, causes cessation of rumen motility.

The consistency of ingesta also has an important influence on rumen motility. Consistency is determined largely by diet type. When the diet consists of succulent plants, grain, or finely chopped forage, there is little material in the solid zone, or rumen mat, and the slurry zone is fluid. This type of fluid ingesta offers little resistance to the movement of the rumen pillars; thus the rumen musculature has to apply relatively little force to mix and circulate the rumen contents. Tension receptors in the reticuloruminal muscle appear to monitor the force necessary to move the pillars through the ingesta. Highly fluid rumen ingesta are associated with low muscle tension and have a negative influence on reticulorumen motility. At the other dietary extreme, when animals are eating dry, long-stem hay, the rumen contents are solid and create a large and highly interwoven rumen mat. Resistance to movement of the pillars through the solid mass of ingesta is high and leads to stimulation of tension receptors, resulting in a positive feedback on motility. The motility rate is directly related to the rate of particle breakdown; this arrangement appears to be a self-regulatory mechanism that increases the rate of particle comminution when animals consume diets with large particle size.

Chemoreceptors in the walls of the rumen and reticulum monitor pH, VFA concentration, and ionic strength (or osmolality). The pH of the reticulorumen is normally slightly acid, reflecting the acidity of the VFAs, but extreme acid conditions are undesirable. Increasing VFA concentrations or decreasing pH results in a suppression of rumen motility. The normal rumen pH is in the range of 5.5 to 6.8, depending on the type of diet. When the rumen pH falls much below 5.0, motility is severely depressed. This response appears to be protective because fermentation tends to be enhanced

by motility-induced mixing; thus suppression of motility slows down fermentation, allowing VFA absorption to catch up with VFA production.

Osmolality may influence rumen motility as well, although motility appears less sensitive to osmotic changes than it does to pH changes. Normal osmolality in the rumen is about 280 mOsm, but the osmolality increases during active fermentation. Osmotically active solutes in the rumen include organic acids and salivary and dietary electrolytes. As organic acid formation increases during fermentation, osmolality also increases, tending to reduce motility. The rumen epithelium creates a relatively impermeable barrier to water, so wide swings in rumen osmolality can occur without large shifts in water between the rumen and the vascular compartment. At abnormally high osmolalities, however, water can be drawn into the rumen.

Omasal Function

Passage of Material From the Reticulum to the Omasum Occurs During Reticular Contraction

The *omasum* is composed of a body and a canal. The body is filled with multiple muscular folds, or leaves, that project from the greater curvature into the lumen. The canal, which is located on the lesser curvature, connects the reticulum to the abomasum. Ingesta move into the omasum during reticular contractions. The reticulo-omasal orifice usually remains open, but dilates during the second phase of the reticular contraction, during which ingesta flow rapidly into the omasal canal. After the reticular contraction, the reticulo-omasal orifice closes briefly as the canal contracts, forcing newly arrived ingesta up into the leaves. Intermittently, the body and leaves of the omasum contract, forcing the material from the body of the organ into the canal and on into the abomasum.

Proper functioning of the omasum and reticulum appears to be particularly important to the passage of ingesta out of the rumen. Occasionally, traumatic injury resulting from ingested foreign bodies causes severe adhesions of the reticulum and omasum to the body wall. In addition, damage to vagal fibers entering the organs may occur. In such cases, motility of the rumen proper may continue normally, but the ability to move food out of the forestomachs and into the abomasum is severely impaired. The rumen becomes greatly distended with finely comminuted feed, and the entire rumen becomes a slurry zone. Despite the distended rumen, little movement of ingesta occurs into the abomasum, and the animals eventually suffer severe inanition. This condition is variably known as *omasal transport failure* and vagal indigestion; usually, little can be done to correct it.

The structure of the omasum, with its many leaves and large mucosal surface area, suggests that it has an absorptive function, but the exact nature of this function is still incompletely understood. One important possibility is that it exists to remove residual VFAs and bicarbonate from ingesta before material is transported to the abomasum. VFAs appear to cause unfavorable reactions in the abomasum, so it is important that a large portion of them be removed before abomasal entry. Also, it appears desirable to absorb, before abomasal entry, any bicarbonate remaining in the ingesta. Bicarbonate remaining in ingesta and entering the abomasum would only neutralize abomasal hydrochloric acid, increasing the work load of the abomasal glands to maintain appropriate abomasal pH.

Volatile Fatty Acid Absorption

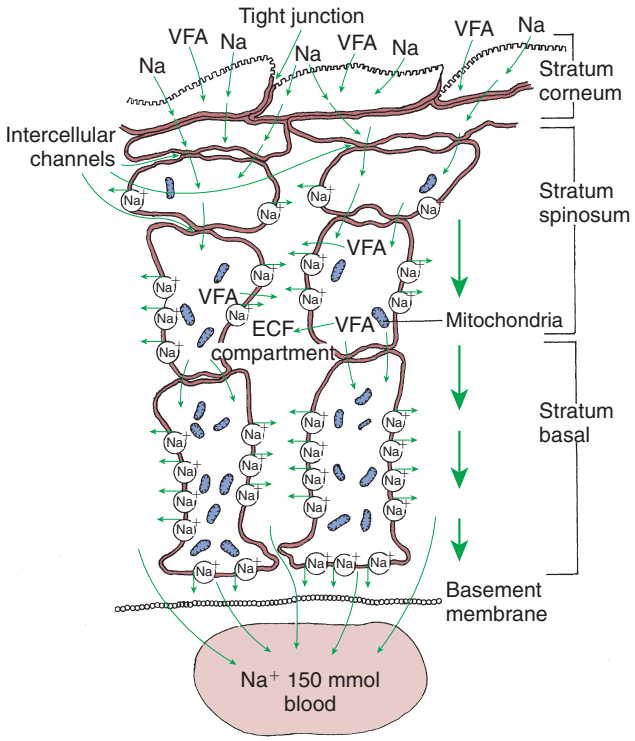
Volatile Fatty Acids, Representing 60% to 80% of the Energy Needs of the Animal, Are Absorbed Directly From the Forestomach Epithelium

VFAs are bacterial waste products and, if allowed to accumulate, will suppress fermentation. Furthermore, the VFAs are extremely important energy substrates for the host, supplying 60% to 80% of the dietary energy to ruminants with most types of diets. Therefore the presence of an efficient and high-capacity mechanism for VFA absorption is important to both digestion and host metabolism. The forestomach epithelium supplies such a system, absorbing a large portion of the VFAs. In addition, the absorptive process helps maintain rumen pH by removing acid from the forestomach ingesta and contributing bicarbonate in the process.

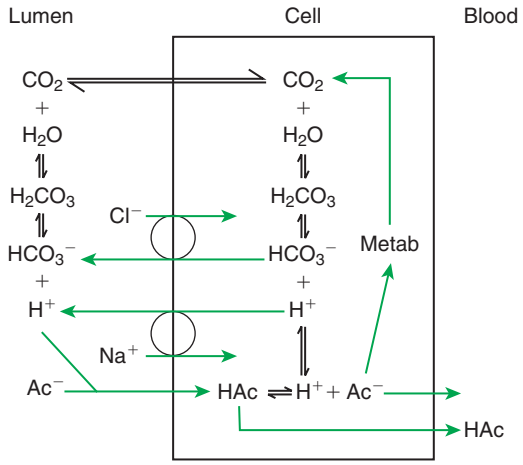
The epithelium responsible for this tremendous absorption is structurally much different from other absorptive epithelia of the GI system. However, the nature of the rumen epithelium may give it functional characteristics similar to those of the absorptive epithelium of the small intestine and colon. The forestomach surface is of the stratified squamous type and, as with the stratified squamous epithelium of the skin and other surfaces, consists of several layers of cells of varying maturity. The deepest layer is the *stratum basale,* from which cells divide and migrate into the stratum spinosum. Cells of the stratum spinosum begin the process of keratinization and continue into the *stratum granulosum,* which is covered by the outermost and most keratinized layer, the *stratum corneum.* Although the forestomach epithelium seems completely different from the columnar epithelium of the small intestine, an interesting similarity between forestomach and intestinal epithelia is noted when the cellular attachments and intercellular spaces of the forestomach are examined (Fig. 31.10).

The cells of the stratum granulosum are tightly joined by junctions that may functionally resemble the tight junctions of the enterocytes (see Chapter 30). Deeper in the epithelium, the cells of the stratum spinosum and stratum basale are separated by intercellular spaces that increase in size as the basement membrane is approached. These intercellular spaces are reminiscent of the lateral spaces of columnar absorptive epithelia. If these observations are combined with the existence of the intercellular bridges that characterize the forestomach epithelium, an interesting analogy to columnar absorptive epithelia can be constructed. VFAs, electrolytes, and water apparently are initially absorbed through the stratum corneum and passed cell to cell by way of intercellular bridges to the cells of the stratum spinosum and stratum basale, from which the absorbed substances are passed into the intercellular spaces before entering the capillaries.

This arrangement of the forestomach epithelium is very similar to the three-compartment characteristics of the columnar absorptive epithelia, with solutes passing from ruminoreticular lumen to cell contents and ultimately to lateral spaces. Although the keratinized cells of the stratum corneum do not appear to retain adequate metabolic machinery (e.g., mitochondria) to maintain appropriate gradients for diffusion, the cells of the stratum spinosum and stratum basale are metabolically active. Because of the intercellular bridges, absorbed solute can be transferred directly from the outer keratinized cells to the deeper, more metabolically active cells. Thus the metabolic activity deep in the epithelium appears to maintain conditions for absorption at the epithelial surface.



• **Fig. 31.10** The stratified squamous epithelium of the rumen, although anatomically much different, shares functional similarities with the columnar epithelium of the intestine. Note the tight junctions of the cells of the stratum corneum and the lateral space–like compartment between adjacent cells of the stratum spinosum and stratum basale. Although the cells of the stratum spinosum are metabolically inactive, the intercellular channels allow the metabolic actions of the stratum basale to be reflected in the more superficial layers. *ECF,* Extracellular fluid; *Na,* sodium; *VFA,* volatile fatty acid. (Modified from Steven DH, Marshall AB. Organization of the rumen epithelium. In: Phillipson AT, ed. *Physiology of Digestion and Metabolism in the Ruminant.* p. 98, Newcastle upon Tyne, UK: Oriel Press; 1970.)



• **Fig. 31.11** VFA absorption is promoted by the conversion of VFA anions *(Ac⁻)* to free acids *(HAc)* in the microenvironment near the epithelial surface. This diagram illustrates two proposed means, one intracellular and one extracellular, by which hydrogen ions could be locally generated to effect the formation of VFA free acids; both mechanisms could exist simultaneously. (From Stevens CE, Argenzio RA, Roberts MC. Comparative physiology of mammalian colon and suggestions for animal models of human disorders. *Clin Gastroenterol.* 1986;15(4):763-785.)

The molecular mechanism of VFA absorption is incompletely understood but seems to involve local alterations in pH near the absorptive surface. Differences in pH can have an important influence on VFA absorption because of shifts in the dissociation state of the VFA molecules. The pKa of the VFA is approximately 4.8, well below the normal pH of the rumen; thus most of the VFAs exist in the rumen in the dissociated, or ionic, form. However, sodium-hydrogen ion exchange by the epithelial cells may decrease the local pH at the absorptive surface. Such a drop in pH would lead to a shift in the VFA from the ionic to the free-acid state. Cell membranes are permeable to VFA free acids, and absorption proceeds because of the concentration gradient between the lumen and cells. The high CO_2 tension in the rumen, caused by the production of fermentation gases, may also enhance the conversion of VFA to the free-acid state. As shown in Fig. 31.11, when one VFA molecule is absorbed, one molecule of bicarbonate (HCO_3^-) is generated in the lumen; thus VFA absorption helps buffer rumen pH both by generating base and by removing acid.

All the VFAs appear to be absorbed by the same mechanism, but they are handled differently within the epithelial cells. Some acetate seems to be completely oxidized within the cells, with the remainder absorbed unchanged. Most propionate is absorbed, but a small portion is converted to lactate by the epithelial cells. Butyrate is modified extensively, and essentially all molecules are changed to β-hydroxybutyrate before absorption. β-Hydroxybutyrate is an important metabolite known as a *ketone body.* Ketone bodies are metabolites that frequently have special medical significance (see Chapter 32). In ruminants, the rumen itself is a significant source of ketone bodies. In monogastric animals, however, ketone bodies arise exclusively from the partial oxidation of long-chain fatty acids.

The rumen epithelium is arranged in *papillae,* finger-like projections that increase the absorptive surface area. Although they serve the same area-expanding function as the villi of the small intestine, papillae are much larger and easily visible to the unaided eye. The size and shape of the papillae are quite dynamic and responsive to changes in diet. Papillary growth is stimulated by VFAs, especially butyrate and propionate. Diets with high digestibility result in high rumen VFA concentrations, which stimulate the growth of long papillae. In contrast, animals receiving little feed or diets of low digestibility have short rumen papillae. It is important to adapt ruminants gradually when changing them from diets of low digestibility to high digestibility; this may allow time for sufficient adjustment of papillary size so that VFA absorption will match VFA production.

Rumen Development and Esophageal Groove Function

Significant Changes in Forestomach Size and Function Occur With Dietary Changes in Early Life

At birth, the forestomach is about equal in size to the abomasum in both lambs and calves, a stark contrast to the normal adult proportions, in which the forestomach accounts for more than 90% of the total stomach volume. Enlargement of the forestomach occurs rapidly after birth, but the rate depends on diet type. When

young ruminants are given access to solid feeds soon after birth, the forestomach development rate is maximal.

In cattle, the period of forestomach development is arbitrarily divided into the *nonruminant period,* from birth to 3 weeks, and the *transitional period,* from 3 to 8 weeks. Approximate adult distribution of stomach proportions is achieved usually by 8 weeks, if calves have access to solid feeds. Calves can be seen eating grain and forage as early as 2 weeks of age and frequently ruminate by 3 weeks, indicating considerable forestomach development by this time. Withholding solid feed dramatically reduces the rate of rumen development. In calves that are given diets of only milk or milk substitute ("replacer"), forestomach development remains rudimentary for 14 to 15 weeks or longer.

Development of forestomach epithelium parallels the general development of the organ. At birth, the epithelium is thin, with small or nonexistent papillae. Exposure of the epithelium to VFAs appears to stimulate papillary development and general organ development as well. Highly digestible feeds, such as concentrates, result in the greatest VFA production and fastest epithelial development. Some dietary forage may aid in muscular development of the forestomachs, but calves and lambs in the transitional period should receive most of their solid feed as grain because their energy needs are high compared with their ability to ferment forages.

The forestomach is sterile at birth but is quickly colonized by environmental bacteria, mostly facultative organisms. As bacterial fermentation proceeds in the anaerobic confines of the forestomach, the electromotive force decreases; the typical reductive environment of the rumen is created by bacterial action. This environment creates conditions necessary for the growth and establishment of the strict anaerobes. The development of forestomach bacterial flora occurs independently of any special inoculation process, and indeed, it is impossible to prevent it from occurring except by raising calves under gnotobiotic conditions. Protozoal inoculation, in contrast to bacterial inoculation, seems to require some exposure to other cattle; calves raised in complete isolation do not develop protozoal fauna. It appears that aerosol spread of protozoa can occur because no direct physical contact among cattle is necessary to establish a protozoal fauna.

The Esophageal Groove Diverts the Flow of Ingested Milk Past the Forestomach and Into the Abomasum

For proper rumen development in the suckling animal, it is important for milk to be diverted away from the developing rumen. This is accomplished by the actions of the *reticular groove* (also called the *esophageal groove*). This structure is a gutter-like invagination traversing the wall of the reticulum from the cardia to the reticulo-omasal orifice. When stimulated, muscles of the groove contract, causing it to shorten and twist. The twisting action causes the lips of the groove to close together, forming a nearly complete tube from the cardia to the omasal canal. When the groove is contracted, milk entering the cardia is directed into the omasum, with 10% or less entering the rumen. Milk quickly traverses the omasum and enters the abomasum.

Reticular groove closure is a reflex action, with brainstem efferent impulses arriving through the vagus nerve. Afferent stimuli arise centrally and from the pharynx. Anticipation of suckling invokes central stimulation of reticular groove closure, which may be considered a cephalic phase. Fluid, especially sodium-containing fluid in the pharynx, stimulates afferent fibers that reinforce the cephalic phase of groove closure. The posture of the calf or lamb when suckling does not appear to have much influence on reticular groove function, but rapid drinking from an open pail, in contrast to suckling from a nipple, frequently results in inefficient groove function and spillage of milk into the rumen. Milk in the rumen results in the formation of improper fermentation patterns.

The reticular groove has its primary function in suckling animals, and the activity of the groove reflex appears to diminish after weaning and with advancing age. However, the groove reflex is stimulated by antidiuretic hormone (ADH; see Chapter 43), indicating that it may have some physiological function in adult life. ADH is secreted by the posterior pituitary in response to dehydration or increases in plasma osmolality. ADH is associated with thirst, and because it stimulates the reticular groove, a large portion of the water consumed by drinking in water-deprived animals may bypass the rumen. This may be a functional mechanism to ensure that water arrives quickly at the site of most rapid absorption, the small intestine.

Function of the Equine Large Hindgut

The Equine Hindgut Has a Great Capacity for Fermentation

A general function of the cecum and colon, as mentioned in Chapter 30, is to recover fluid and electrolytes from ingesta leaving the ileum. In many herbivorous species, this function has been expanded to include fermentative digestion. Absorptive and fermentation functions complement each other in the colons of nonruminant herbivores. This arrangement leads to an elegantly interactive system of fermentation and absorption; however, it also results in an interdependence between the two processes, meaning that disturbances in fermentation can result in important abnormalities in absorption, and vice versa.

The Types of Substrate and Fermentation Patterns Are Essentially Identical for Forestomach and Hindgut Fermentation

Structural and nonstructural carbohydrates and proteins form the major substrates for hindgut fermentation. However, the passage of material through the stomach and small intestine before its arrival at the cecum and colon may have some important effects on fermentative digestion. First, hindgut fermentation may be aided by prior gastric action. The effects of soaking and acid exposure on plant particles in the stomach may increase their susceptibility to microbial attack and thus raise their rate of digestion in the hindgut. Second, some of the readily available carbohydrate, particularly sugars and starches, may be digested and absorbed before the other material arrives in the cecum. Most evidence indicates, however, that glandular digestion of carbohydrate in the horse is not extremely efficient and that substantial amounts of starch and sugars reach the cecum. Further, cell-wall carbohydrate appears to interfere with the digestion or absorption of nonstructural carbohydrate, so that diets high in cell-wall content result in relatively little starch digestion and absorption in the equine small intestine. Even with a high-grain diet, up to 29% of dietary starch may reach the cecum and colon.

Protein and carbohydrate is absorbed in the small intestine, potentially leading to a deficiency of nitrogen for colonic microbes. However, there is extensive urea recycling into the colon and cecum, similar to that occurring in the rumen (see Fig. 31.6). Thus urea plus protein escaping small-intestinal digestion supplies the nitrogen

needs of the microbes. In contrast to ruminants, horses do not have an efficient means of recovering the microbial protein synthesized in the hindgut, and most of it passes out in the feces. Some experiments have shown a small amount of amino acid absorption from the equine cecum or colon, but the amount does not compare with microbial protein availability in the ruminant.

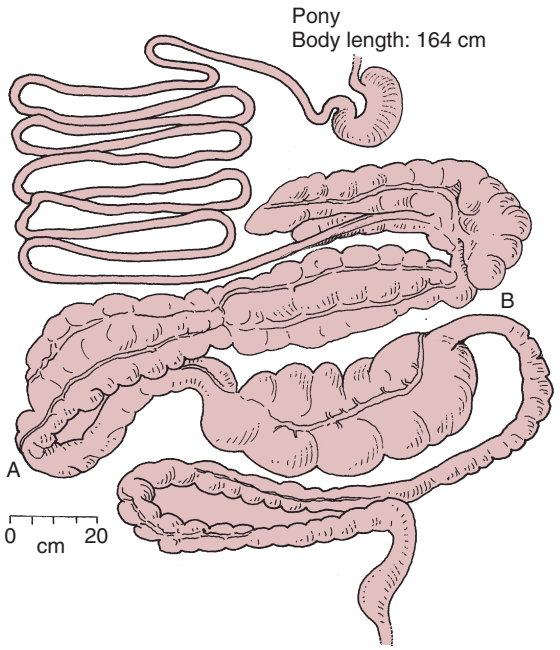
The Motility Functions of the Cecum and Colon Retain Material for Fermentation and Separate Particles by Size

The functions of the equine hindgut in maintaining fermentation are similar to those of the rumen: favorable conditions must be maintained to support optimal fermentation. As in the rumen, these conditions are (1) substrate supply, (2) control of pH and osmolality, (3) anaerobiosis, (4) retention of fermenting material, and (5) continual removal of waste products and the residue of spent fermentation substrate. Separation of fermenting material from residue appears to be accomplished by selective retention of particles according to size, just as in the rumen; however, the means by which the cecum and colon accomplish size separation and discriminate passage are quite different from those of the forestomach. Anatomical characteristics and motility patterns in the cecum and colon are responsible for selective retention of long particles, allowing sufficient exposure for microbial digestion to occur. In general, the fermentative digestive process in the horse is not as efficient as that in the ruminant, and digestible energy values for forages are usually lower for horses than for cattle.

Before the motility of the equine cecum and colon is discussed, a brief review of the anatomy of the equine hindgut is important. Fig. 31.12 shows the equine digestive system, separated from its mesenteric attachments and laid out in linear fashion. The hindgut commences with the cecum, which is separated from the large colon by a well-defined orifice. The large colon is folded on itself three times, forming four major anatomical divisions: the *right ventral* and *left ventral* and the *left dorsal* and *right dorsal* colon segments. Ingesta enter the right ventral colon and course to the left ventral colon, from which the material enters the left dorsal portion through the *pelvic flexure.* From the left dorsal colon, material moves to the right dorsal colon before entering the small colon. (Anatomy textbooks describe the arrangement of the large colon in the abdomen.) For the purposes of physiological study, the reader should note in Fig. 31.12 the tremendous size and volume of the cecum and colon compared with the small intestine. The differences in diameter that occur throughout the colon should also be noted, particularly the reductions in diameter that occur at the pelvic flexure and at the junction of the large and small colons. The saclike evaginations that occur in the wall of the cecum and most segments of the colon are called *haustra.* Functionally, the equine hindgut can be divided into four sections: cecum, ventral colon, dorsal colon, and small colon.

Ingesta reach the cecum after a relatively short time in the stomach and small intestine. A large portion of soluble ingesta usually reaches the cecum by 2 hours after ingestion, whereas solids take somewhat longer, depending on particle size and consistency. The material in the cecum and throughout the large colon has a high water content and a slurry-like consistency.

The majority of cecal motility is of a mixing nature, with frequent low-amplitude contractions that transport ingesta from haustrum to haustrum and back in a mixing pattern. The mixing action of



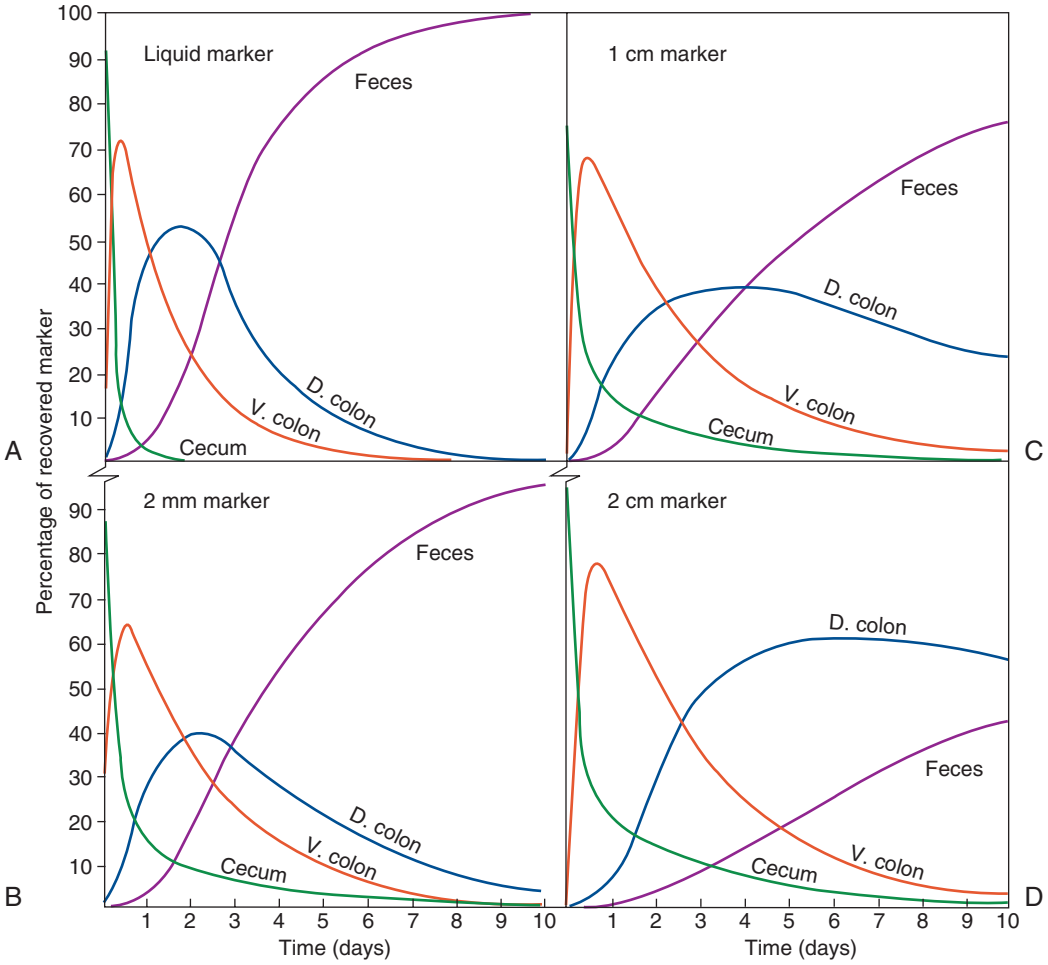
• **Fig. 31.12** Equine gut. Note the tremendous development of the colon compared with the small intestine. Note also the relative areas of constriction at the junctions of the (A) ventral and dorsal colons (B), and the large and small colons. (From Stevens CE. Comparative physiology of the digestive system. In: Swenson MJ, ed. *Dukes' Physiology of Domestic Animals.* 9th ed. p. 221 Ithaca, NY: Cornell University Press; 1977.)

the cecum maintains the cecal contents in a homogeneous state. About once every 3 to 4 minutes, there is a strong contraction of cecal muscles in a mass movement type of action (see Chapter 28 for a description of *mass movement*) in which the *body* and *apex* of the organ shorten and constrict, lifting ingesta into the *base.* Constriction of the base forces material through the *cecocolic orifice* and into the right ventral colon. The motility pattern functionally separates the cecum from the ventral colon. In contrast to some other species, in horses there is no retrograde flow of material from the colon back into the cecum, so the composition of ingesta in these two organs usually differs somewhat.

Three types of motility patterns exist in the right and left ventral colon: *haustral segmentation, propulsive peristalsis,* and *retropulsive peristalsis.* Segmentation serves a mixing function that aids in promoting fermentation and bringing VFAs in contact with the mucosa for absorption. Mixing occurs throughout the ventral colon, and the right and left segments may be regarded as one functional unit with homogeneous ingesta. Propulsive activity, or aboral peristalsis, in the ventral colon originates near the cecum and appears to occur as a continuation of the cecal mass movements. Peristaltic activity in the proximal ventral colon propels ingesta distally into the left ventral colon. In the left ventral colon, retropulsive or antiperistaltic movements resist the flow of ingesta and result in the retention and mixing of material in the ventral colon, allowing time for microbial digestion and preventing the washout of microbes. In addition, the retropulsive actions of the left ventral colon aid in creating differential flow rates of liquid and particulate matter through the colon. The antiperistaltic motility appears to originate from a pacemaker in the *pelvic flexure,* the area of restricted diameter where the left ventral and left dorsal colons meet.

The motility of the ventral colon can be roughly compared with that of the stomach, with the pelvic flexure and distal left ventral colon acting as the pylorus and antrum, respectively. The pumping action of cecal mass movements, combined with the propulsive action of the proximal ventral colon, continually moves ingesta toward the pelvic flexure. In the distal ventral colon, however, antiperistaltic activity and the narrow diameter of the pelvic flexure impede the movement of material, causing it to be retained in the ventral colon. The squeezing action of the pelvic flexure mimics the action of the pylorus in selectively retaining relatively large particulate matter while allowing liquid and small particles to pass. As particle size is reduced by fermentative action and the mixing activity of the colon, particles eventually become small enough to flow with the fluid phase and leave the colon. The action of the pelvic flexure is not as efficient as that of the pylorus, and some large particles do escape the ventral colon. In addition, there are periods during which propulsive movements occur in the left ventral colon and pelvic flexure. These factors allow the movement of particulate matter into the left dorsal colon.

The actions of the dorsal colon appear to mimic those of the ventral colon. Impedance to ingesta flow is created by the size restriction at the junction of the right dorsal colon and small colon. In addition, retropulsive motility may originate in the area of the distal right dorsal colon, near the junction with the small colon. These actions tend to impede the movement of ingesta through the dorsal colon, subjecting the material to another round of fermentative digestion, as occurred in the ventral colon. The delay in the flow of ingesta created by the combined actions of the ventral and dorsal colons results in significant retention of material, with most particulate matter taking from 24 to 96 hours to pass the large colon. The efficiency of the large colon in retaining and separating ingesta of different particle sizes can be understood from Fig. 31.13.



• **Fig. 31.13** Retention of liquid and particles of various sizes in the compartments of the equine large intestine. Marker liquid (A) and marked particles of various sizes (B, 2 mm; C, 1 cm; D, 2 cm) were placed in the cecum of ponies, and the distribution of marked materials among colon segments was measured at 2-hour intervals. The lines of the graph were mathematically fitted to the data. Each line indicates the percentage of marker in a given segment at any time. Note in graph A that at 7 days after infusion, almost all the liquid marker had been recovered in the feces, with little or none remaining in the intestinal segments. Increasing particle size has a relatively small effect on the movement of particles out of the cecum. In contrast, as particle size increases, there is significant retention of material in the colon and slow passage to the feces. *D,* Dorsal; *V,* ventral. (From Argenzio RA, Lowe JE, Pickard DW, et al. Digesta passage and water exchange in the equine large intestine. *Am J Physiol.* 1974;226(5):1035-1042.)

Understanding the motility of the equine colon is important because problems of colon impaction in horses are common. Impactions usually occur near or within the pelvic flexure, probably because the pelvic flexure is a site of flow restriction and differential flow of solid and liquid material. One can easily appreciate how the normal motility pattern could allow solid material to accumulate in this area and cause obstructions.

Although general understanding of small colon motility is limited, it appears to consist primarily of segmentation and propulsion. The characteristic fecal balls of horses are formed by segmentation within the small colon.

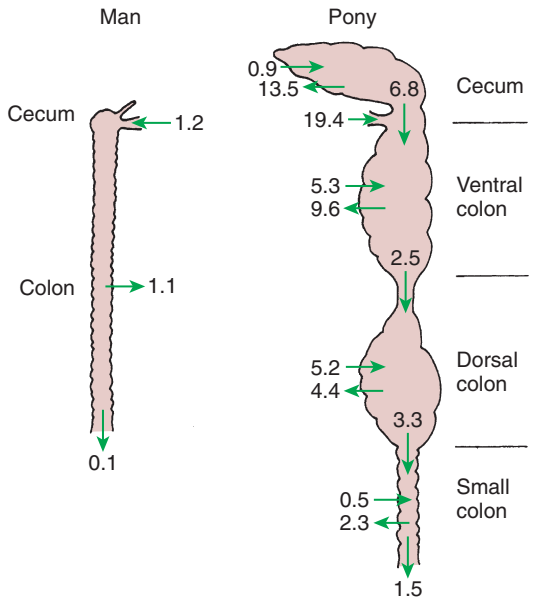
The Rate of Fermentation and Volatile Fatty Acid Production in the Equine Colon Is Similar to That in the Rumen

In the equine colon, efficient means of buffering and VFA absorption must be present. Salivary buffering, as occurs in ruminants, cannot aid in buffering the colon. In the horse, large quantities of fluid, rich in bicarbonate and phosphate buffers, are secreted by the ileum and transferred to the cecum, thus mimicking the actions of the salivary glands in ruminants. Also, because of the glandular nature of the colonic mucosa, bicarbonate and other electrolytes are added more directly to the lumen fluid in the cecum and colon than in the rumen.

Large fluxes of water traverse the cecal and colonic mucosa during the course of digestion. When horses are meal-fed, feed starts to enter the cecum about 2 hours after eating, and VFA production rapidly commences. As ingesta are transported from the cecum, VFA production continues in the large colon. During the period of active VFA production, large quantities of water enter the hindgut from the blood through the mucosa. Although this water flux may be a response to increased osmolality created by the generation of osmotically active VFA molecules, it is more likely a response to direct fluid secretion from the crypts of the colonic epithelium (see Chapter 30). Secretion of fluid containing sodium, bicarbonate, and chloride from the colonic mucosa appears to occur in response to high concentrations of VFA in the lumen. This secretory response, in combination with the ileal secretions, is responsible for buffering of the lumen contents. Fig. 31.14 illustrates the magnitude of water fluxes that occur during hindgut digestion in the pony. Note that considerable inward and outward movement of water occurs across the mucosa in each of the major fermentation compartments, ventral and dorsal colons, and cecum. Inward (into the lumen) water movement results from mucosal secretion, whereas outward water movement occurs in association with absorption of VFA.

The molecular mechanisms of VFA absorption in the equine colon appear to be identical to those in the rumen (see Fig. 31.11). Note in Fig. 31.11 that sodium absorption accompanies VFA absorption and that bicarbonate is generated in the lumen. The absorption of VFA and sodium leads to osmotic absorption of water, probably through the transcellular pathway. (See Chapter 30 for a review of the dynamics of water and electrolyte absorption in the gut.)

The function of the small colon is to recover water, electrolytes, and VFAs that were not absorbed in the large colon. VFA production appears minimal in the small colon, but considerable absorption of water, sodium, and phosphate occurs there.



• **Fig. 31.14** Net movement of water through the large intestine of a 70-kg man and a 160-kg pony. Values are in liters per day. Note the relatively large amount of fluid delivered to the pony's colon from the ileum (19.4 L/day) compared with the human colon. Also note the inward and outward movement of fluid in the various compartments of the pony's large intestine. (From Argenzio RA, Lowe JE, Pickard DW, Stevens CE. Digesta passage and water exchange in the equine large intestine. *Am J Physiol.* 1974;226(5):1035-1042.)

The large water and electrolyte fluxes in the colon make horses vulnerable to colonic diseases, resulting in fluid and electrolyte losses that are more characteristic of small-intestinal disease in many other animals.

Hindgut Anatomy and Function Vary Greatly Among the Many Species of Veterinary Interest

All the variations in hindgut anatomy and function among the species are beyond the scope of this discussion, but you should remember that rabbits, rats, guinea pigs, swine, and some large birds, in addition to Equidae, depend on hindgut fermentation for a significant portion of their energy needs. Also, ruminants have a reasonably extensive hindgut, and fermentative digestion occurs there, even after material has been through the rumen.

Scientific understanding of colonic function in general is not as advanced as that of small-intestinal function. This dichotomy probably exists because in humans, diseases of the small intestine are more frequent and severe than diseases of the colon. However, interest in colonic physiology and pathophysiology has increased among basic scientists and physicians. Veterinary physiologists have long been interested in colonic function and have become leaders in the exploration of this area.

CLINICAL CORRELATIONS

Grain Engorgement Toxemia

History

In mid-January, a cattle feeder has asked you to examine a feedlot full of 400-kg steers. The steers have been on free-choice grain from a self-feeder for several weeks. Three days ago, a blizzard prevented the caretaker from delivering feed to the feeders, and they were empty for 36 hours. Yesterday the feeders were filled, and all the steers ate ravenously. Today, two of the 40 steers are dead, and many appear lethargic and uncoordinated, and have diarrhea.

Clinical Examination

Two steers are isolated for physical examination; they are lethargic and must be coaxed to move. Their heart rates are all above 100 beats/min (normal, <80 beats/min), and their body temperatures are less than 101.0° F (normal, 101.5° F to 103.0° F). Their rumens appear distended, and there is no evidence of rumen motility. The eyes are sunken in the orbits, and the oral mucosa is dry and sticky, indicating clinical dehydration. Necropsy examination of a dead steer reveals a greatly distended rumen filled with grain and fluid. A pH-paper test indicates that the rumen fluid pH of the dead animal is below 4.5 but above 3 (normal, 5.5–7.0).

Comment

Ruminants may be fed large amounts of grain if they are accustomed to it and receive it on a regular and frequent basis. In this case even though the steers had been accustomed to a high-grain diet, the lack of grain for more than a day, followed by a large grain intake, set up conditions for *grain engorgement toxemia.*

In grain engorgement, there is an abundant supply of starch, leading to the rapid growth and proliferation of rumen streptococci. These bacteria produce VFA rapidly, causing the rumen pH to diminish. As the rumen pH becomes lower, conditions become unfavorable for the growth and survival of cellulose-fermenting organisms and favorable for the growth of lactic acid–producing bacteria. This leads to the accumulation of lactic acid, a stronger acid than the VFAs. Thus rumen pH becomes even lower, killing many of the normal microflora. Some of the lactic acid is absorbed, leading to a reduced blood pH and a life-threatening situation. Moreover, the large ruminal concentration of lactic acid and VFA results in a high osmotic pressure, drawing water out of the vascular fluid compartment and into the rumen. This leads to systemic hypovolemia, which may proceed to hypovolemic shock.

Treatment

This is a life-threatening situation, and the farmer will likely lose additional steers. Treatment is aimed at expanding the intravascular fluid volume, correcting the systemic acidosis, and reestablishing a normal rumen environment. Severely affected steers should be evaluated to determine whether their prognosis is good enough to warrant the expense of therapy; if not, euthanasia should be employed. Initial treatment should consist of rapid intravenous administration of large quantities of alkalizing fluid. After the correction of fluid and acid-base disturbances, ideally the rumen should be emptied either by intubation with a large-bore stomach tube or by rumenotomy. In some cases, oral administration of an antifermentation agent, such as oil of turpentine, mineral oil, or an antibiotic, along with an alkalizing agent, is an acceptable alternative to emptying the rumen. After the rumen environment is brought back to normal, it may be helpful to reinoculate the rumen with material taken from the rumen of a normal animal.

Impaction Colic

History

You are presented with a 20-year-old gelding that has been showing signs of abdominal discomfort (colic) for 16 hours. When left alone in the stall, the horse lies down, frequently preferring to lie on his back. There is little fresh manure in the stall. When taken out of the stall, the gelding leads normally, but then lies down and rolls on release from the lead rope.

Clinical Examination

The heart rate is slightly elevated, at 60 beats/min; respiratory rate and temperature are normal. The hydration state, and the color and perfusion of the mucous membranes, is normal. Simple laboratory evaluation reveals the packed cell volume to be 41% (normal, 35%–45%) and the plasma total solids to be 7.8 g/dL (normal, 6.5–8.0 g/dL). Borborygmi (intestinal sounds) are softer and less frequent than normal, especially on the left side. Examination by rectal palpation reveals the pelvic flexure to be firm with a doughlike consistency; normally, the contents of the pelvic flexure have a fluid consistency. When you examine the teeth, you find that the molar surfaces are irregular, and one of the molars has a crack extending from the table surface to below the gum line.

Comment

The pelvic flexure is a site of flow restriction and particle size separation. As water moves through the pelvic flexure, large forage particles accumulate and are retained for further fermentation and mixing in the ventral colon. A horse with poor teeth may not chew its forage adequately; as a result, many large particles may be swallowed. These particles tend to accumulate in the pelvic flexure and may cause an impaction and obstruction, as occurred in this case.

Treatment

Treatment involves the oral administration of softening agents, such as mineral oil. Drugs such as dioctyl sodium sulfosuccinate, which stimulate water secretion from the intestinal mucosa, are also beneficial. Prevention in this case involves correction of the dental problems so that forage is more thoroughly chewed. Feeding pelleted feeds may also be beneficial.

Practice Questions

1. In which of the following respects is fermentative digestion *different* from glandular digestion?
- **a.** Enzymes are not involved in fermentative digestion.
- **b.** Chemical bonds are not split by hydrolysis in fermentative digestion.
- **c.** Only carbohydrates are digested by fermentative digestion.
- **d.** Substrates are more extensively altered in fermentative digestion than in glandular digestion.
- **e.** Proteins are digested to amino acids by fermentative digestion and to dipeptides by glandular digestion.

2. In a comparison of hindgut fermentation and forestomach fermentation, which of the following statements is *true?*
- **a.** The microbial populations are considerably different, but the products of digestion are the same.
- **b.** The microbial populations are the same, but the products of digestion are considerably different.
- **c.** Both the microbial populations and the digestion products are similar.
- **d.** Structural carbohydrates of plants are not digested by hindgut fermentation.
- **e.** A nitrogen source is not required by the microbes of the hindgut.

3. The three VFAs—acetate, propionate, and butyrate—are:
- **a.** Net-reaction products of the fermentative action of the entire rumen biomass.
- **b.** The individual products of cellulose, starch, and hemicellulose digestion, respectively.
- **c.** The individual products of bacterial, protozoal, and fungal digestion, respectively.
- **d.** Volatile products that leave the rumen with the gas phase during eructation.
- **e.** Intermediate metabolites that are passed between microbial species.

4. Matching protein and energy availability in the rumen is an important nutritional goal in feeding ruminants. Which of the following completions of this statement concerning protein and energy availability in the rumen is *false?* Diets well matched in available protein and energy result in:
- **a.** The most efficient use of energy for microbial growth.
- **b.** Maximal delivery of protein to the host.
- **c.** Minimal breakdown of protein in the rumen.
- **d.** Loss of a minimal amount of dietary amino acids because of formation of excess ammonia.
- **e.** Optimal rumen ammonia concentrations.

5. Which of the following statements is *true* of both methane and propionate?
- **a.** They are waste products of anaerobic fermentation but contain potential energy that is recoverable by the host.
- **b.** They are highly oxidized molecules.
- **c.** They are eructed from the rumen.
- **d.** Their formation results in the generation of NAD from NADH.
- **e.** They are toxic to monogastric animals.

Bibliography

Entry-Level Supplements

Millen DD, Arrigoni MDB, Pacheco RDL. *Rumenology*. Springer International Publishing Switzerland: 2016 [Chapters 1-3].

Advanced Reading

Aschenbach JR, Penner GB, Stumpff F, Gäbel G. Ruminant nutrition symposium: role of fermentation acid absorption in the regulation of ruminal pH. *J Anim Sci*. 2011;89(4):1092–1107.

Cheeke PR, Dierenfeld ES. *Comparative Animal Nutrition and Metabolism*. Wallingford, UK: CAB International; 2010.

Forbes JM, France J, eds. *Quantitative Aspects of Ruminant Digestion and Metabolism*. Wallingford, UK: CAB International; 1993.

Gäbel G, Aschenbach JR, Müller F. Transfer of energy substrates across the ruminal epithelium: implications and limitations. *Anim Health Res Rev*. 2002;3(1):15–30.

Khan MA, Bach A, Weary DM, von Keyserlingk MAG. Invited review: transitioning from milk to solid feed in dairy heifers. *J Dairy Sci*. 2016;99(2):885–902.

Penner GB, Steele MA, Aschenbach JR, McBride BW. Ruminant nutrition symposium: molecular adaptation of ruminal epithelia to highly fermentable diets. *J Anim Sci*. 2011;89(4):1108–1119.

Reynolds CK, Kristensen NB. Nitrogen recycling through the gut and the nitrogen economy of ruminants: an asynchronous symbiosis. *J Anim Sci*. 2008;86(14 suppl):E293–E305.

Russell JB. *Rumen Microbiology and Its Role In Ruminant Nutrition*. Ithaca, NY: Cornell University Press; 2002.

Sejrsen K, Hvelplund T, Nielsen MO, eds. *Ruminant Physiology: Digestion, Metabolism and Impact of Nutrition on Gene Expression, Immunology and Stress*. Wageningen, The Netherlands: Wageningen Academic Publishers; 2006.

Stevens CE, Hume ID. *Comparative Physiology of the Vertebrate Digestive System*. 2nd ed. Cambridge, UK: Cambridge University Press; 1995.

Van Soest PJ. *Nutritional Ecology of the Ruminant*. 2nd ed. Ithaca, NY: Cornell University Press; 1994.

32

Postabsorptive Nutrient Utilization

THOMAS H. HERDT

KEY POINTS

1. Homeostatic mechanisms balance the supply and demand of almost all nutrients.

The Furnace

1. The tricarboxylic acid (or Krebs) cycle is the major energy-yielding pathway of fuel utilization in the body.

The Fuels

1. The major metabolic fuels consist of glucose, amino acids, fatty acids, and ketone bodies.
2. Glucose is the central fuel in the energy metabolism of most animals.
3. Amino acids are important fuels in addition to being the building blocks of protein.
4. Fatty acids are the major form of energy storage in the animal body.
5. Ketone bodies are fat-derived, water-soluble metabolites that serve as glucose substitutes.

Nutrient Utilization During the Absorptive Phase

1. During the absorptive phase, the liver takes up glucose and converts it into glycogen and triglyceride.
2. The conversion of glucose to fatty acids is an irreversible process.
3. Transport of fatty acids out of the liver is through chylomicron-like particles known as very-low-density lipoproteins.
4. Amino acids can be classified into groups on the basis of metabolic characteristics.
5. Amino acids are extensively modified during absorption.
6. Many amino acids are removed by the liver on "first pass," never reaching the systemic circulation.
7. Some amino acids taken up by the liver are used for protein synthesis.
8. Most amino acids taken up by the liver are converted to carbohydrates.
9. Not all amino acids are subject to hepatic destruction.
10. Metabolism at the tissue level is coordinated with hepatic metabolism and results in the deposition of fuel into storage tissues during the absorptive period.
11. Insulin promotes the synthesis of protein and the deposition of glycogen in muscle.
12. Insulin-stimulated uptake of amino acids by muscle results in a net increase in muscle protein synthesis.
13. During the absorptive phase, triglyceride accumulation in adipose tissue occurs by two mechanisms: uptake from very-low-density lipoproteins and direct lipid synthesis from glucose.

Nutrient Utilization During the Postabsorptive Phase

1. Hepatic metabolism switches from glucose utilization to glucose production during the postabsorptive phase.
2. Fuel mobilization in peripheral tissues occurs when the blood insulin concentration declines.
3. Muscle reacts to a metabolic demand for glucose by mobilizing amino acids to support hepatic gluconeogenesis.
4. Muscle release of amino acids is related to reduced glucose and amino acid uptake.
5. The complex pattern of muscle amino acid catabolism and release is necessary to accommodate the liver's limited capacity for uptake of branch-chain amino acids and to facilitate the removal of amino nitrogen from the muscle.
6. The reaction of adipose tissue during the postabsorptive phase is to mobilize fatty acids.

Nutrient Utilization During Prolonged Energy Malnutrition or Complete Food Deprivation

1. During prolonged periods of fasting or undernutrition, glucose and amino acids are conserved by extensive utilization of fats and ketone bodies for energy production.
2. A large portion of the fatty acids released from adipose tissue are taken up directly by the liver.
3. Hepatic ketone body formation is promoted by low glucose availability, a high glucagon/insulin ratio, and a ready supply of fatty acids.
4. Glucagon plays an important role in the excessive production of ketone bodies in diabetes mellitus.
5. Fatty acids cannot be used for glucose synthesis.
6. Ketone bodies are formed in the mitochondria from acetyl coenzyme A.
7. Hepatic very-low-density lipoproteins may be synthesized from adipose-derived fatty acid and from newly synthesized fatty acid.
8. Hormonal conditions direct the distribution of very-low-density lipoprotein fatty acids in the body.
9. Changes in growth hormone concentrations may aid in shifting peripheral fuel utilization from glucose and amino acids to ketone bodies and fatty acids.

Special Fuel Considerations of Ruminants

1. Ruminants exist in a perpetual state of gluconeogenesis because of their unique digestive process.

The rate of absorption of nutrients from the gut is not constant but rather fluctuates greatly with food intake. Meals are digested at a rate dependent on their chemical composition, regardless of the animal's nutrient needs. The nature of digestion dictates that nutrient absorption from the gut is rapid during digestion and then ceases during interdigestive periods. In other words, the gut is not a storehouse for nutrients, and digestion is not modulated by the animal's nutritional demands. Nutrient needs are not well matched to the wide fluctuations that occur in nutrient absorption from the gut. In fact, a vital need exists for a constant, steady supply of fuel-providing nutrients to maintain the basal metabolic functions of the body. In addition, the periods when the animal's metabolic needs are greatly elevated often do not coincide with times of rapid absorption of nutrients from the gut. Therefore animals must have a sophisticated system for maintaining the supply of nutrients, particularly energy-supplying nutrients, and buffering both the short-term and the long-term "feast or famine" effects associated with the absorptive and postabsorptive periods of digestion.

Homeostatic Mechanisms Balance the Supply and Demand of Almost All Nutrients

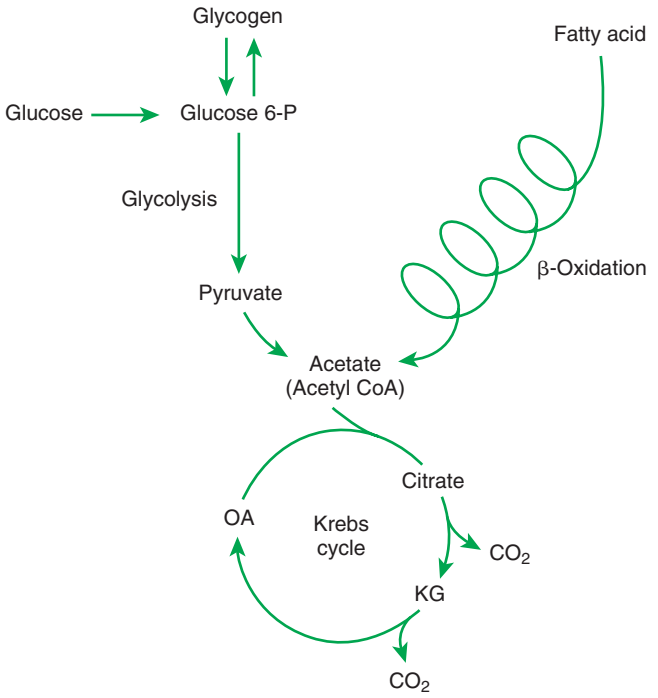
This chapter focuses on supply regulation of the major energy-supplying nutrients; however, other nutrients, including vitamins and minerals, also are subject to homeostatic regulatory mechanisms. Although many of these mechanisms directly involve the digestive system, a complete discussion is beyond the scope of this book. The homeostatic mechanisms regulating the supply of minerals and vitamins are described in select references of the Bibliography.

Energy-supplying nutrients are referred to as *metabolic fuels,* and the physiological mechanisms for maintaining the supply of fuels and matching it to demand constitute *fuel homeostasis.* Fuel homeostasis is maintained by several mechanisms: the insulin-glucagon axis, the hypothalamic-pituitary axis, and the central nervous system (CNS). This chapter discusses some of the ways in which fuel is stored during the absorptive period of digestion and subsequently mobilized when needed to supply energy requirements. You may want to review Chapter 1 and the section on insulin and glucagon in Chapter 34 before reading this chapter.

The Furnace

The Tricarboxylic Acid (or Krebs) Cycle Is the Major Energy-Yielding Pathway of Fuel Utilization in the Body

The Krebs cycle is the pathway through which all body fuels are eventually "burned." In this cycle, carbon compounds from the various body fuels are completely oxidized to carbon dioxide and water. Much of the energy released in this process is captured as ATP, the direct source of chemical energy for most physiological processes. The substrate for oxidation in the Krebs cycle is *acetate,* a two-carbon compound that enters the Krebs cycle in an activated form known as *acetyl-CoA.* Fig. 32.1 illustrates that glucose and fatty acids are the sources of acetate for oxidation. Selecting between these two sources of acetate is a major function of fuel homeostasis.



• **Fig. 32.1** Relationship of the three major oxidative, catabolic pathways. *CO_2,* carbon dioxide; *KG,* α-ketoglutarate; *OA,* oxaloacetic acid; *P,* Phosphate.

The Fuels

The Major Metabolic Fuels Consist of Glucose, Amino Acids, Fatty Acids, and Ketone Bodies

Glucose, amino acids, fatty acids, and ketone bodies are the compounds that can be directed into the Krebs cycle for energy production. Fuel homeostasis is the coordinated process by which these fuels are stored, mobilized, and interconverted to assure a continuous supply of energy for the body.

Glucose Is the Central Fuel in the Energy Metabolism of Most Animals

Glucose, the digestion product of carbohydrate, is the basic metabolic fuel during periods of adequate nutrition in omnivorous monogastric animals, such as dogs and rats. Although there are other important fuels in the body, glucose has special significance because, under most conditions, it is the only fuel consumed by the CNS. Therefore maintaining a steady supply of glucose for brain metabolism is of paramount importance to the body. It is not surprising that an elegant system of homeostasis exists to regulate the availability of glucose to the brain and other tissues. This system of maintaining glucose availability is a major focus of this chapter.

Glucose can be stored in the body as glycogen, a highly branched starch found in liver and skeletal muscle. Glycogen is the only direct storage form of glucose in the body. Directing glucose to and from glycogen depots is a major function of fuel homeostasis. Glucose is released from glycogen through the process of *glycogenolysis.*

The first step through which glucose is used as a fuel is through *glycolysis,* the series of biochemical steps that initiate the oxidation

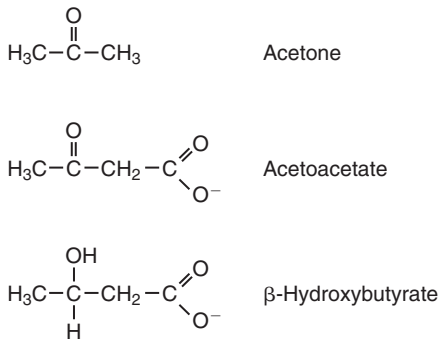
of glucose. Through glycolysis, each molecule of glucose is converted to two molecules of *pyruvate,* a key molecule that stands at the junction of several metabolic pathways. When glucose is completely oxidized for fuel, pyruvate is converted to acetate (acetyl CoA) and passes into the *Krebs cycle,* the site of its complete oxidation. For the study of fuel homeostasis, you should appreciate that the conversion of glucose to pyruvate is reversible. Therefore any metabolic pathway that can lead to the production of pyruvate can also lead to the creation of glucose. Fig. 32.1 illustrates a key relationship between pyruvate, acetate, and oxaloacetic acid. The conversion of pyruvate to acetate is irreversible whereas the conversion of pyruvate to oxaloacetate can flow in either direction. Thus any metabolic pathway that can lead to the creation of pyruvate or oxaloacetate can lead to the synthesis of glucose. Creation of glucose through such pathways is called *gluconeogenesis.* The process of gluconeogenesis occurs in the liver and to a small extent also in the kidneys. It occurs in no other tissues. Another pathway for glucose oxidation is the *pentose-phosphate pathway.* This is a quantitatively minor pathway that does not have great effect on fuel homeostasis. However, it is an important metabolic pathway in erythrocytes (RBCs), which have an absolute need for glucose, although these cells' overall need for energy is small compared with the rest of the body.

Amino Acids Are Important Fuels in Addition to Being the Building Blocks of Protein

Amino acids are important fuels. These monomers, the building blocks of proteins, are also carbon-containing compounds that can provide energy to the body. In addition, they are important substrates for gluconeogenesis, indicating that most amino acids can be converted to glucose when the available glucose supply is limited. Although it is sometimes said that there is no storage site of amino acids in the body, the protein of skeletal muscle could be considered to have an amino acid storage function in addition to its locomotor functions.

Fatty Acids Are the Major Form of Energy Storage in the Animal Body

Fatty acids are stored in adipose tissue in the form of *triglycerides* (also called *triacylglycerols*), which consist of three fatty acid molecules linked to a glycerol molecule by ester bonds (see Fig. 30.24). Triglycerides are an ideal form of energy storage for animals. They are highly reduced molecules (there is little oxygen compared with the amount of carbon and hydrogen), which means they are a concentrated energy source, having more than twice the caloric value per gram than carbohydrates or amino acids. In addition, adipose tissue contains little water compared with muscle or liver tissue, the sites of storage for amino acids and glycogen, the other two potential fuels. Thus adipose tissue is undiluted by bulky water, allowing it to be a concentrated form of energy storage that permits animals to carry with them a maximal amount of energy at a minimal amount of weight. Fats, however, have a metabolic disadvantage; they are not water soluble. Therefore special transport systems are needed to enable fats to be distributed among the tissues through the blood and lymph systems. In addition, fatty acids cannot be converted to glucose, so they cannot, under usual circumstances, contribute to the energy supply of the CNS. However, fatty acids can be converted to *ketone bodies.*



• **Fig. 32.2** Physiological ketone bodies.

Ketone Bodies Are Fat-Derived, Water-Soluble Metabolites That Serve as Glucose Substitutes

Although glucose cannot be formed from fat, the fat-derived ketone bodies do have some glucose-like attributes. For example, ketone bodies can pass the blood-brain barrier. During prolonged periods of dietary energy deprivation, they can provide a large portion of the energy supply to the CNS, at least in some species. It does appear, however, that ketone bodies cannot totally replace glucose in this function, and that a small amount of glucose is always needed by the CNS.

In monogastric species, ketone bodies are formed exclusively in the liver and are used by a wide variety of tissues. Some tissues, including cardiac muscle, use ketone bodies instead of glucose. In ruminants, the ketone body β-hydroxybutyrate is formed from butyrate in the rumen epithelium. Thus in ruminants, ketone bodies are not only products of fatty acid metabolism but also products of normal digestion. Elevated serum concentrations of ketone bodies are characteristic of several diseases associated with abnormalities of fuel homeostasis. This fact might lead one to conclude that ketone bodies are abnormal, or even toxic, metabolites. In fact, when present in physiological concentrations, ketone bodies are important fuels that occupy an integral part of the scheme of fuel homeostasis. Fig. 32.2 illustrates the chemical structure of the three major ketone bodies.

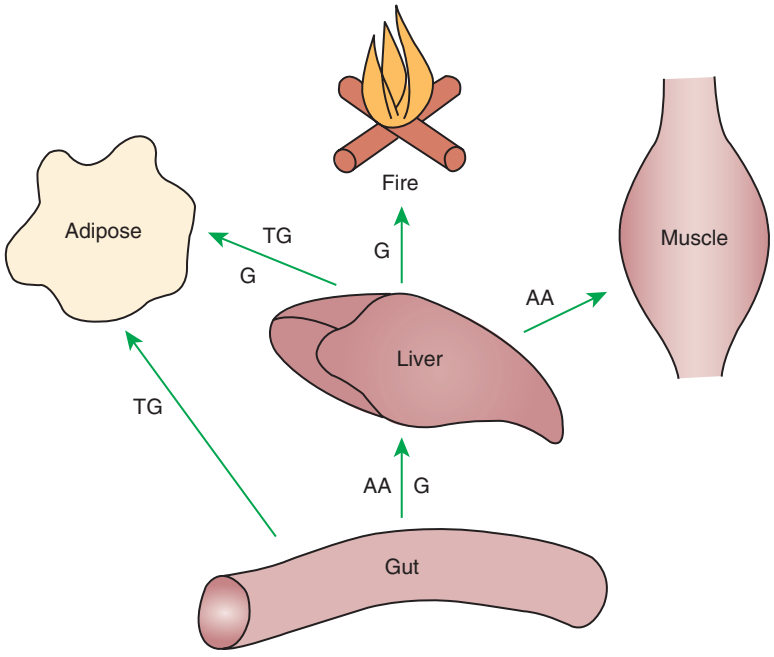
Nutrient Utilization During the Absorptive Phase

The discussions in this chapter will divide fuel metabolism into three phases: (1) an absorptive phase associated with the active digestion and absorption of nutrients from the gut, (2) a postabsorptive phase that occurs during intermeal intervals when nutrients are not being absorbed from the gut, and (3) a prolonged energy deficiency or food deprivation phase.

During the absorptive phase as nutrient absorption takes place, metabolic events in the liver and peripheral organs are coordinated to direct nutrients into storage molecules and storage sites. Fig. 32.3 illustrates the general scheme of metabolism during the absorptive phase.

During the Absorptive Phase, the Liver Takes Up Glucose and Converts It Into Glycogen and Triglyceride

When a meal is ingested, insulin secretion begins even before maximal absorption of glucose is achieved. This secretion is



• **Fig. 32.3** Metabolism during the absorptive period is characterized by the movement of potential fuels into depot sites and the utilization of glucose (*G*) as a fuel. *AA,* Amino acid; *TG,* triglyceride.

stimulated by the action of gastric inhibitory peptide (see Chapter 27) and perhaps other enteric hormones. Early insulin secretion ensures that the liver and other tissues will be "primed" and ready for the arrival of glucose from the gut. A large portion of the glucose absorbed postprandially is taken up by the liver as portal blood traverses the hepatic sinusoids. Under the influence of insulin, glucose in the liver is directed into glycogen synthesis. The net effect is that glucose from the digestion and absorption of carbohydrate is stored in the liver during absorptive periods. This process moderates the flow of glucose from the gut to the general circulation and keeps blood glucose concentrations from becoming excessively high during the digestion of a carbohydrate meal. Insulin exerts its stimulatory effect on hepatic glycogen synthesis by stimulating intracellular metabolic pathways that lead to the formation of glycogen. These effects are discussed further in reference to the counterbalancing effects of glucagon.

The amount of glycogen that can be stored in the liver is limited and, under normal conditions, probably never exceeds 10% of the total weight of the liver. In humans, this represents about 100 g of glycogen, and a proportionately similar limit likely exists for the storage of glycogen in the livers of other species. This amount of glycogen does not account for all the glucose taken up by the liver during the digestion and absorption of a large carbohydrate meal; therefore some additional mechanism must exist for the disposal of excess glucose. If there were no such alternatives for glucose disposal, blood glucose concentrations could rise out of control after glycogen stores had reached their maximum. Fatty acid synthesis offers an alternative mechanism for glucose removal.

The Conversion of Glucose to Fatty Acids Is an Irreversible Process

The synthesis of fatty acids from glucose begins with glycolysis. This pathway leads to the production of two pyruvate molecules for each molecule of glucose consumed. Pyruvate can then enter the mitochondria to be activated to *acetyl coenzyme A* (acetyl CoA) for entry into the Krebs cycle, as discussed previously. However, the Krebs cycle is for energy generation, and during the absorptive period, there is more than enough acetyl CoA and Krebs cycle activity to provide for energy needs; therefore the excess acetyl CoA must be shunted away from the Krebs cycle. The excess acetyl CoA combines with *oxaloacetate* to form *citrate* in what is essentially the first reaction of the Krebs cycle. Instead of continuing through the Krebs cycle reactions, however, during the absorptive period much of the citrate is transported out of the mitochondria into the cytosol. When in the cytosol, each citrate molecule contributes two carbons toward the synthesis of fatty acids. The remaining portion of the citrate molecule cycles back into the mitochondria for further use. Citrate serves as a carrier molecule to transport two-carbon units out of the mitochondria because acetyl CoA cannot pass the mitochondrial membrane directly (Fig. 32.4).

Several important steps in this conversion of glucose to fatty acids are promoted by insulin and are discussed in detail later. It is important to recognize that the conversion of glucose to fatty acids is irreversible; thus carbohydrate can form fat, but fat cannot form carbohydrate. The discussion here concerns hepatic metabolism, and the liver is an important site of fatty acid synthesis in several species. Direct synthesis of fatty acids also occurs in adipose tissue. The relative importance of liver and adipose tissue as sites of fatty acid synthesis varies with species, as discussed later.

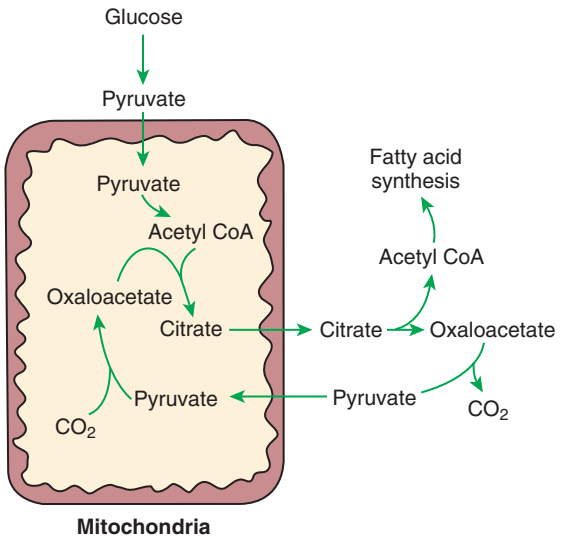
Transport of Fatty Acids out of the Liver Is Through Chylomicron-Like Particles Known as Very-Low-Density Lipoproteins

When formed in the liver, fatty acids must be transported either to adipose tissue for storage or to other tissues (e.g., muscle) for direct utilization for energy production. Because fatty acids are insoluble in blood, some special transport mechanism for their distribution is necessary. This mechanism is through the hepatic formation of triglyceride-rich serum lipoproteins, also known as *very-low-density lipoproteins* (VLDLs); these triglyceride-rich lipoproteins are much less dense than other lipoproteins in blood serum. In the synthesis of VLDLs, fatty acids are first esterified to form triglycerides, and the triglycerides are wrapped in a coat of phospholipid, cholesterol, and specific proteins (Fig. 32.5). This

is essentially the same mechanism by which fatty acids are transported out of the enterocytes after absorption from the gut. In the latter case, the lipoproteins are called *chylomicrons.* The VLDLs of the liver are smaller than chylomicrons but have a similar structure and function. The mechanisms by which VLDLs and chylomicrons deliver fatty acids to peripheral tissues are further discussed in relation to peripheral tissues.

Amino Acids Can Be Classified Into Groups on the Basis of Metabolic Characteristics

The discussion of amino acid absorption and metabolism becomes complicated because not all amino acids are subject to the same reactions. For this discussion, the amino acids are divided into two groups, each containing two subgroups (Table 32.1). The major groups are the "nutritionally dispensable" amino acids and the "nutritionally indispensable" amino acids. Within the dispensable amino acid group, glutamate, aspartate, alanine, glutamine, and asparagine are separated out as *transport amino acids;* within the indispensable amino acid group, leucine, isoleucine, and valine form a special subgroup known as the *branch-chain amino acids* (BCAAs). The transport amino acids are used in several reactions in which amino groups are transferred from molecule to molecule or organ to organ.



• **Fig. 32.4** Hepatic synthesis of fatty acid from carbohydrate requires the passage of carbohydrate carbons through the mitochondria. Citrate forms a shuttle to transport the carbons of acetyl coenzyme A (*acetyl CoA*) out of the mitochondria because acetyl CoA cannot pass directly through the mitochondrial membrane. The formation of citrate from oxaloacetate and acetyl CoA is the first reaction of the Krebs cycle; thus fatty acid formation is an alternative to Krebs cycle oxidation when there is more than enough acetyl CoA to provide cellular energy through Krebs cycle activity.

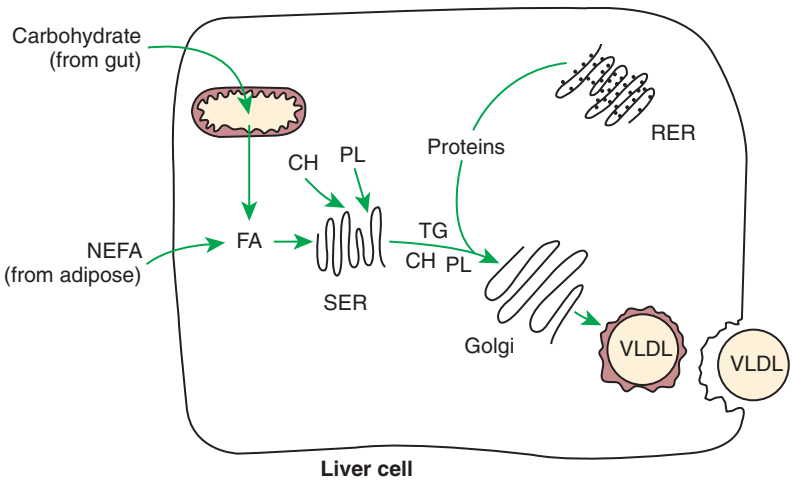
Amino Acids Are Extensively Modified During Absorption

The profile of amino acids in the portal vein is considerably different from that of the diet, indicating that amino acid destruction and transformation occur during the absorptive process. Essentially all the glutamate and much of the aspartate in the diet are removed by the intestinal epithelial cells during absorption, so the portal blood is almost devoid of glutamate and contains little aspartate. Much of the nitrogen from glutamate and aspartate is transferred to pyruvate to form the amino acid alanine, which is present in high concentrations in portal blood. The metabolism of the transport amino acids in the intestinal epithelium is a good example of both the way in which amino groups can be gained and lost and how the metabolism of amino acids interfaces with the metabolism of carbohydrate. Glutamate and aspartate are similar to two Krebs

TABLE 32.1 Metabolic Classification of Amino Acids

Indispensable Amino Acids		Dispensable Amino Acids	
Branch-Chain Amino Acids	Others	Transport Amino Acids	Others
Leucine	Arginine[a]	Alanine	Cysteine
Isoleucine	Histidine	Glutamine	Glycine
Valine	Lysine	Glutamate (Glutamic acid)	Proline
	Methionine	Asparagine	Tyrosine[b]
	Phenylalanine	Aspartate (Aspartic acid)	Serine
	Threonine		
	Tryptophan		

[a]Indispensable for cats, but not required in the diets of many other species.
[b]Dietary adequacy depends on a supply of phenylalanine.

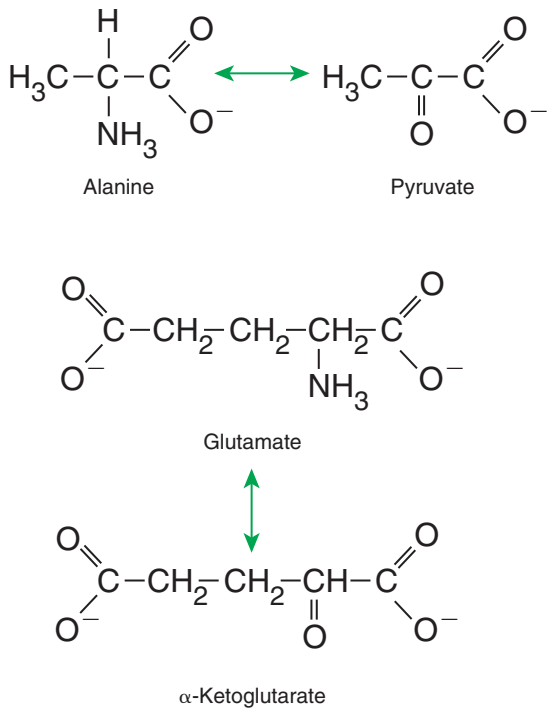


• **Fig. 32.5** Formation of very-low-density lipoprotein (*VLDL*). Fatty acids (*FA*) for triglyceride (*TG*) formation may come from synthesis by carbohydrate or amino acids or from adipose tissue FA arriving at the liver in the form of nonesterified fatty acids (*NEFA*). Note the similarity to chylomicron formation (see Fig. 30.27). *CH,* Cholesterol; *PL,* phospholipid; *RER,* rough endoplasmic reticulum; *SER,* smooth endoplasmic reticulum.

cycle intermediates, α-ketoglutarate and oxaloacetate, differing only by the presence of an amino group or a keto-oxygen. Carbohydrates and amino acids having this relationship are said to be *analogs;* thus α-ketoglutarate is the keto-analog of glutamate, and pyruvate is the keto-analog of alanine (Fig. 32.6). All amino acids can form keto-analogs, and all keto-analogs can be readily converted back to their parent amino acids.

Many Amino Acids Are Removed by the Liver on "First Pass," Never Reaching the Systemic Circulation

The hepatic-portal circulation is arranged in such a way that all nutrients leaving the gut via the blood pass through the liver before entering the systemic circulation (see Fig. 30.23). This arrangement places the liver in a "sentinel" position, from which it can modify the nutrient composition of portal blood before the blood is distributed to other tissues. The function of the liver in modifying portal blood composition is well illustrated in the case of amino acid absorption. Many of the amino acids absorbed into portal blood are removed as the blood passes the liver, so they never reach the general circulation. Fig. 32.7 illustrates that, in the dog, only about 23% of the amino acids reaching the liver during the absorptive period pass into the general circulation; the liver thus helps keep blood amino acid concentrations stable during periods of amino acid absorption. The blood amino acid concentration, as with the blood glucose concentration, is usually kept relatively constant.



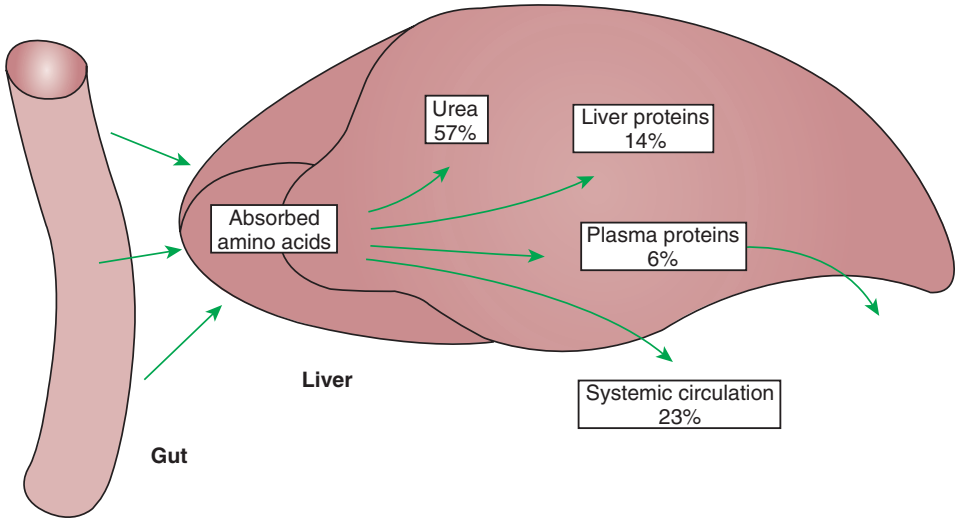
• **Fig. 32.6** Example of amino acids and their keto-analogs. All amino acids can reversibly form keto-analogs.

Some Amino Acids Taken Up by the Liver Are Used for Protein Synthesis

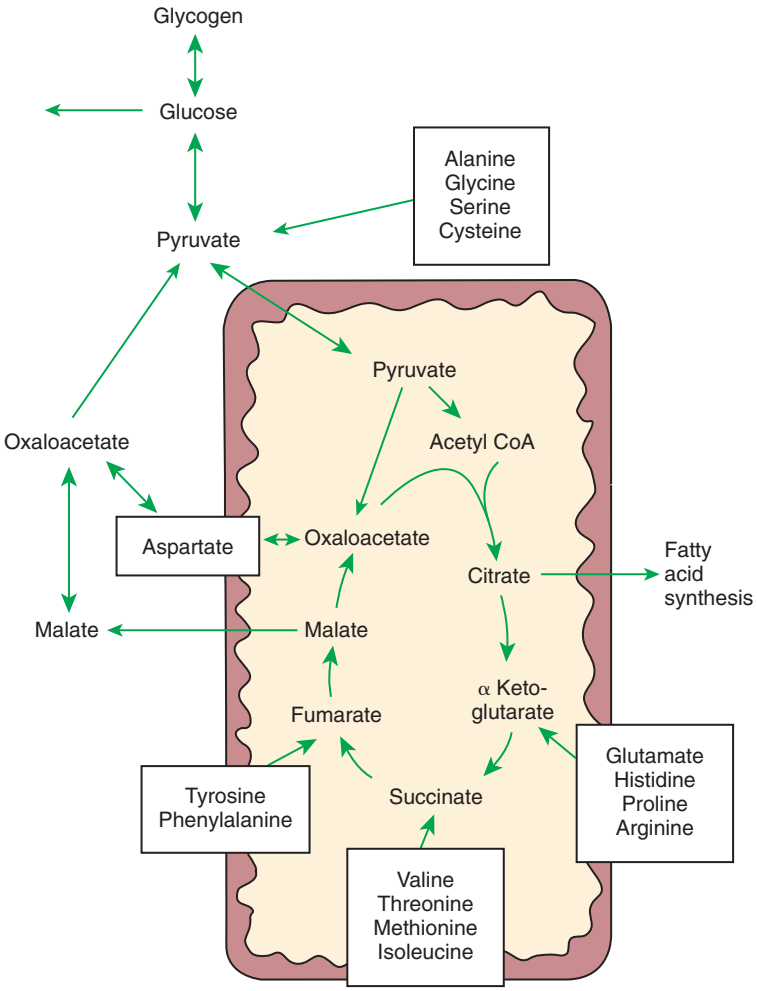
The liver is an important site of protein synthesis, and thus its priority position for amino acid uptake seems reasonable. Fig. 32.7 shows that approximately 20% of the portal blood amino acid supply is used for protein synthesis in the liver, although this proportion varies with dietary protein intake. Almost all the serum proteins are synthesized in the liver, including such critical proteins as albumin and the blood-clotting factors. Although the liver-derived serum proteins serve many important functions, one function they *do not* serve is that of amino acid transport. The direct amino acid supply for protein synthesis in nonhepatic tissue comes from free amino acids in the blood, not from preformed serum proteins.

Most Amino Acids Taken Up by the Liver Are Converted to Carbohydrates

Most amino acids entering the liver undergo *deamination,* which means the amino groups are removed and the molecules converted to their keto-analogs. The keto-analogs enter the pathways of carbohydrate metabolism, from which they may be completely metabolized for energy, converted to glucose or glycogen, or shunted to fatty acid synthesis. All these reactions proceed in the same manner as previously described for carbohydrate metabolism. Fig. 32.8 illustrates the sites at which the various amino acids enter the carbohydrate pathways.



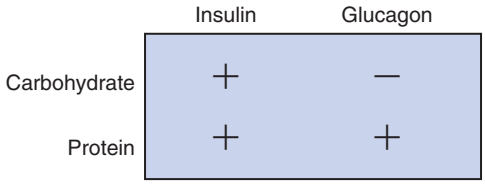
• **Fig. 32.7** Fate of dietary amino acids reaching the canine liver.



• **Fig. 32.8** Sites of entry of various amino acids into the scheme of carbohydrate metabolism. This figure illustrates the means by which glucose can be synthesized from amino acids in the process of gluconeogenesis. In the case of the dispensable amino acids, the reactions are reversible, allowing amino acid production from carbohydrate.

Deamination of amino acids for the production of carbohydrate or energy may seem like a waste of expensive dietary protein; in some species, however, the deamination of amino acids is important for homeostasis of glucose and other fuels. For example, the natural diets of the true carnivores (e.g., cats, mink) contain a large portion of protein and little carbohydrate, but their glucose needs are no less than those of other animals, so it is extremely important they synthesize glucose from amino acids. Ruminants are in a similar situation because most of the carbohydrates they consume undergo fermentative digestion and are absorbed as volatile fatty acids (VFAs) rather than glucose. As with carnivores, ruminants depend on amino acids for some of their glucose needs, although a large portion of ruminant glucose requirements may be met through conversion of propionate.

To allow carbohydrate production and the deamination of excess amino acids, the endocrine reactions to high-protein meals are somewhat different from those to meals containing substantial amounts of carbohydrate. During the digestion of high-protein meals, insulin and glucagon secretion does not occur in its usual reciprocal pattern. Insulin secretion is stimulated by amino acids and by glucose. Glucagon secretion, which is inhibited by glucose, is stimulated by amino acids as long as glucose concentrations are moderately low. This relationship means that during the digestion of a high-protein, low-carbohydrate meal, there is simultaneous secretion of insulin and glucagon. One of the effects of insulin is the greater cellular uptake of amino acids and glucose. Thus the effect of insulin in this situation is to increase transport of amino acids into tissues.



• **Fig. 32.9** Influence of dietary carbohydrate and protein on insulin and glucagon secretion.

If insulin secretion were the only action stimulated by amino acid absorption, however, the animal would risk insulin-stimulated hypoglycemia when it consumed a high-protein, low-carbohydrate diet. An important action of glucagon is to stimulate gluconeogenesis through the deamination of amino acids in the liver. This process ensures that adequate glucose will be available to counterbalance the effects of amino acid–stimulated insulin secretion. Fig. 32.9 illustrates the relationship between insulin secretion and glucagon secretion during the absorption of diets with different carbohydrate and protein concentrations.

Not All Amino Acids Are Subject to Hepatic Destruction

During the absorptive period, amino acids for peripheral (nonhepatic) protein synthesis must come from that portion of amino acids that escape hepatic destruction. As seen from Fig. 32.7, this portion amounts to only about 23% of the amino acids absorbed from the gut. Although this may seem like a meager portion of amino acids to be allocated for protein synthesis by all body tissues except the liver, two considerations make it more appropriate. First, amino acids are selectively taken up by the liver, so the distribution of individual amino acids in blood leaving the liver is not the same as that in blood reaching the liver. The indispensable amino acids, especially the BCAAs, are not avidly extracted by the liver, whereas some of the dispensable amino acids (e.g., alanine) are extensively taken up by hepatic tissue. The dispensable amino acids can be synthesized by protein-producing tissues; thus the relatively low concentration of serum amino acids resulting from hepatic amino acid removal is not rate limiting for tissue protein synthesis. Second, the proportion of amino acids taken up by the liver, and the fate of the amino acids that are taken up, is not constant and can be adjusted according to the body's protein needs. Low-protein diets lead to reductions in hepatic amino acid uptake, protein synthesis, and amino acid destruction by the liver.

Metabolism at the Tissue Level Is Coordinated With Hepatic Metabolism and Results in the Deposition of Fuel Into Storage Tissues During the Absorptive Period

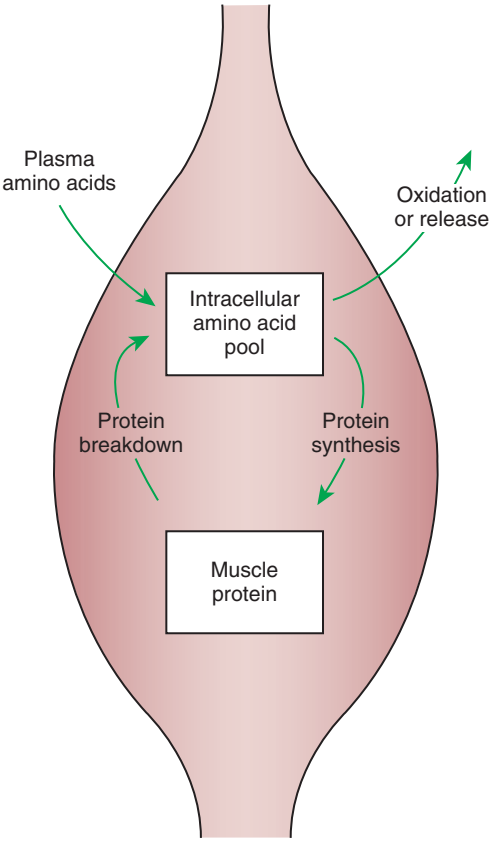
The overall effects of hepatic metabolism during the absorption of a meal are the removal of glucose and amino acids and the synthesis of protein and fat. Complementary changes occur in peripheral tissues, so additional glucose and amino acids are removed by skeletal muscle and adipose tissue. In addition, fatty acids secreted by the liver as VLDL triglycerides are deposited in adipose tissue, as are the triglycerides of chylomicrons.

Insulin Promotes the Synthesis of Protein and the Deposition of Glycogen in Muscle

The absorptive period is dominated by the effects of insulin. In skeletal muscle, the largest tissue mass of the body, insulin promotes the uptake of glucose and amino acids and thus tends to moderate the increase in blood concentration of these nutrients during absorption of a meal. The uptake of glucose by muscle is associated with glycogen synthesis, just as in the liver. Muscle glycogen, in contrast to liver glycogen, cannot be made directly available to augment blood glucose concentrations during periods of low glucose availability. Muscle glycogen is primarily for metabolism in the muscle.

Insulin-Stimulated Uptake of Amino Acids by Muscle Results in a Net Increase in Muscle Protein Synthesis

The term *net increase* is used in reference to muscle protein synthesis because muscle protein is in a state of dynamic equilibrium, that is, a constant state of flux. Protein molecules are continuously being broken down and their amino acids added to an intracellular amino acid pool. Simultaneously, new proteins are constantly being made, deriving their amino acids from the same pool (Fig. 32.10). The size of the amino acid pool depends on the relative rates of entry and exit of amino acids. Amino acids enter the pool from the blood during the absorptive phase and at all times from the breakdown of body protein. Exit of amino acids from the pool results from protein synthesis and oxidative catabolism. In the absorptive phase of digestion, the amino acid pool is large because amino acids are being taken up from the blood. In addition, few of the amino acids leaving the pool are directed toward oxidative catabolism because sufficient glucose is available for oxidation and energy generation. Therefore the amino acid pool is large, and a high proportion of amino acids is directed to protein synthesis. When the rate of protein synthesis exceeds the rate of protein breakdown, there is a net increase in the amount of muscle protein. Thus during the absorptive phase, amino acids are stored in muscle protein, protein that has a functional role not only for locomotion and posture but also for amino acid storage.



• **Fig. 32.10** Intracellular amino acid pool. The size of the pool depends on the rates of amino acid uptake from plasma and muscle protein relative to the rates of amino acid loss resulting from oxidation, export to plasma, and protein synthesis.

During the Absorptive Phase, Triglyceride Accumulation in Adipose Tissue Occurs by Two Mechanisms: Uptake From Very-Low-Density Lipoproteins and Direct Lipid Synthesis From Glucose

Triglyceride fatty acids are transferred from chylomicrons and VLDLs to adipose tissue by the action of *lipoprotein lipase* (LPL). This enzyme resides on endothelial surfaces of capillaries and,

when activated, binds to chylomicrons and VLDLs, catalyzing the hydrolysis of fatty acids from their core triglycerides and allowing the transfer of those fatty acids to the surrounding tissues. The sensitivity of LPL to specific hormones varies in different tissues. Adipose tissue LPL is stimulated by insulin; thus during the absorptive phase, fatty acids from chylomicrons and VLDLs are selectively transferred to adipose tissue. Therefore under the influence of insulin, excess carbohydrate and amino acids are converted to fatty acids in the liver, and those fatty acids are subsequently transported via VLDLs to the adipose tissue. Similarly, chylomicron triglyceride arising from intestinal fatty acid absorption is also selectively transported to adipose tissue, under the influence of insulin.

Adipose tissue fatty acids may also arise from direct synthesis in addition to uptake from chylomicrons and VLDLs. Adipose tissue cells are metabolically active, and under the influence of insulin, they take up glucose. Within the adipocytes, glucose can be converted to fatty acids by the same metabolic mechanisms by which fatty acids were synthesized in the liver. In addition, acetate from fermentative digestion also can serve as a substrate for fatty acid synthesis in adipose tissue (see later discussion of the special fuel considerations of ruminants). Thus there are two major sites of fatty acid synthesis in the body: liver and adipose tissue. The relative importance of these sites varies with species.

Nutrient Utilization During the Postabsorptive Phase

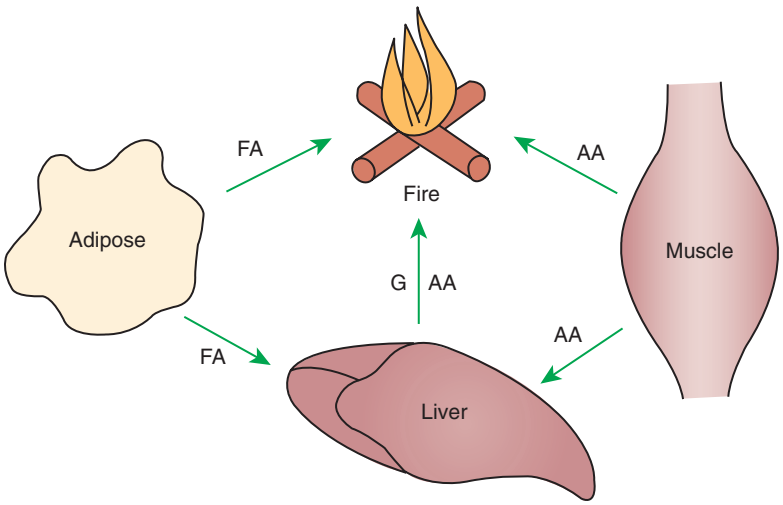
The *postabsorptive phase* is the relatively brief period (usually a few hours) between meals in well-fed animals. It is characterized by short-term changes that mobilize nutrients from storage pools to maintain fuel availability for metabolically active tissue. Fig. 32.11 illustrates the general scheme of postabsorptive metabolism.

Hepatic Metabolism Switches From Glucose Utilization to Glucose Production During the Postabsorptive Phase

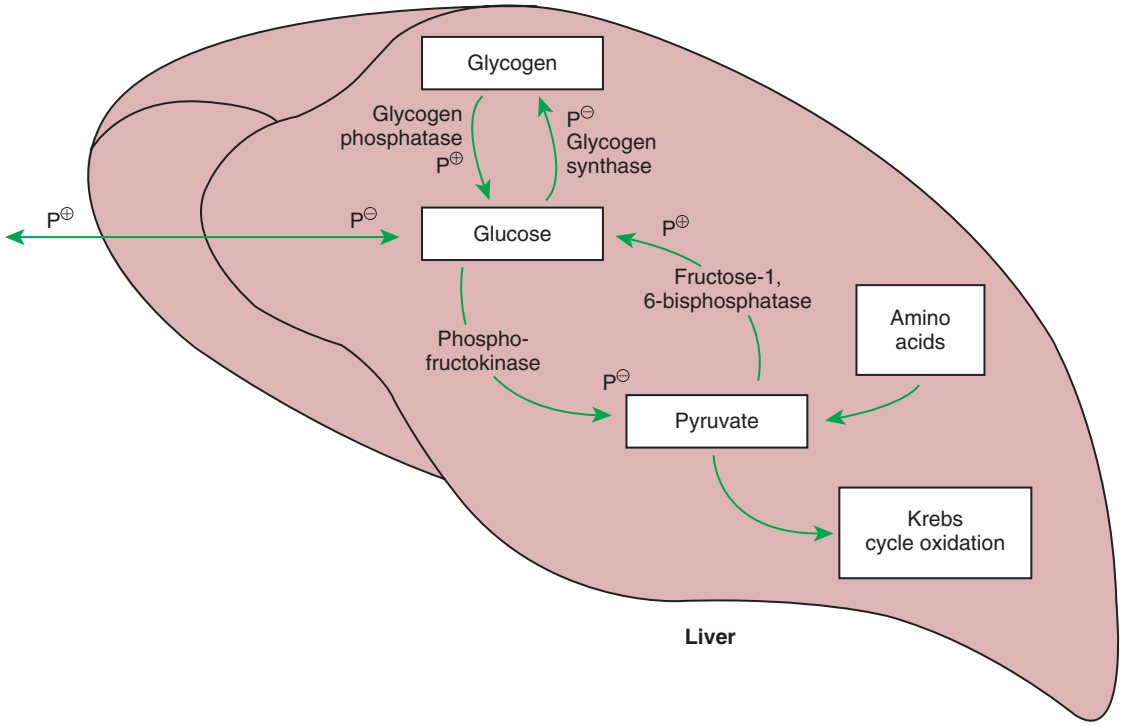
As the absorption of a meal is completed, the rate of glucose absorption from the gut wanes, and the blood glucose concentration diminishes, removing the stimulus for insulin production. As blood glucose concentrations decline, glucagon secretion is stimulated. The primary target organ of glucagon is the liver, in which glucagon creates marked metabolic changes. Through stimulation of specific cell surface receptors on hepatocytes, glucagon activates adenyl cyclase, leading to the phosphorylation of numerous cellular enzymes (see Chapter 1). Some enzymes are activated by phosphorylation, whereas others are inactivated, and unless the overall scheme of substrate flow is considered, the whole phosphorylation-dephosphorylation system appears to be quite random and to make little sense. Considering the actions of the individual enzymes in light of their effect on the flow of energy substrate through the liver, however, reveals that the system is an elegant and incredibly well-orchestrated mechanism for the maintenance of fuel homeostasis.

The enzymes that stimulate mobilization and utilization of fuels are activated by phosphorylation, whereas those stimulating storage of fuels are inactivated by phosphorylation. It must be understood that many enzymes of intermediary metabolism serve a passive role, catalyzing reactions that can go in either direction, depending on substrate concentrations. A relatively small number of regulatory enzymes usually stand at the head of metabolic pathways and determine the substrate concentrations to which the other, unregulated enzymes are exposed. Through its effect on several key regulatory enzymes, glucagon (a stimulator of phosphorylation) places the liver in a fuel-mobilization state. In contrast, insulin (an inhibitor of phosphorylation) promotes a hepatic metabolic pattern that favors fuel storage, as previously discussed.

The opposing actions of insulin and glucagon on hepatic metabolism are evident from their actions on two key regulatory enzyme pairs: *glycogen synthase* and *glycogen phosphatase,* and *phosphofructokinase* and *fructose-1,6-bisphosphatase.* The first of these pairs regulates glycogen synthesis and breakdown, whereas the second regulates glycolysis and gluconeogenesis, respectively. Fig. 32.12 illustrates the actions of these enzymes and their regulatory effects. Glycogen synthase and phosphofructokinase are inhibited by phosphorylation and thus are stimulated by insulin. Glycogen phosphatase and fructose-1,6-bisphosphatase are stimulated by phosphorylation and thus stimulated by glucagon. The actions of insulin and glucagon on these antagonistic enzyme pairs emphasize the importance of the insulin/glucagon ratio to which the liver is exposed. Neither hormone elicits an "all-or-none" reaction, but rather alters the balance of opposing reactions by influencing the relative activity of antagonistic enzymes. Thus the fuel-mobilizing



• **Fig. 32.11** Postabsorptive metabolism is characterized by movement of fuels out of depot sites for immediate use. Glucose (*G*) arising from glycogenolysis or gluconeogenesis is a major fuel, although some fatty acid (*FA*) is consumed as well. Amino acid *(AA)* forms the substrate for gluconeogenesis.



• **Fig. 32.12** Effects of phosphorylation on four key enzymes of glucose production and utilization. All four enzymes are phosphorylated under the influence of cyclic adenosine monophosphate (*cAMP*). Note, however, that the enzymes that favor glucose formation are stimulated by phosphorylation (P^+), whereas those that favor glucose utilization and storage are inhibited by phosphorylation (P^-).

or fuel-storing activity of the liver depends on which hormone is most dominant. For this reason, the insulin/glucagon ratio appears to be more important to liver metabolism than the absolute concentration of either hormone.

Under the influence of glucagon, glycogen phosphatase is activated by phosphorylation, promoting glycogenolysis and the elevation of intracellular glucose concentrations. As glucose accumulates, it is prevented from cycling back into glycogen because the major enzyme catalyzing that reaction, glycogen synthase, is blocked by phosphorylation. In addition, the flow of glucose into glycolysis is also blocked by phosphorylation inhibition of phosphofructokinase (see Fig. 32.12). Thus the normal pathways for glucose utilization within the hepatocyte are all inhibited by glucagon, allowing glucose from glycogen breakdown to accumulate in the cells. Eventually, intracellular glucose escapes into the extracellular fluid and on into the blood. In this manner, hepatic glycogen is mobilized to elevate and maintain blood glucose concentrations when they begin to decline.

The liver stores of glycogen are relatively limited and cannot maintain blood glucose concentrations for a long period. Estimates in humans are that hepatic glycogen will serve blood glucose needs for 6 to 12 hours under conditions of light exertion and for only about 20 minutes under conditions of heavy exertion. Values for animals are probably similar. Therefore in addition to glycogen mobilization, some other means must exist for maintaining the body's glucose supply during periods of exertion and prolonged periods between meals. Under these conditions of greater demand, glucose is provided by gluconeogenesis. Gluconeogenesis is promoted by the phosphorylation-stimulated enzyme fructose-1,6-bisphosphatase. This enzyme essentially puts the glycolytic pathway into reverse, leading to glucose production from the same molecules that are intermediates in its oxidative destruction. Important substrates are pyruvate and all the intermediates of the Krebs cycle.

At this point, it is important to remember that most of the Krebs cycle intermediates or pyruvate can be supplied by the deamination of amino acids. The entry point of the various amino acids into the scheme of carbohydrate metabolism is illustrated in Fig. 32.8. Pyruvate and all the Krebs cycle intermediates can flow backward through the oxidative pathway (not all the reactions of gluconeogenesis are the exact reverse of the corresponding reactions in glycolysis, but the net result of gluconeogenesis is the reverse of glycolysis), resulting in the production of glucose. Thus amino acids provide a large store of precursors for glucose formation. The end result of glucagon stimulation is to promote the production of glucose through glycogenolysis and gluconeogenesis, turning the liver into a glucose-synthesizing organ.

Fuel Mobilization in Peripheral Tissues Occurs When the Blood Insulin Concentration Declines

The pattern of metabolism in the peripheral tissues changes in the postabsorptive phase to support the liver's capacity to maintain fuel supplies.

Muscle Reacts to a Metabolic Demand for Glucose by Mobilizing Amino Acids to Support Hepatic Gluconeogenesis

Mobilization of amino acid from muscle appears to be stimulated largely by a relative lack of insulin; thus mobilization occurs when blood glucose concentrations are low. Amino acids mobilized from

skeletal muscle come from the intracellular amino acid pool cited earlier (see Fig. 32.10). However, the mobilizing reactions are complex, and the distribution of amino acids leaving the muscle does not reflect the distribution of amino acids in the intracellular pool, as explained later.

Muscle Release of Amino Acids Is Related to Reduced Glucose and Amino Acid Uptake

The postabsorptive decline in the serum insulin concentration has a twofold effect on muscle: the entry of amino acids from the serum into the intracellular amino acid pool is diminished and the entry of glucose into muscle cells for energy production declines. Reduced amino acid entry results in conditions favoring net protein degradation to maintain the cellular amino acid pool size. Reduced glucose entry results in increased utilization of amino acids from the pool for energy production.

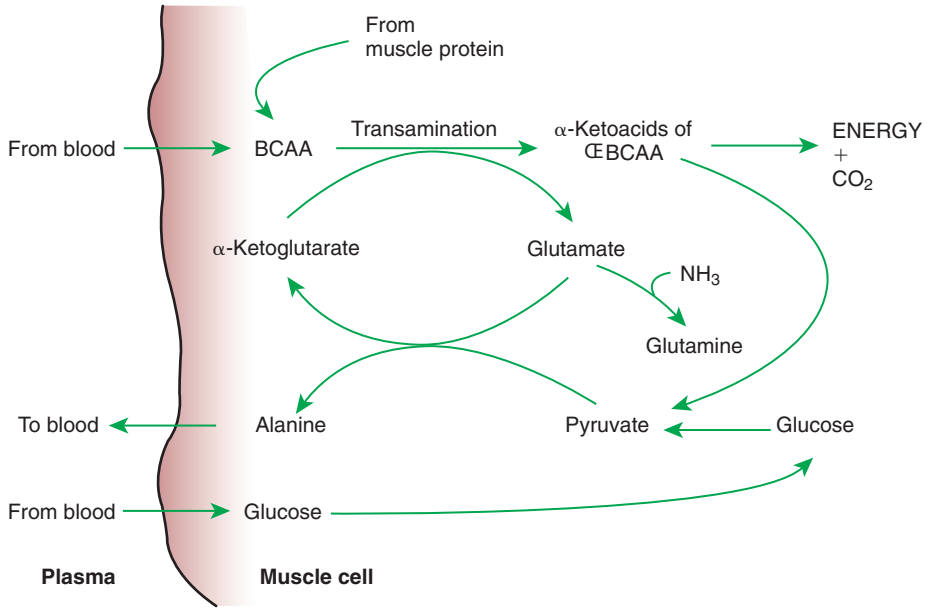
The pattern of utilization of amino acids for energy by muscle may at first seem unnecessarily complex, involving selective use and extensive transformation of amino acids. BCAAs serve as primary sources of energy in muscle cells during the postabsorptive phase because these amino acids account for approximately one-third of all muscle amino acid. Catabolism of BCAAs begins with deamination and the formation of the α-keto-acid of the BCAA. The α-keto-acids then enter the Krebs cycle for energy production. Deamination of the BCAA requires that some acceptor be available to receive the amino group, and this acceptor is ultimately pyruvate, resulting in the formation of alanine. The source of pyruvate can be muscle glycogen, blood glucose, or the metabolic products of BCAA α-keto-acids. When metabolism of BCAA α-keto-acids serves as the supply of pyruvate for alanine synthesis, the net reaction is conversion of BCAA to alanine (Fig. 32.13). Thus the overall metabolic activity in muscle during the postabsorptive phase is the destruction of BCAAs and the formation of alanine. The alanine formed is released from the muscle cells into the blood, from which it may be taken up by the liver for gluconeogenesis.

The Complex Pattern of Muscle Amino Acid Catabolism and Release Is Necessary to Accommodate the Liver's Limited Capacity for Uptake of Branch-Chain Amino Acids and to Facilitate the Removal of Amino Nitrogen From the Muscle

It might appear that a simpler system of amino acid transfer to the liver would suffice. Why are amino acids not just released from muscle cell amino acid pools into the blood and transported to the liver for glucose synthesis? The answer lies in the limited uptake capacity of the liver for BCAAs and the need to transport amino nitrogen out of the muscle. BCAAs, the predominant amino acids of skeletal muscle, are not taken up readily by the liver; thus if BCAAs were not transformed to alanine, amino acid transfer to the liver would be limited.

In addition, alanine is a convenient means by which nitrogen from the deamination of muscle amino acid can be transported to the liver. This is important because free amino groups liberated by the catabolism of amino acids in muscle, if not removed, could lead to the formation of toxic levels of ammonia. Ammonia is detoxified in the body by the formation of urea, but urea formation occurs only in the liver. Thus alanine forms a gluconeogenic precursor that also transports nitrogen to the liver for urea synthesis. Figs. 32.13 and 32.14 illustrate the role of alanine in the transport of amino acid nitrogen and carbon to the liver for synthesis of urea and glucose, respectively.

The regulation of muscle protein mobilization is influenced to a large extent by the lack of insulin. However, the adrenocortical hormone cortisol has an important effect of stimulating protein breakdown and amino acid mobilization. Through the mobilization of muscle protein and stimulation of hepatic gluconeogenesis, cortisol exerts one of its major effects, raising blood glucose concentration. Under normal conditions, glucagon, the other major gluconeogenic hormone, exerts its effects on the liver and does not appear to have a direct effect on muscle.

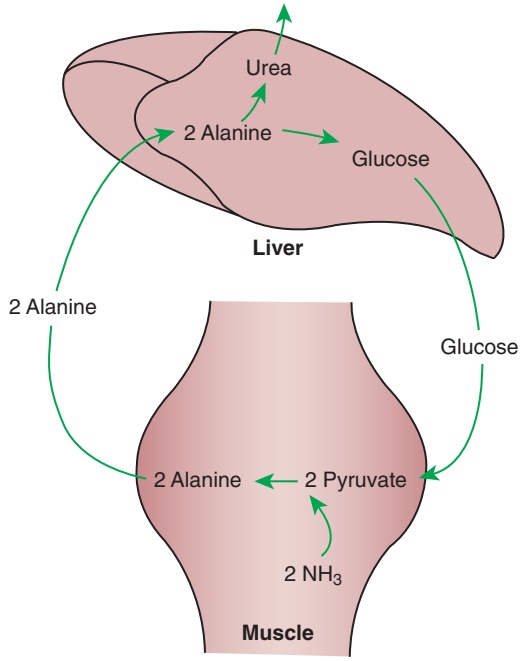


• **Fig. 32.13** Catabolism of branch-chain amino acids (*BCAAs*) by muscle cells. The pyruvate for export of amino groups may be derived from glucose or the amino acids themselves.

The Reaction of Adipose Tissue During the Postabsorptive Phase Is to Mobilize Fatty Acids

Fatty acids are released from adipose tissue because of the action of the phosphorylation-stimulated enzyme *hormone-sensitive lipase* (HSL). This enzyme is stimulated by the relative lack of insulin in the postabsorptive period; insulin suppresses the action of HSL by promoting its dephosphorylation. Glucagon may have some adipose tissue activity in promoting triglyceride breakdown by stimulating the phosphorylation and activation of HSL. More likely, however, glucagon's effects are restricted to the liver, and the normal stimulation of HSL comes from epinephrine or norepinephrine; norepinephrine originates from sympathetic nerves in the adipose tissue. The exact means by which sympathetic nerve activity in adipose tissue is coordinated with body fuel availability is not well established, but the catecholamine hormones and neuroregulators appear to be the primary positive stimulus for breakdown of adipose triglyceride. However, the negative stimulus provided by the absence of insulin may be the most important regulator of adipose fat mobilization.

Stimulation of HSL in the postabsorptive state leads to the release of fatty acids from adipose tissue into the blood. Fatty acids in blood are reversibly bound to albumin because they are not otherwise soluble in water. Albumin-bound fatty acids in blood are usually referred to as *nonesterified fatty acids* (NEFAs) to distinguish them from triglyceride fatty acids in chylomicrons and lipoproteins. NEFAs in blood may be used directly for energy by many tissues. However, many NEFAs are taken up by the liver and used for either ketone body production or VLDL synthesis, as discussed in the next section.

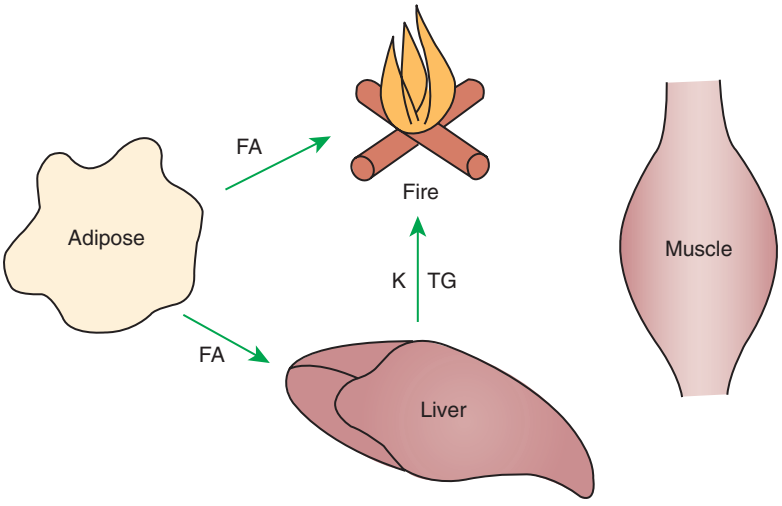


• **Fig. 32.14** Alanine arising from branch-chain amino acid (BCAA) catabolism in muscle is converted to glucose and urea in the liver. The glucose produced can potentially return to the muscle for alanine production. Thus the cycle of alanine to glucose forms a shuttle to transport nitrogen from the muscle to the liver for urea synthesis. NH_3, Ammonia.

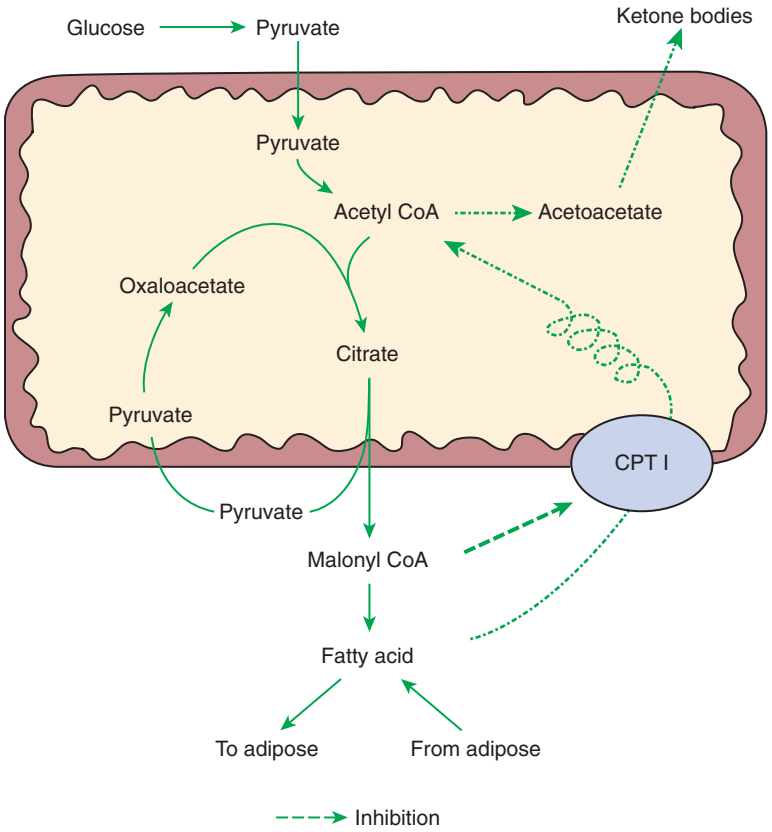
Nutrient Utilization During Prolonged Energy Malnutrition or Complete Food Deprivation

During Prolonged Periods of Fasting or Undernutrition, Glucose and Amino Acids Are Conserved by Extensive Utilization of Fats and Ketone Bodies for Energy Production

From the previous discussion of postabsorptive metabolism, you can appreciate that amino acids form an important depot for glucose precursors and energy-producing substrate. During prolonged fasting or undernutrition, however, it would not be advantageous for animals to rely heavily on their skeletal muscle for energy and glucose production; doing so would soon lead to severe weakness as the skeletal muscle protein was consumed. Thus protective mechanisms have developed by which skeletal muscle is preserved during periods of insufficient energy intake. In utilization of stored fuels, shifts away from glucose and toward adipose fat stores are necessary for protein sparing. Fig. 32.15 illustrates



• **Fig. 32.15** During prolonged periods of food deprivation or energy deficiency, the ketone bodies (*K*), fatty acids (*FA*), and triglycerides (*TG*) become the major fuels. Glucose oxidation becomes minor, thus sparing muscle protein that otherwise would be needed for gluconeogenesis.



• **Fig. 32.16** The liver is a site of both destruction and synthesis of fatty acids. To keep both processes from occurring simultaneously, fatty acid destruction is inhibited during periods of fatty acid synthesis. *Solid lines,* pathway of fatty acid synthesis; *irregular broken line,* fatty acid destruction. Oxidative destruction is suppressed by the action of malonyl CoA, an intermediate in the synthesis of fatty acid. Malonyl CoA blocks the transport of fatty acids into the mitochondria at the translocation enzyme carnitine palmitoyltransferase I (*CPT I*).

the general scheme of metabolism during prolonged catabolic periods.

A Large Portion of the Fatty Acids Released From Adipose Tissue Are Taken Up Directly by the Liver

During prolonged periods of undernutrition, low glucose availability leads to rapid mobilization of adipose fatty acids in the form of NEFAs. Although NEFAs are metabolized by various tissues, many are extracted from the blood by the liver, which receives much of the total blood flow and has an efficient hepatic NEFA extraction mechanism. When the NEFAs are in the hepatocytes, they may follow any of three potential metabolic paths. The first is complete oxidation for energy production; however, the hepatic requirements for energy are such that only a small amount of the total fatty acid supply during adipose mobilization needs to be used for complete oxidation. The second pathway is esterification leading to triglyceride formation, and the third is production of ketone bodies. Triglyceride synthesis is discussed later; the focus here is ketone body production.

Hepatic Ketone Body Formation Is Promoted by Low Glucose Availability, a High Glucagon/Insulin Ratio, and a Ready Supply of Fatty Acids

Ketone body formation occurs within the hepatic mitochondria, and the rate of ketone body synthesis is controlled by the regulated transport of fatty acids across the mitochondrial membrane (Fig. 32.16). Fatty acids enter mitochondria in combination with a molecule known as *carnitine,* and transport depends on an enzyme known as *carnitine palmitoyltransferase I* (CPT-I). The activity of this enzyme, along with the availability of fatty acid, is the primary determinant of the rate of ketone body formation. CPT-I activity is regulated in an interesting manner, being inhibited by an intermediate of the fatty acid synthesis pathway, *malonyl CoA.* Malonyl CoA concentrations are high when the liver is responding to insulin and glucose is being used for fatty acid synthesis. When glucagon concentrations are high relative to insulin, little fatty acid is synthesized in the liver. Thus malonyl CoA concentrations are low, and CPT-I is fully active when the insulin/glucagon ratio is low. Ketone body synthesis is stimulated under these hormonal conditions.

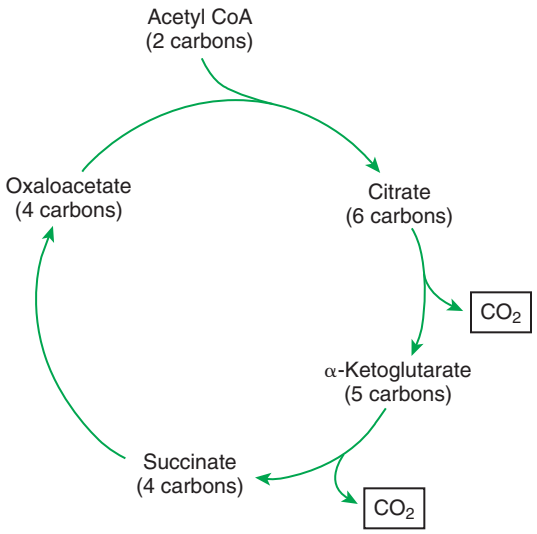
Under conditions when CPT-I is active, most available fatty acid is transported into the mitochondria for ketone body synthesis. This well-orchestrated but somewhat complex regulatory system is important because the liver can both produce and consume fatty acids. If there were not a way of "turning off" fatty acid destruction during periods of synthesis, a futile cycle of synthesis and destruction would occur. The inhibition of CPT-I by malonyl CoA provides a system that blocks the metabolic destruction of newly synthesized fatty acid while still providing a mechanism for the utilization of fatty acids derived from adipose tissue. The overall pattern of metabolism results in a reciprocal relationship between glucose availability and ketone body production. Although ketone bodies are produced in the liver, they cannot be used there for energy production. Therefore all ketone bodies are transported to peripheral tissues for utilization. When the concentration of ketone bodies in the blood becomes abnormally high, some are excreted in urine.

Glucagon Plays an Important Role in the Excessive Production of Ketone Bodies in Diabetes Mellitus

If untreated, diabetes mellitus in animals, especially dogs, leads to high concentrations of ketone bodies in the blood. Diabetes mellitus occurs because of a lack of insulin, but the hepatic production of ketone bodies results from the unrestrained action of glucagon. Even though serum concentrations of glucose are high in diabetes mellitus, the inability of the pancreas to secrete insulin leads to a low insulin/glucagon ratio; thus the liver is functioning solely under the direction of glucagon. Glucagon inhibits fatty acid production from glucose, so malonyl CoA concentrations are low and CPT-I activity is high. Because of the lack of insulin to suppress adipose HSL, blood NEFA concentrations are high. The combination of high NEFA availability and unrestrained CPT-I activity results in rapid transport of fatty acids into the mitochondria with extensive ketone body production, even though blood glucose concentrations are high.

Fatty Acids Cannot Be Used for Glucose Synthesis

It is important to understand that the metabolism of fat within the mitochondria cannot contribute directly to gluconeogenesis. When they cross the mitochondrial membrane, fatty acids undergo β-oxidation, which leads to the successive removal of two-carbon acetyl CoA units from the carbon chains of the fatty acids. The resulting acetyl CoA can enter the Krebs cycle through condensation with oxaloacetate. Because any of the Krebs cycle intermediates can lead to glucose production, it may first appear that acetyl CoA from fatty acid β-oxidation could lead to the production of glucose. However, this is not the case; there is no *net* production of oxaloacetate associated with the consumption of acetyl CoA by the Krebs cycle (Fig. 32.17). Existing oxaloacetate combines with acetyl CoA to form citrate in the initial step of the cycle. At the end of the cycle, the original oxaloacetate is reformed while the two carbons from the acetyl CoA are converted to carbon dioxide. No new oxaloacetate can be produced by this process.



• **Fig. 32.17** Oxidation of acetyl CoA (from acetate) by the Krebs cycle. The two carbons of acetyl CoA result in the formation of carbon dioxide; there is no net synthesis of oxaloacetate. Because oxaloacetate forms the precursor for glucose synthesis, acetyl CoA (and thus acetate) cannot lead to glucose formation.

Ketone Bodies Are Formed in the Mitochondria From Acetyl Coenzyme A

Not all mitochondrial acetyl CoA must enter the Krebs cycle. In fact, when fatty acids are rapidly entering the mitochondria, there is much more acetyl CoA available than necessary for Krebs cycle activity. It is this excess acetyl CoA, originating from fatty acids, from which the ketone bodies are synthesized (see Fig. 32.16). Ketone bodies are able to leave the mitochondria freely.

Ketone bodies affect fuel homeostasis in peripheral tissues, where they may serve as a substitute for glucose. In this way, they conserve available glucose and reduce the need for gluconeogenesis.

Hepatic Very-Low-Density Lipoproteins May Be Synthesized From Adipose-Derived Fatty Acid and From Newly Synthesized Fatty Acid

The section on absorptive-phase metabolism discusses the hepatic production of VLDL. During the absorptive phase, triglyceride for VLDL synthesis comes from fatty acids synthesized from glucose. During catabolic periods, VLDLs may continue to be produced, but fatty acids derived from serum NEFAs are used for VLDL synthesis (see Fig. 32.5). This may initially appear to be an unnecessary and inefficient metabolic step. Why should fatty acids from adipose tissue be transported to the liver for VLDL formation when they can be directly metabolized for energy by the tissues? The need for VLDL synthesis occurs because of the need for a better transport system. The capacity of the serum to transport NEFA is limited because NEFA must circulate bound to albumin. The NEFA-binding capacity of albumin is finite and may become almost saturated during periods of rapid adipose mobilization. VLDLs provide a transport system for fatty acids that is independent of serum albumin.

Hormonal Conditions Direct the Distribution of Very-Low-Density Lipoprotein Fatty Acids in the Body

During the absorptive phase, VLDLs are directed to adipose tissue by the action of adipose tissue LPL, an insulin-stimulated enzyme. LPL also exists in muscle tissue but does not depend on insulin stimulation for activity. Thus during periods of low glucose availability, adipose tissue LPL is inhibited because of a lack of insulin, but muscle tissue LPL is fully active. This situation leads to the selective direction of VLDL fatty acids to muscle tissue during times of adipose mobilization.

Changes in Growth Hormone Concentrations May Aid in Shifting Peripheral Fuel Utilization From Glucose and Amino Acids to Ketone Bodies and Fatty Acids

The fat mobilization–induced changes in hepatic metabolism are effective in conserving protein only because of changes that occur in glucose and amino acid utilization in peripheral tissues. As ketone bodies, NEFAs, and VLDL triglycerides become the major energy supplies, there is less tissue demand for glucose or amino

acids as energy substrates. Endocrine alterations, in addition to low insulin concentrations, may aid in promoting this switch in peripheral fuel utilization. In several species, growth hormone concentrations rise during a prolonged period of energy deprivation. Growth hormone is antagonistic to insulin, thus promoting an increase in the serum glucose concentration even in the presence of normal or near-normal serum insulin levels. In addition, growth hormone may have some direct effect on conserving protein and mobilizing lipid.

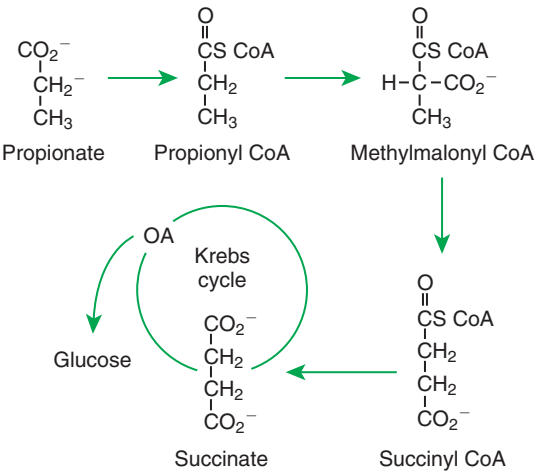
Special Fuel Considerations of Ruminants

Ruminants Exist in a Perpetual State of Gluconeogenesis Because of Their Unique Digestive Process

Most carbohydrate digestion in ruminants occurs in the forestomach through fermentative digestion. The result is that on diets with a relatively high proportion of forage, almost no digestible carbohydrate enters the intestine for glandular digestion and absorption as glucose. Therefore ruminants exist in a constant state of potential glucose deficiency. To cope with this situation, ruminants have developed efficient systems of both production and conservation of glucose.

Essentially, all the glucose available to ruminants with typical diets originates from gluconeogenesis. Quantitatively, the most important glucose precursor is the VFA propionate. Propionate contributes to glucose synthesis after entering the Krebs cycle at the level of succinate (Fig. 32.18). Note that succinate is a four-carbon Krebs cycle intermediate that can lead to net formation of oxaloacetate, the entry metabolite for gluconeogenesis. The other VFAs, acetate and butyrate, also enter the Krebs cycle, although they enter as acetyl CoA. As previously discussed, acetyl CoA cannot lead to the net production of oxaloacetate or glucose. Therefore of the ruminant's major energy sources—acetate, propionate, and butyrate—only propionate can support glucose production. Almost all propionate absorbed from the rumen is extracted from the portal blood by the liver, never entering the systemic circulation.

In addition to constant gluconeogenesis, ruminants also support their glucose needs by efficiently conserving glucose. Fatty acids are synthesized in the liver of some animals (e.g., primates, rats, dogs) but only in the adipose tissue of ruminants. Furthermore, glucose is essentially not used for fatty acid synthesis. Rather, fatty acids are synthesized from acetate, which is the most abundant energy source in ruminants. The only glucose used by adipose tissue is for the synthesis of the glycerol backbone for triglycerides. In lactating animals, fatty acids produced in the udder for milk fat are synthesized from either acetate or ketone bodies, never from glucose.



• **Fig. 32.18** Gluconeogenesis from propionate involves its initial conversion to succinate. Succinate is a four-carbon Krebs cycle intermediate that can lead to net glucose synthesis.

Some important metabolic diseases of ruminants occur during periods when their system of glucose homeostasis is stressed. Dairy cows are especially vulnerable in early lactation because the synthesis of lactose (milk sugar) requires glucose. In high-producing cows, nearly all the glucose they produce goes to lactose synthesis, whereas the remaining tissues function on alternative fuels. Sheep experience a similar stress on glucose synthesis in late gestation. The energy needs of the fetus and placenta can be met only by glucose (or glucose-derived lactate) and amino acids. Compared with many other animals, sheep have a high ratio of fetal mass to body size; thus their fuel homeostatic mechanisms are particularly stressed by pregnancy. Failure of the glucose homeostatic mechanism frequently occurs under these circumstances, resulting in conditions known as *lactational ketosis* in dairy cows and *pregnancy toxemia* in ewes.

CLINICAL CORRELATIONS

Hepatic Lipidosis in a Cat

History

You are asked to examine a 3-year-old intact female cat. She apparently was normal, even fat and happy, until 2 weeks ago, when she disappeared from her owner's apartment for 4 days. When she returned, the cat seemed depressed and would not eat. Over the next few days, she became progressively more listless, almost somnolent.

Clinical Examination

The cat has a normal pulse, temperature, and respiratory rate, but she is depressed and responds little to handling. The ocular sclerae (whites of the eyes) appear icteric, or jaundiced. The latter physical sign leads you to suspect liver disease, so you submit blood samples for biochemical analysis. Analysis of blood taken from the jugular vein reveals a higher-than-normal concentration of bile acids and bilirubin, confirming a diagnosis of liver disease. A needle aspiration biopsy of the liver reveals hepatocytes that are distended with large droplets of nonstaining material, probably fat.

Comment

The presence of significant concentrations of bile acids in blood, other than in the hepatic-portal circulation, is evidence of reduced liver function. Recall that bile acids are absorbed from the ileum into the portal vein, in which they return to the liver. The normal liver extracts bile acids from portal blood efficiently, allowing only small amounts to escape into the systemic circulation; thus elevated concentrations of bile acids in jugular blood indicate liver disease.

Hepatic lipidosis, or "fatty liver," is a common disease of cats, initiated by a period of stress combined with either an unwillingness to eat or a lack of available food. In either situation, the cat begins to mobilize large quantities of fat to support metabolic energy needs. Normally, much of the mobilized

Continued

CLINICAL CORRELATIONS—cont'd

NEFA would be expected to be taken up by the liver and converted to VLDL for export to energy-using tissues. In cats that experience fatty liver, the hepatic influx of NEFA appears to overwhelm the liver's capacity to synthesize and secrete VLDL, so fat accumulates in the liver. When the fat accumulation becomes severe, hepatic function is compromised, and the cat becomes systemically ill. Appetite becomes severely depressed, leading to a downward spiral of events in which the hepatic lipidosis becomes more and more severe.

Treatment

Treatment consists of reversing the state of negative energy balance by force-feeding. Various methods of force-feeding exist; the most practical approach is placement of an indwelling gastric tube. The tube is often passed through the nostrils but may be placed by a number of different techniques, including direct intubation through the wall of the abdomen. The latter technique is facilitated by the use of a fiberoptic gastroscope. When the cat is in positive energy balance, adipose mobilization ceases, and the liver eventually clears of fat. Tube feeding may need to continue for several days before the cat begins to eat on its own. Tube feeding has greatly improved the prognosis for hepatic lipidosis, although it is still a life-threatening condition.

Hyperlipemia in a Horse

History

A client calls you concerned about his 25-year-old Morgan mare that is depressed and has not been eating or drinking well for the past few days. It is the middle of the winter. The mare has poor teeth, and she has been losing weight this winter, but the client think the weight loss has been more dramatic in the past few weeks.

Clinical Examination

The mare's temperature, pulse, and respiratory rate are normal. She appears mentally dull, and she is dehydrated. The feces look somewhat dry. Her mucous membranes are yellow (icteric). She has decreased gastrointestinal borborygmi in all quadrants. On palpation per rectum, she has some dry feces, but no other abnormalities are detected. On nasogastric intubation, she does not have any reflux. Based on her age, inappetance, and weight loss, you are concerned that she might be hyperlipidemic/hyperlipemic. You submit blood work for complete blood count (CBC) and biochemical profile including triglycerides. You give her some water via the nasogastric tube, and you float her remaining teeth. You recommend that they feed her an equine senior diet (which is a highly digestible feed).

Laboratory Results

Her clinical chemistry demonstrates increased glucose (170 mg/dL), triglycerides (TG) (550 mg/dL), and bilirubin (2.5 mg/dL) concentrations. A CBC reveals an increased white blood cell count (leukocytosis) characterized by increased neutrophils (neutrophilia), monocytosis (increased monocytes), and decreased lymphocytes (lymphopenia).

Comment

Negative energy balance as a result of decreased intake with current or increased calorie demands results in increased lipolysis of the adipose tissue. This results in the mobilization of adipose fat in the form of NEFAs. Circulating NEFA may be used for energy by a wide variety of tissues, but a large proportion is extracted from serum by the liver. In the liver, NEFA can potentially be used for complete oxidization, ketone body synthesis, or reesterification to triglycerides. In horses, a large portion of NEFA arriving at the liver are reesterified. The resulting TG are either stored in the liver or released to the circulation as components of VLDLs. Removal of VLDL from the blood is regulated by LPL, an enzyme that, especially in adipose tissue, is insulin dependent. When the rate of hepatic VLDL secretion exceeds their removal, TG accumulates in the blood serum. Various terms including *hyperlipidemia, hypertriglyceridemia,* and *hyperlipemia* are used to describe this condition. The use of the specific terms is usually dependent on the severity of the serum lipid accumulation.

Treatment

If there is a primary cause of the inappetance (i.e., colic, neoplasia), which can create a catabolic state, it needs to be identified and treated first. For this horse, it was likely the poor teeth, which decreased the ability of the horse to chew hay, and thus the horse was not eating sufficient calories to equal its metabolic needs. Additionally, during the winter, many horses will drink less; the horse may have more easily become dehydrated. This could lead to an impaction colic, which would be an additional reason why the horse was not eating sufficient calories. Basal metabolism may also be increased as it is winter. Thus all of these factors (i.e., teeth, decreased drinking, colic) are likely causes of the decreased intake and increased caloric needs. It is also possible that the mare has neoplasia. Regardless of the primary cause, because of the decreased intake, it led to the catabolic state, thus precipitating the hyperlipidemia/hyperlipemia. Treatment is aimed at treating both the primary cause and the hyperlipidemia/hyperlipemia. The horse can be given good quality grass when available and the senior diet. Based on her blood work, there are no overt signs of neoplasia. She could be treated based on the possible issues of teeth, decreased intake, and increased metabolic demands for thermoregulation as the primary causes, and her progress should be monitored. For the hyperlipidemia/hyperlipemia, she can be given insulin and maybe glucose. Insulin increases gluconeogenesis, both by inhibiting HSL, which causes lipolysis of adipose tissue, and increasing LPL. Thus adipose storage in the liver will be decreased. Depending on the severity of disease and initial causes, this treatment may be sufficient. Many horses require in-clinic care with intravenous fluids and heparin to promote insulin sensitivity through increased activity of LPL, while decreasing activity of HSL. Overall, these horses need careful monitoring, whether on the farm or in the clinic, as it can take some period of time to correct hyperlipidemia/hyperlipemia, including addressing the caloric/nutritional needs and addressing the primary cause.

Practice Questions

1. All the following are metabolites that can be oxidized for fuel in the animal body. Which one is *not* important in the transport of energy between organs and organ systems?
- **a.** Triglyceride
- **b.** Ketone bodies
- **c.** Oxaloacetic acid
- **d.** Nonesterified fatty acids
- **e.** Amino acids

2. Which of the following reactions is not characteristic of the absorptive phase of digestion?
- **a.** Hepatic synthesis of glycogen
- **b.** Hepatic uptake of glucose
- **c.** Destruction of dietary amino acid
- **d.** Utilization of muscle-derived amino acid for gluconeogenesis
- **e.** Hepatic synthesis of triglyceride from glucose

3. Which of the following reactions in the liver could be expected to occur during both the digestive phase and a prolonged fast?
a. Glycogen synthesis
b. Fatty acid synthesis
c. Ketone body synthesis
d. Ketone body oxidation
e. Triglyceride synthesis from fatty acids

4. Which of the following statements is true of both ketone bodies and nonesterified fatty acids?
a. They are water soluble.
b. They provide energy for muscle metabolism.
c. They circulate in blood bound to albumin.
d. They can provide energy to the brain.
e. They are formed exclusively in the liver.

5. Which of the following amino acids is not extensively catabolized by the liver?
a. Valine
b. Alanine
c. Glutamine
d. Asparagine
e. Glycine

Bibliography

Entry-Level Supplements

Engelking LR. *Textbook of Veterinary Physiological Chemistry*. 2nd ed. Sections I, III, V, VI. Elsevier; 2015.

Advanced Reading

Aschenbach JR, Kristensen NB, Donkin SS, et al. Gluconeogenesis in dairy cows: the secret of making sweet milk from sour dough. *IUBMB Life*. 2010;62(12):869–877.

Barton MH, Divers TJ. Disorders of the liver. In: Reed SM, Bayly WM, Sellon DC, eds. *Equine Internal Medicine*. 4th ed. St Louis: Elsevier; 2018.

Bauman DE, Currie WB. Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorrhesis. *J Dairy Sci*. 1980;63(9):1514–1529.

Brody T. *Nutritional Biochemistry*. 2nd ed. San Diego: Academic Press; 1999.

Cheeke PR, Dierenfeld ES. *Comparative Animal Nutrition and Metabolism*. Cambridge, Mass: CABI; 2010.

Hand MS, Thatcher CD, Remillard RL, et al. *Small Animal Clinical Nutrition*. 5th ed. Topeka, Kan: Mark Morris Institute; 2010.

Herdt TH. Fuel homeostasis in the ruminant. *Vet Clin North Am Food Anim Pract*. 1988;4(2):213–231.

Kaneko JJ, Harvey JW, Bruss ML. *Clinical Biochemistry of Domestic Animals*. 6th ed. Burlington, MA: Elsevier, 2008.

Keane K, Newsholme P. Metabolic regulation of insulin secretion. *Vitam Horm*. 2014;95:1–33.

McCue MD. Starvation physiology: reviewing the different strategies animals use to survive a common challenge. *Comp Biochem Physiol A Mol Integr Physiol*. 2010;156(1):1–18.

McKenzie HC III. Equine hyperlipidemias. *Vet Clin North Am Equine Pract*. 2011;27(1):59–72.

Nordlie RC, Foster JD, Lange AJ. Regulation of glucose production by the liver. *Annu Rev Nutr*. 1999;19:379–406.

Puchalska P, Crawford PA. Multi-dimensional roles of ketone bodies in fuel metabolism, signaling, and therapeutics. *Cell Metab*. 2017;25(2):262–284.

Storey KB, ed. *Functional Metabolism: Regulation and Adaptation*. Ottawa: John Wiley & Sons; 2004.

Wang T, Hung CC, Randall DJ. The comparative physiology of food deprivation: from feast to famine. *Annu Rev Physiol*. 2006;68:223–251.

33

The Endocrine System

BRIAN K. PETROFF AND DEBORAH S. GRECO

KEY POINTS

General Concepts

1. Hormones are chemicals produced by specific tissues that are transported by the vascular system to affect other tissues at low concentrations.
2. The endocrine and nervous systems are integrated in their control of physiological processes.

Synthesis of Hormones

1. Protein hormones are initially synthesized as preprohormones and then cleaved in the rough endoplasmic reticulum to form prohormones, and in the Golgi apparatus to form the active hormones, which are stored in granules before being released by exocytosis.
2. Steroids are synthesized from cholesterol, which is synthesized by the liver; steroids are not stored but are released as they are synthesized.

Transport of Hormones in the Blood

1. Protein hormones are hydrophilic and carried in the plasma in dissolved form.
2. Steroids and thyroid hormones are lipophilic and carried in plasma in association with both specific and nonspecific binding proteins; the amount of unbound, active hormone is relatively small.

Hormone-Cell Interaction

1. Protein hormones have specific receptors on target tissue plasma membranes, whereas steroids have specific receptors within the cytoplasm or nucleus.

Postreceptor Cell Responses

1. Steroids interact directly with the cell nucleus through the formation of a complex with their cytosolic or nuclear receptors, whereas protein hormones need a messenger because they cannot enter the cell.

Metabolism of Hormones

1. Steroid hormones are metabolized by conjugation with sulfates and glucuronides, which makes steroids water soluble.

Feedback Control Mechanisms

1. The most important feedback control for hormones is the negative-feedback system, in which increased hormone concentrations result in less production of the hormone, usually through an interaction with the hypothalamus or pituitary gland.
2. Endocrine secretory patterns can be influenced by factors such as sleep or light and can produce circadian rhythms.

The Hypothalamus

1. The hypothalamus coordinates the activity of the pituitary gland through the secretion of amines and peptides, particularly releasing hormones.

The Pituitary Gland (Hypophysis)

1. The neurohypophysis has cell bodies that originate in the hypothalamus, with cell endings that secrete oxytocin and vasopressin (antidiuretic hormone).
2. Oxytocin and vasopressin are synthesized in cell bodies within the hypothalamus and are carried by axon flow to the posterior lobe, where they are released.
3. The main effects of oxytocin are on the contraction of smooth muscle (mammary gland and uterus); the effects of vasopressin are primarily on the conservation of water (antidiuresis) and secondarily on blood pressure.
4. Plasma osmolality controls the secretion of vasopressin.
5. The adenohypophysis produces growth hormone, prolactin, thyroid-stimulating hormone, follicle-stimulating hormone, luteinizing hormone, and corticotropin.
6. Adenohypophyseal activity is controlled by hypothalamic releasing hormones, which are released into the portal system, which in turn connects the median eminence of the hypothalamus and the anterior pituitary gland.

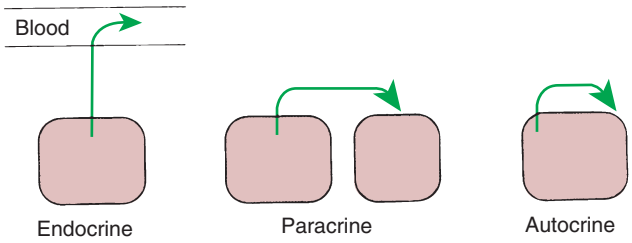
General Concepts

Hormones Are Chemicals Produced by Specific Tissues That Are Transported by the Vascular System to Affect Other Tissues at Low Concentrations

The endocrine system has evolved to allow physiological processes to be coordinated and regulated. The system uses chemical messengers called *hormones*. Hormones have traditionally been defined as "chemicals that are produced by specific endocrine organs, are transported by the vascular system, and are able to affect distant target organs in low concentration." Although this definition is useful from a veterinary medical point of view, it should be recognized that some substances, such as prostaglandins and somatomedins, are produced by many other tissues and are still considered hormones.

Other types of control systems use chemical substances that are not transported in the vascular system to influence distant cell activity. These systems serve as means of local integration among or between cells, as follows:

- *Paracrine effectors,* in which the messenger diffuses through the interstitial fluids, usually to influence adjacent cells; if the messenger acts on the cell of its origin, the substance is called an *autocrine effector* (Fig. 33.1).
- *Neurotransmitters*, which affect communication between neurons, or between neurons and target cells; the substances are limited in the distance traveled and the area of the cell influenced (Fig. 33.2).
- *Exocrine effectors,* such as hormones produced by the pancreas, are released into the gastrointestinal tract.



• **Fig. 33.1** Types of cell communication via chemical mediators. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

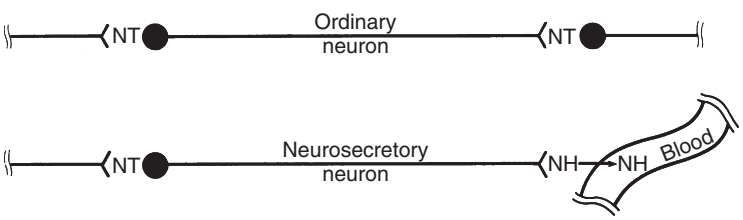
The Endocrine and Nervous Systems Are Integrated in Their Control of Physiological Processes

The endocrine system interacts with the other main regulatory system, the *nervous system*, which coordinates activities that require rapid control. An example of the close interaction of the two systems is the reflex in which suckling causes the release of milk. Suckling initiates the transmission of nerve impulses from the mammary gland to the hypothalamus (by way of the spinal tract). Neurosecretory neurons within the supraoptic and paraventricular nuclei are stimulated to synthesize *oxytocin*. Oxytocin is transported along axons of these nerves and is released from nerve endings in the posterior pituitary into the blood vascular system. Oxytocin is then carried to the mammary gland where it causes contraction of myoepithelial cells. These cells surround the smallest unit of milk-secreting cells, called an *alveolus*. This results in the movement of milk into the large cisternae adjacent to the teat and subsequently into the teat.

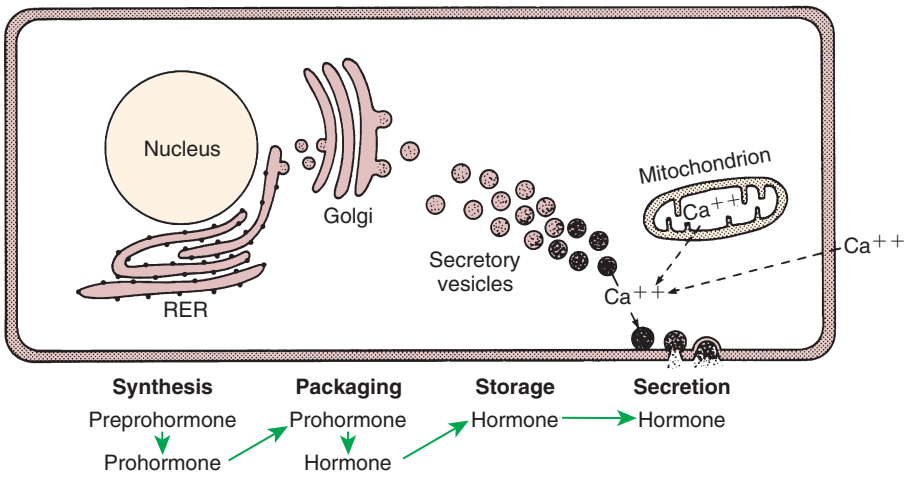
The interaction between the nervous and endocrine systems can be even more direct. For example, endocrine cells of the adrenal medulla are directly controlled by preganglionic neurons of the adrenal medulla, and the medullary hormones are released immediately in response to stressful stimuli. The endocrine and nervous systems also share transmitters; substances such as epinephrine, dopamine, histamine, and somatostatin are found in both endocrine and neural tissues.

The endocrine system is involved in control of physiological functions, including metabolism, growth, and reproduction. *Metabolism* can be divided into two forms: energy and mineral. The hormones that control *energy metabolism* include insulin, glucagon, cortisol, epinephrine, thyroid hormone, and growth hormone. The hormones that control *mineral metabolism* include parathyroid hormone, calcitonin, aldosterone, angiotensin, and renin. The hormones that control *growth* include growth hormone, thyroid hormone, insulin, estrogen and androgen (both reproductive hormones), and a large number of growth factors. The hormones that control *reproduction* include estrogen, androgen, progesterone, luteinizing hormone (LH), follicle-stimulating hormone (FSH), prolactin (PRL), and oxytocin.

One of the important characteristics of the endocrine system is the *amplification* of the signal. The action of one steroid molecule to activate a gene can result in the formation of many messenger ribonucleic acid (mRNA) molecules, and each of these can induce the formation of many enzyme molecules. Also, one protein molecule can influence the formation of many cyclic adenosine



• **Fig. 33.2** Comparison of functional arrangements of an ordinary neuron releasing its neurotransmitter (*NT*) into a synapse and a neurosecretory neuron releasing its neurohormone (*NH*) into a blood vessel. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)



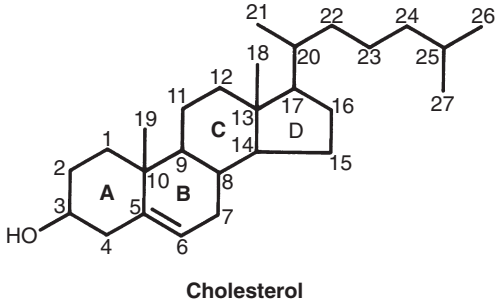
• **Fig. 33.3** Subcellular components of peptide hormone synthesis and secretion. *RER,* Rough endoplasmic reticulum. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

3′,5′-monophosphate (cAMP) molecules, and each of these can activate many enzymes. Amplification is the basis for the sensitivity of the endocrine system, which allows small amounts of hormones in plasma (10^{-11}–10^{-12} mol) to produce significant biological effects. Hormone action also influences rates of existing enzyme reactions, implying that there are certain basal levels of enzyme activities even in the absence of hormones. Hormone action is relatively slow and prolonged, with the effects of hormones lasting minutes to days. This contrasts with the nervous system, in which the response is rapid and short (milliseconds to seconds).

Synthesis of Hormones

Protein Hormones Are Initially Synthesized as Preprohormones and Then Cleaved in the Rough Endoplasmic Reticulum to Form Prohormones, and in the Golgi Apparatus to Form the Active Hormones, Which Are Stored in Granules Before Being Released by Exocytosis

The major classes of hormones include proteins (e.g., growth hormone, insulin, corticotropin [also called adrenocorticotropic hormone, or ACTH]); peptides (e.g., oxytocin and vasopressin); amines (e.g., dopamine, melatonin, epinephrine); and steroids (e.g., cortisol, progesterone, vitamin D). The protein and peptide hormones are initially synthesized on ribosomes as larger precursor proteins, which are referred to as *preprohormones* (Fig. 33.3). Synthesis of protein hormones begins in ribosomes, with the "pre" portion immediately attaching to the rough endoplasmic reticulum (RER), which pulls the ribosomes into close apposition with the RER. During synthesis, the preprohormone is secreted into the interior of the RER. The presence of a peptidase within the wall of the RER allows the "pre" portion of the molecule to be rapidly removed and the prohormone to leave the RER in vesicles that have been pinched off from the RER. These vesicles then move to the Golgi apparatus, where they coalesce with Golgi membranes to form secretory granules. The prohormone is cleaved during this process, so most of the hormone is in its final form within the Golgi apparatus, although some prohormone can also be found.



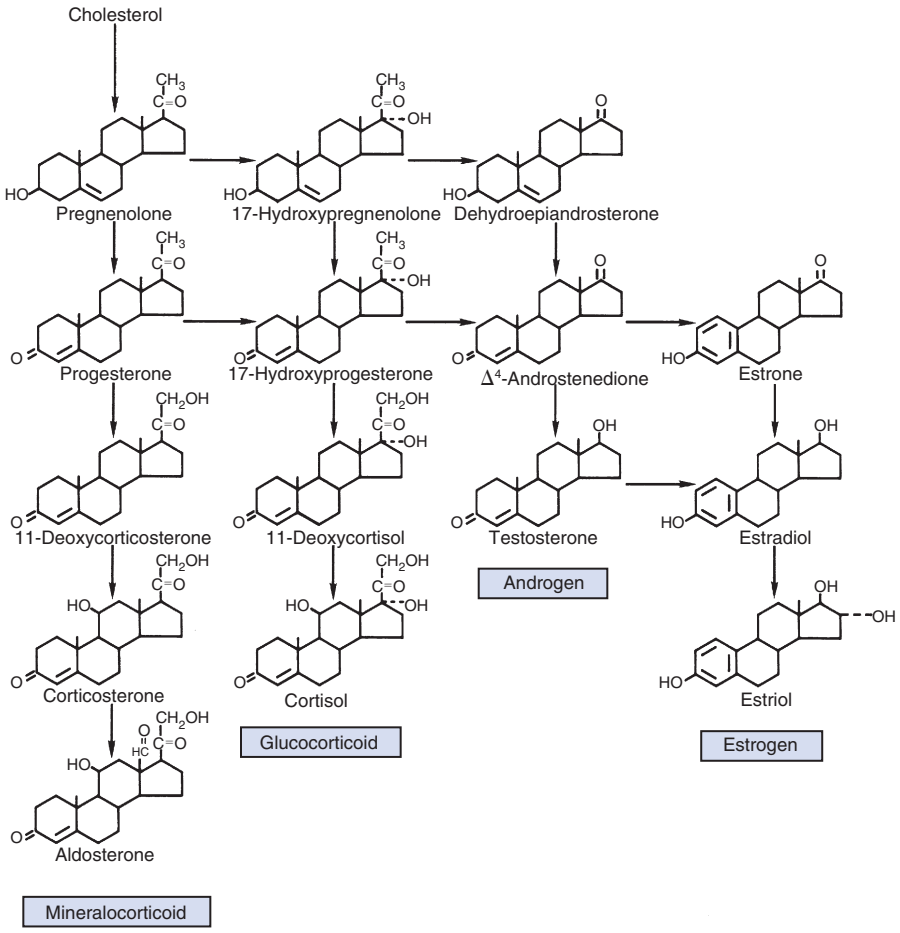
Cholesterol

• **Fig. 33.4** The ring structure and numbering system of the carbon atoms in steroid hormones, illustrated for the cholesterol molecule. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

Protein hormones are stored in granules within the gland until needed for release. Although some of the hormone is secreted on a continuous basis, most is secreted through the process of *exocytosis* of granules in response to a specific signal. The process of exocytosis requires adenosine triphosphate (ATP) and calcium (Ca^{2+}). Increased cytoplasmic calcium results from intracellular release of Ca^{2+} from mitochondria, or endoplasmic reticulum, or from the influx of extracellular Ca^{2+}.

Steroids Are Synthesized from Cholesterol, Which Is Synthesized by the Liver; Steroids Are Not Stored but Are Released as They Are Synthesized

Steroids represent a class of hormones that, unlike protein hormones, are lipophilic. In general, they belong to one of two categories: *adrenocortical hormones* (glucocorticoids, mineralocorticoids) and *sex hormones* (estrogens, progesterone, androgens). They have a common four-ring, 17-carbon skeleton derived from *cholesterol* (Fig. 33.4). Although the steroids can be synthesized *de novo* within the cell from the two-carbon molecule acetate, the majority of steroids are formed from cholesterol, which is synthesized by the liver (Fig. 33.5). Low-density lipoproteins (LDLs) enter



• **Fig. 33.5** Pathways involved in the production of the major steroid hormones. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

steroid-producing cells through interaction with a membrane receptor. Cholesterol is released through the degradation of LDLs by lysosomal enzymes. Cholesterol is either used immediately for steroid synthesis or stored in granules in an ester form within the cell. The first step in the synthesis of all steroid hormones from cholesterol involves cleavage of the side chain of cholesterol to form *pregnenolone;* this step occurs within the mitochondrion. Subsequent modifications of the steroid molecule may occur within the mitochondrion or may involve movement to other compartments of the cell (Fig. 33.6). The control of movement of steroids among cell compartments during the synthesis process is not well understood.

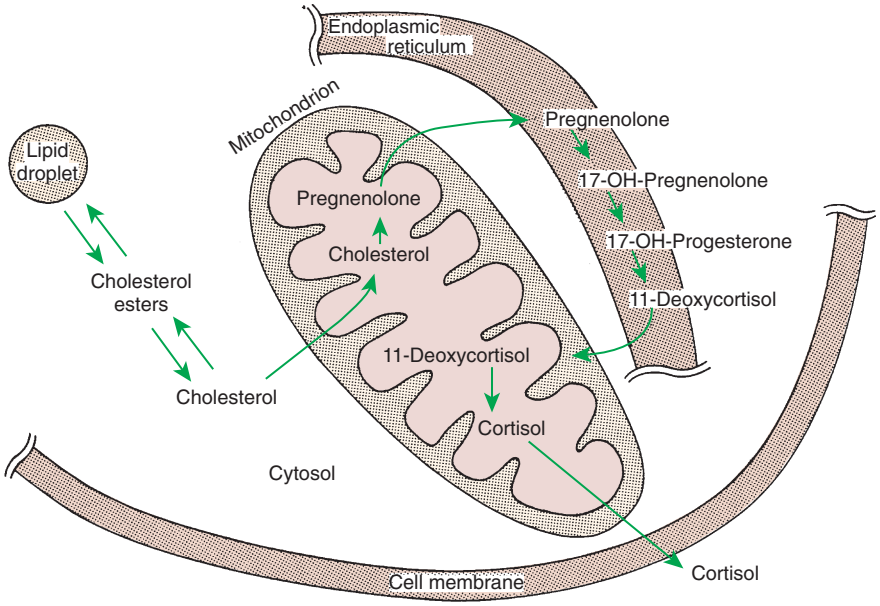
The type of steroid hormone that is eventually synthesized depends on the presence of specific enzymes within the particular cell. For example, only cells of the adrenal cortex contain enzymes (hydroxylases) that result in hydroxylation of the eleventh and twenty-first carbon molecules, a process essential for the production of glucocorticoids and mineralocorticoids. The pattern for sex steroid biosynthesis is for pregnenolone to be modified in a sequence that involves progesterone, androgens, and finally estrogens. Cells that synthesize androgens (e.g., Leydig cells of the testis) have the enzymes required for the formation of pregnenolone and progesterone, as well as for the modification of progesterone to androgen, but lack the enzymes necessary to modify androgens into estrogens. Although the sex steroid–forming cells do not have enzymes that allow the formation of adrenocortical hormones, the adrenal cortex contains the enzyme systems necessary for the formation of both adrenocortical hormones and sex hormones, although the former are emphasized. As a result, the adrenal cortex normally produces small amounts of sex steroids and larger amounts in certain pathophysiological conditions.

There is no provision for the storage of steroid hormones within the cell; they are secreted immediately after formation by simple diffusion across the cell membrane because of their lipophilic structure. Thus synthesis and secretion of steroid hormones occur in a tightly coupled manner, whereby the rate of hormone secretion is controlled by the rate of synthesis. The only storage form of steroids within these cells involves that of the precursor molecule, cholesterol, as an ester.

Transport of Hormones in the Blood

Protein Hormones Are Hydrophilic and Carried in the Plasma in Dissolved Form

This chapter focuses mainly on hormones transported to target tissues in the vascular system. The means by which hormones are transported in the blood varies according to the solubility of the hormone. Protein and peptide hormones are hydrophilic and are carried in the plasma in dissolved form. The protein hormones may circulate in *monomeric* (single-unit) or *polymeric* (multiple-unit) form (e.g., insulin). Hormones that have subunits can appear in



• **Fig. 33.6** Subcellular compartmentalization of cortisol biosynthesis. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

the circulation in subunit form, although this reduces the biological potency of the molecule.

Steroids and Thyroid Hormones Are Lipophilic and Carried in Plasma in Association With Both Specific and Nonspecific Binding Proteins; the Amount of Unbound, Active Hormone Is Relatively Small

The transport of steroid and thyroid hormones is more complicated than that of protein hormones, because the steroid and thyroid hormones are lipophilic and thus have limited solubility in aqueous solutions including blood. These hormones are transported in the blood through association with various types of proteins. Some of the proteins that bind steroids have a high affinity for a particular steroid; for example, a globulin, *transcortin,* has a high affinity for cortisol and corticosterone but also serves as an important transport vehicle for progesterone, even though it has a lower affinity for this hormone. The carrier proteins that have high affinities have low capacity because of their low plasma concentration. In contrast, the plasma proteins called *albumins* have low affinities for steroid hormones but have a high capacity for steroid transport because of their high concentration in plasma.

A hormone must be in the *free,* or unbound, form before it can penetrate a target cell and elicit biological activity. This is accomplished by the establishment of equilibrium between bound and free hormone levels in the plasma. The free form usually represents only about 1% of the total amount of hormone in the plasma (up to 10% of cortisol may be in the free form). The system is responsive to use of the free form, and the free form is replenished quickly by dissociation of bound hormone from the protein. The *total* amount of the hormone is usually measured, with the exception of thyroid hormone, for which attempts are usually made to estimate the amounts of bound and free. As indicated for steroid hormones, synthesis and release are tightly linked, and because metabolic clearance rates are usually constant, concentrations of steroids in plasma are usually a good reflection of the secretion rate. Under certain physiological conditions, such as pregnancy in humans, metabolism of estrogens can change because of the increased production of estrogen-binding proteins.

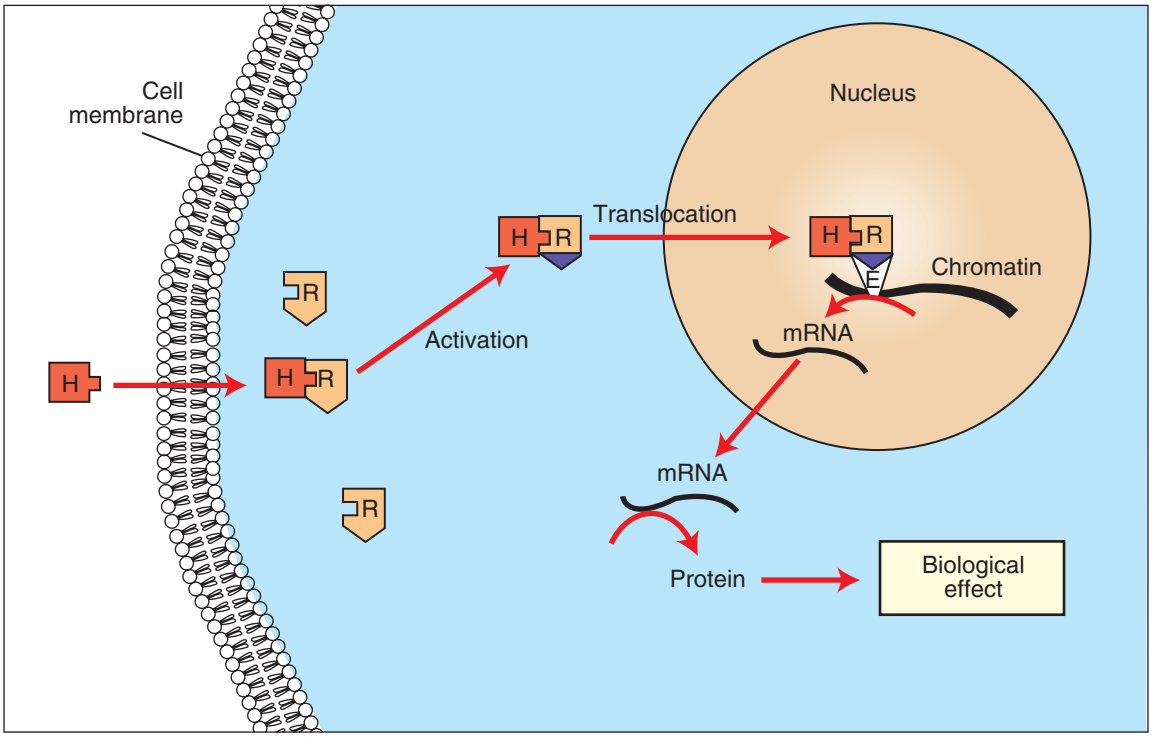
Hormone-Cell Interaction

Protein Hormones Have Specific Receptors on Target Tissue Plasma Membranes, Whereas Steroids Have Specific Receptors Within the Cytoplasm or Nucleus

A central question in endocrinology is how hormones and target cells of a particular tissue interact in a specific manner. The problem seems almost overwhelming for steroids because they are lipid soluble and able to permeate all cells of the body. The solution is that target cells have receptors that are *specific* for a particular hormone. For steroids, the receptors are located in the cytoplasm or nucleus of the target cells, whereas receptors for protein and peptide hormones are located on the plasma membrane of the cell. In addition to specificity, receptors have a high *affinity* for their respective hormone. These characteristics of the receptor allow hormones to be in low concentration in the blood but effective in producing significant tissue response.

The greater the affinity of the receptor for the hormone, the longer is the biological response. Termination of the action of a hormone usually requires *dissociation* of the hormone from the receptor. This occurs most often as a result of a decrease in plasma concentrations of the hormone; the binding of receptor and hormone is noncovalent, and declining hormone concentrations favor a chemical equilibrium of dissociation over association. Termination of hormone action can also result from *internalization* of the receptor-hormone complex through the process of *endocytosis.* The hormone is degraded by lysosomal enzymes, whereas the receptor, protected because of its association with the vesicle membrane, can be recycled to the plasma membrane.

Receptors are present on cells in much greater numbers than required for the elicitation of a biological response. Occupancy



• **Fig. 33.7** Subcellular mechanism of action of a lipophilic hormone (*H*) via an intracellular receptor (*R*). The H-R complex induces messenger ribonucleic acid (*mRNA*) synthesis by binding to a response element (*E*) on the chromatin. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

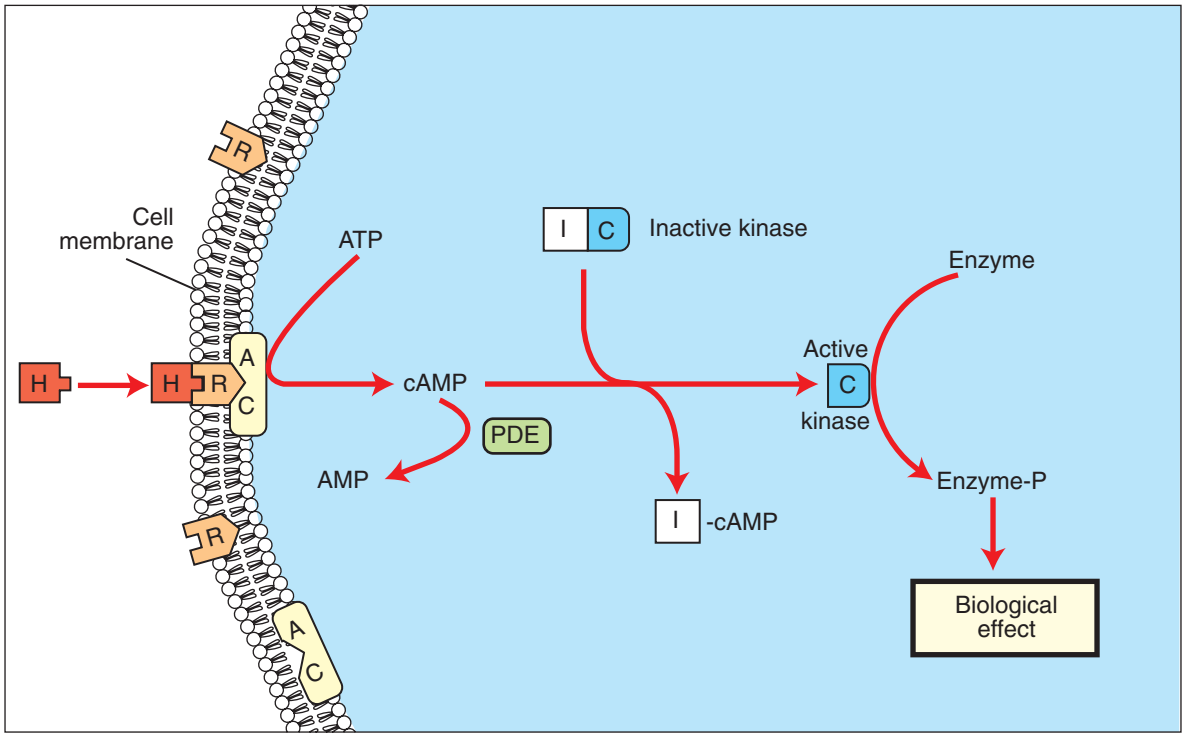
by a hormone of considerably less than 50% of the receptors usually elicits a maximal biological response. Even so, changes in receptor numbers that affect the sensitivity of the cell, although not its maximal responsiveness, can occur. Changes in receptor number affect the probability that an interaction will occur between receptor and hormone. Receptor synthesis can be stimulated by a hormone that is different from the hormone that interacts with the receptor. For example, predominant gonadotropin receptors on ovarian granulosa cells change from FSH to LH receptors late in the ovarian follicle phase because of the influence of FSH. This allows the control of the ovarian follicle to pass from FSH to LH, which facilitates ovulation and luteinization (the formation of a corpus luteum). Conversely, receptor numbers can decrease in conjunction with continued interaction of receptor and hormone. This often occurs when an agonist that has great affinity for the receptor is administered or when amounts of hormone are pathologically elevated. The receptor numbers are downregulated in this situation. The end result is that the animal becomes resistant to continued therapy with the hormone in question.

Postreceptor Cell Responses

Steroids Interact Directly With the Cell Nucleus Through the Formation of a Complex With Their Cytosolic or Nuclear Receptors, Whereas Protein Hormones Need a Messenger Because They Cannot Enter the Cell

The events that follow binding of the hormone and receptor depend on whether a steroid, protein, or peptide hormone is involved. With steroids, the hormone is able to interact within the cell because of its ability to penetrate the lipoprotein plasma membrane (Fig. 33.7). Some steroids (e.g., estrogen) also penetrate the nucleus to complex with their nuclear steroid receptor. The interaction of receptor and steroid hormone in the cytosol (e.g., androgen, glucocorticoid) results in activation of the complex and subsequent translocation to the nucleus, where it interacts with specific sites on the chromatin (response elements). The result is the production of mRNA, which, when translocated to the ribosomes, directs synthesis of proteins that produce the desired biological result.

Protein or peptide hormones require an intermediary to act in their behalf because they are not able to penetrate the plasma membrane of the cell; the intermediary substance is known as a *second messenger* (Fig. 33.8). The best-documented second messenger is cAMP, which is produced by the activation of an enzyme, adenylyl cyclase, through interaction of the hormone and receptor in the plasma membrane. The activation of adenylyl cyclase and the production of cAMP result in the phosphorylation of protein kinases, which are responsible for the biological response. Other second messengers include cytosolic calcium and its associated phosphodiesterase, calmodulin, as well as inositol triphosphate (IP_3) and diacylglycerol, both of which are products of phosphatidylinositol metabolism. An important action of IP_3 is the stimulation of intracellular calcium release. One important response to diacylglycerol is the activation of phospholipase A and the formation of arachidonic acid, which leads to formation of members of the prostaglandin family of molecules. The biological response to a protein or peptide hormone–receptor interaction is often more rapid than that to steroids; preexisting enzymes are activated, whereas the biological response to steroid generally requires the synthesis of enzyme protein.



• **Fig. 33.8** Subcellular mechanism of action of a hydrophilic hormone (*H*) via a membrane receptor (*R*), adenyl cyclase (*AC*), and cyclic adenosine monophosphate (*cAMP*). *ATP,* Adenosine triphosphate; *I* and *C,* inhibitory and catalytic subunits of the kinase, respectively; *PDE,* phosphodiesterase. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

Metabolism of Hormones

Steroid Hormones Are Metabolized by Conjugation With Sulfates and Glucuronides, Which Makes Steroids Water Soluble

Hormone activity is limited by the metabolism of hormones. The metabolism of steroids usually involves reduction of the molecule, followed by conjugation with sulfates and glucuronides, which increases the water solubility of the steroids, allowing them to be excreted in urine. The liver is the main organ responsible for this process. Iodine molecules are removed from thyroid hormones during metabolism. Protein hormones are cleaved by peptidases; this is preceded by reduction of disulfide bonds if this is a characteristic of the molecule. Although a metabolite is usually less biologically potent than the original molecule, some evidence suggests that conjugates of steroids can have significant biological activity. This raises the question of whether the conversion of hormones intracellularly, such as testosterone to dihydrotestosterone, represents metabolism, because dihydrotestosterone is more potent biologically than testosterone. Another example, the conversion of estradiol-17β to estrone by peripheral tissues, including adipose cells, is described as a form of metabolism; however, estrone is a natural, and relatively potent, estrogen.

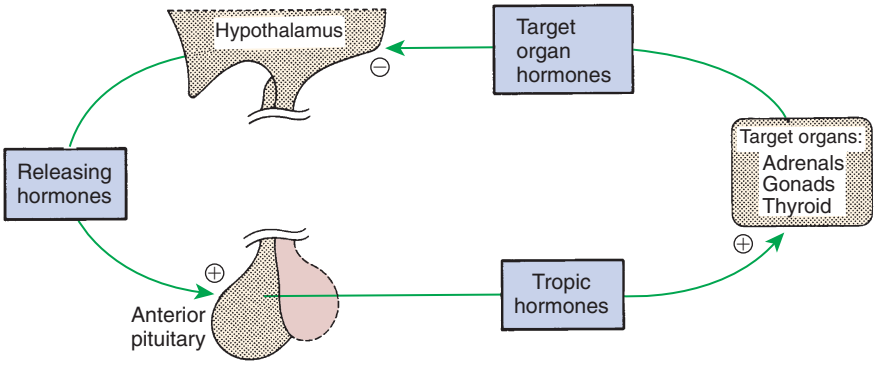
Although in some situations the *rate of clearance* of a hormone can change (e.g., decrease as a result of increased hormone-binding plasma proteins during pregnancy, or increase as the result of decreased hormone-binding plasma proteins in conjunction with liver disease), *metabolism* of hormones is relatively constant, and the concentration of a hormone usually reflects the other determinant of hormone activity: the *rate of synthesis* of the hormone.

Feedback Control Mechanisms

The Most Important Feedback Control for Hormones Is the Negative-Feedback System, in Which Increased Hormone Concentrations Result in Less Production of the Hormone, Usually Through an Interaction With the Hypothalamus or Pituitary Gland

The effects of hormones are proportional to their concentrations in blood, and therefore control of these concentrations is an important aspect in ensuring that physiological function is normal. As indicated previously, the primary factor affecting hormone concentrations in blood is the secretion rate by a particular organ. Feedback-loop control systems have evolved in which concentrations of hormones are monitored at the controlling point either to increase or to decrease secretion of a hormone by an endocrine organ. The most common feedback system by far is *negative feedback,* in which continuous monitoring allows the system to counteract changes in hormone secretion or to maintain a relatively constant environment.

An example of systems in which negative-feedback control involves both the endocrine and the nervous system is shown in Fig. 33.9. The hypothalamus, which controls secretion of tropic hormones in the anterior pituitary through the secretion of peptide-releasing hormones, has cells with a certain set point by which they compare concentrations of hormone in the blood with the output of releasing hormones. If blood concentrations fall below the physiological set point, releasing-hormone output increases; this in turn increases production of tropic hormones by the anterior



• **Fig. 33.9** Negative feedback of tropic and releasing hormones by target organ hormones. The *plus signs* indicate stimulation, and the *minus sign* indicates inhibition. In some cases, such inhibition occurs at the pituitary gland. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

pituitary and subsequently the secretion of the hormone by the target organ. Conversely, if the hormone concentration increases above acceptable physiological limits, a shutdown of releasing-hormone production occurs within the hypothalamus, tropic hormone secretion by the anterior pituitary decreases, and production of the hormone by the target organ decreases. This type of control system is not an "all-or-none" arrangement because changes and adjustments are being continuously made to maintain an optimal concentration of hormone.

In the negative-feedback system, an increase in secretion of hormone results in a decrease in tropic hormone secretion. It is also possible to have a negative-feedback system in which an increase in a physiological substance, such as glucose, causes an increase in a hormone, in this case insulin, which plays an important role in glucose metabolism. This is considered to be a negative-feedback system because blood glucose concentrations are being "dampened," or returned toward normal levels, through the action of insulin.

Positive-feedback systems also exist, although they are much less common than negative-feedback systems. One example is the preovulatory release of LH, in which the pulsatile rate of LH secretion greatly increases during the late stages of ovarian follicular development because of increased estrogen production by the follicle. In this situation, there is a definitive end point; ovulation results in a decline in the stimulus, estrogen, although the duration of the LH surge is probably determined within the hypothalamus, and therefore the LH response to estrogen is modulated.

Endocrine Secretory Patterns Can Be Influenced by Factors Such as Sleep or Light and Can Produce Circadian Rhythms

Endocrine secretion patterns can occur outside the control of negative-feedback inhibition. Hormone patterns can change on an approximately 24-hour basis, a process referred to as a *diurnal rhythm,* or *circadian rhythm. Circadian* is the preferred term because *diurnal* refers to activity in the daytime; *nocturnal* should be used for the rhythms that are active at night. Most of the daily rhythms have some aspects of light, or lack of light, as a major influence in the rhythm. Rhythmic changes in hormone patterns that occur at shorter intervals, often in the range of an hour, are called *ultradian rhythms.*

The Hypothalamus

The Hypothalamus Coordinates the Activity of the Pituitary Gland Through the Secretion of Amines and Peptides, Particularly Releasing Hormones

As indicated previously, the two major controlling systems are the nervous and the endocrine systems. The interface between these systems occurs, for the most part, in the hypothalamus. The hypothalamus is an area of the diencephalon that forms the floor of the third ventricle and includes the optic chiasma, tuber cinereum, mammillary bodies, and the median eminence. Often not included in this classification are the infundibulum and the neurohypophysis (stalk of the posterior lobe and the posterior lobe, respectively), although both tissues represent extensions of the hypothalamus into the pituitary gland.

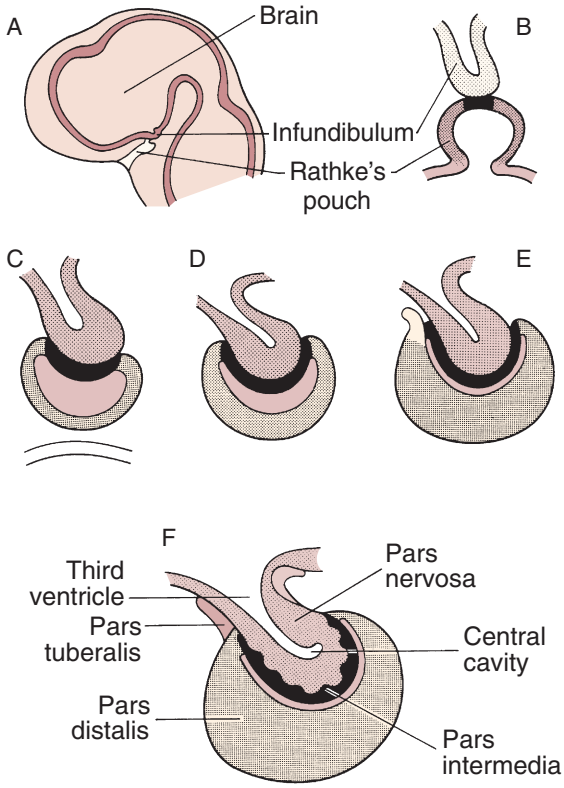
The hypothalamus produces peptides and amines that influence the pituitary gland to produce (1) tropic hormones (e.g., ACTH), which in turn influence the production of hormones (e.g., cortisol) by peripheral target endocrine tissue, or (2) hormones that directly cause a biological effect in tissues (e.g., PRL). The hypothalamus is also the center for the control of a large number of autonomic nervous system control pathways.

The Pituitary Gland (Hypophysis)

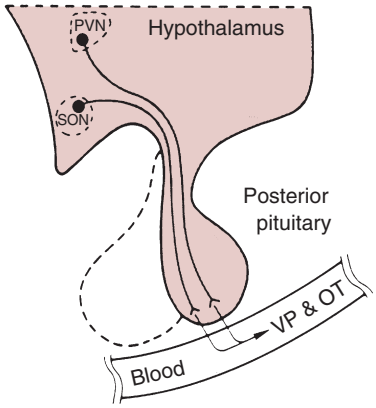
The pituitary gland, or hypophysis, is composed of the *adenohypophysis* (pars distalis, or anterior lobe), the *neurohypophysis* (pars nervosa, or posterior lobe), the *pars intermedia* (intermediate lobe), and the *pars tuberalis* (Fig. 33.10). The adenohypophysis is formed from an area of the roof of the embryonic oral ectoderm called *Rathke's pouch,* which extends upward to meet the neurohypophysis, which extends downward as an outpouching of neural ectoderm from the floor of the third ventricle.

The Neurohypophysis Has Cell Bodies That Originate in the Hypothalamus, With Cell Endings That Secrete Oxytocin and Vasopressin (Antidiuretic Hormone)

The neurohypophysis is composed of axons whose neural origin is largely within the supraoptic and paraventricular nuclei of the



• **Fig. 33.10** Diagrams showing progressive stages in the embryonic development of the pituitary gland. Rathke's pouch becomes detached from the oral epithelium at stage C. (From Villee CA, Walker WF Jr, Smith FE. *General Zoology*. 2nd ed. Philadelphia: Saunders; 1963; Turner CD, Bagnara JT. *General Endocrinology*. 6th ed. Philadelphia: Saunders; 1976.)



• **Fig. 33.11** The hypothalamoneurohypophyseal system, which secretes vasopressin (*VP*) and oxytocin (*OT*). *PVN,* Paraventricular nucleus; *SON,* supraoptic nucleus. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology,* Philadelphia: Saunders; 1987.)

hypothalamus. The neurohypophysis is an extension of the hypothalamus into the pituitary; that is, the cell bodies are in the hypothalamus. The axons form the stalk of the posterior lobe, and the nerve endings are in the lobe proper (Fig. 33.11).

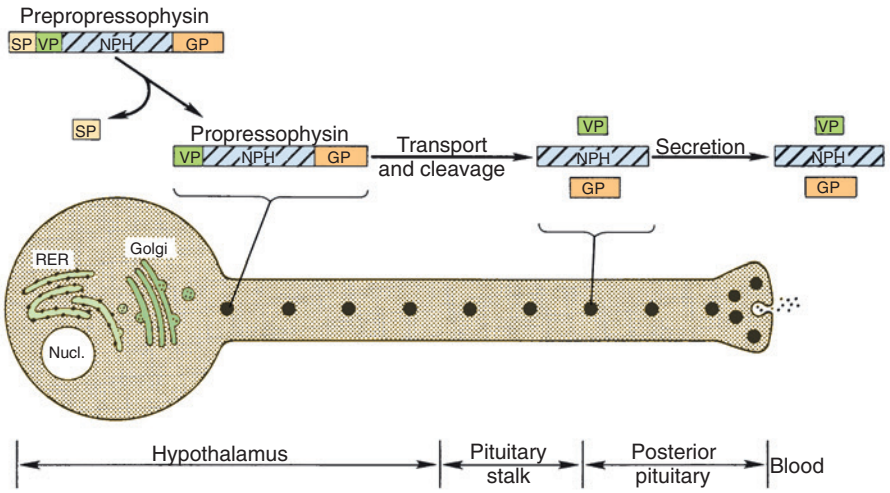
The endocrine-secretory neurons that constitute the neurohypophysis differ from neurons involved in the transmission of neural signals in several ways: (1) neurosecretory neurons do not innervate other neurons, even though they are innervated; (2) the secretory product of neurosecretory neurons is secreted into the blood; and (3) the secretory product can act at distances greatly removed from the neuron. Also, in contrast to anterior pituitary hormones, which influence other tissues to produce hormones, posterior lobe hormones can directly cause the desired tissue response.

The first indication of the physiological activity of the neurohypophyseal lobe was the finding of Oliver and Schafer in 1895 that the injection of whole pituitary extracts caused a rise in blood pressure. This effect was soon found to be associated with the *pars nervosa.* This action represents the effects of one of the main neurohypophyseal hormones, *vasopressin or antidiuretic hormone.* The existence of the other main neurohypophyseal hormone, *oxytocin,* was first indicated in 1915, when Gaines showed that the injection of posterior pituitary gland extracts caused milk ejection. In 1941, Ely and Peterson showed that a denervated mammary gland could eject milk if the gland was perfused with blood that had been enriched with posterior pituitary extract. Both neurohypophyseal hormones were isolated and sequenced by du Vigneaud in 1954. These were some of the first proteins whose amino acid sequences were documented.

Oxytocin and Vasopressin Are Synthesized in Cell Bodies Within the Hypothalamus and Are Carried by Axon Flow to the Posterior Lobe, Where They Are Released

As indicated, the two important hormones produced by the neurohypophysis are vasopressin and oxytocin. Although it was previously thought that the two hormones were produced in separate nuclei, evidence now indicates that both hormones are produced in both the supraoptic and the paraventricular nucleus. The cell bodies that synthesize the hormones are large and thus the neurons are called *magnocellular neurons.* The synthesis of vasopressin and oxytocin, as described previously for protein and peptide hormones, involves first the production of a preprohormone, prepropressophysin for vasopressin and preprooxyphysin for oxytocin, at the level of the cell body within the hypothalamus (Fig. 33.12). The "pre" portion of the molecule is cleaved before the molecules are packaged into granules. During passage of the granules down the axon, the prohormone is cleaved to produce either oxytocin or vasopressin; the remaining peptide fragments are called neurophysin I or neurophysin II, respectively. Neurophysin I, which is released into the vascular system along with oxytocin, has been quantified as an alternative means of following the release of oxytocin. At present, the physiological function of the neurophysins is not known.

The release of the posterior lobe peptide hormones is initiated in the hypothalamus as a result of depolarization of the cell body because of stimulation by neural afferents. The action potential generated extends down the axon to the nerve terminal, where the secretory granules containing the hormone are stored. The depolarization of the nerve cell membrane allows the influx of calcium ions, which initiates the release of hormone through the process of exocytosis.



• **Fig. 33.12** Diagram of a vasopressin-secreting neuron illustrating the subcellular components involved in synthesis and secretion. This process begins with the synthesis of prepropressophysin, which consists of (1) a signal peptide (*SP*), (2) vasopressin (*VP*), (3) neurophysin (*NPH*), and (4) a glycoprotein (*GP*). The production and release of oxytocin is identical except that no glycoprotein is involved. *Nucl.,* Nucleus; *RER,* rough endoplasmic reticulum. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

The Main Effects of Oxytocin Are on the Contraction of Smooth Muscle (Mammary Gland and Uterus); the Effects of Vasopressin Are Primarily on the Conservation of Water (Antidiuresis) and Secondarily on Blood Pressure

The main effects of oxytocin involve the contraction of the myoepithelial cells, which surround the alveoli in the mammary gland and the myometrium of the uterus (see Chapters 38 and 39).

The main activity of vasopressin belies its name because its main effect is antidiuretic, the enhancement of water retention by the kidney. As a consequence, the hormone is often called *antidiuretic hormone* (ADH) (Fig. 33.13). Vasopressin is the most important hormone for the control of water balance. Vasopressin also has a pressor effect, which involves the contraction of smooth muscle of the vascular system and therefore has an effect on blood pressure. The main form of vasopressin in most species is arginine vasopressin, whereas in pigs it is lysine vasopressin, and in birds it is arginine vasotocin.

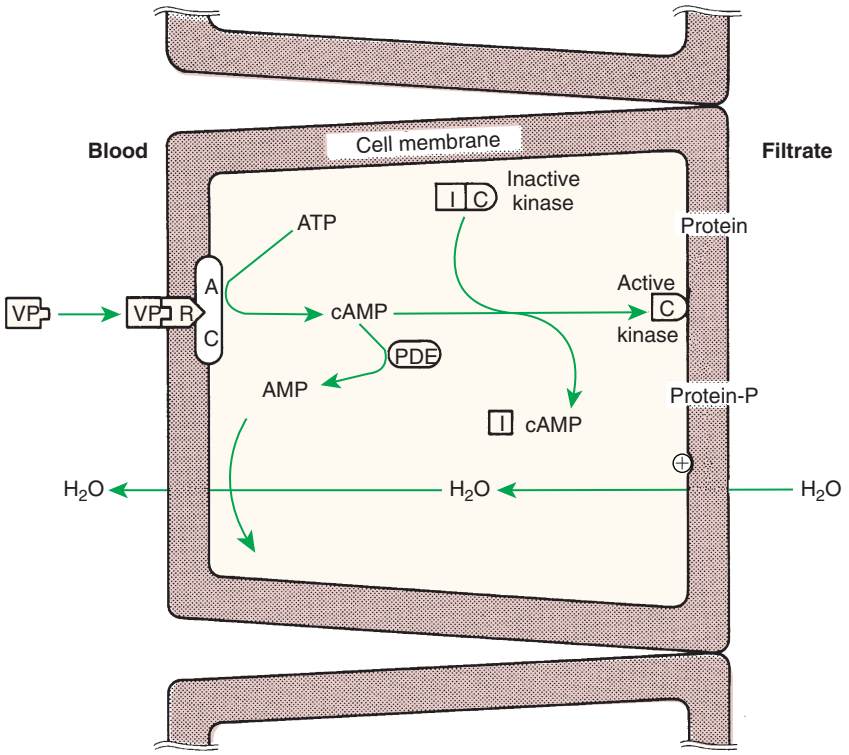
Plasma Osmolality Controls the Secretion of Vasopressin

The control of vasopressin secretion as a result of changes in plasma osmolality is through *osmoreceptors* located in the hypothalamus as well as through receptors located in the esophagus and stomach that immediately sense water intake (Fig. 33.14). An increase in osmolality of body fluids increases the rate of action potential firing in the osmoreceptors, which in turn activates hypothalamic cells that synthesize vasopressin. This negative-feedback system is sensitive to changes in osmolality, and the solute-to-water ratio is maintained within 1% to 2% of the normal values. The regulation of the pressor effect of vasopressin—that is, through blood volume—is achieved by increasing the number of action potentials in stretch receptors located in the atria. A decrease in blood volume activates the stretch receptors, which inhibit activity of neurons, vagal in origin, that inhibit the osmoreceptor cells. Blood volume changes that decrease blood pressure also affect vasopressin release through activation of *baroreceptors* in the carotid sinus and aortic arch.

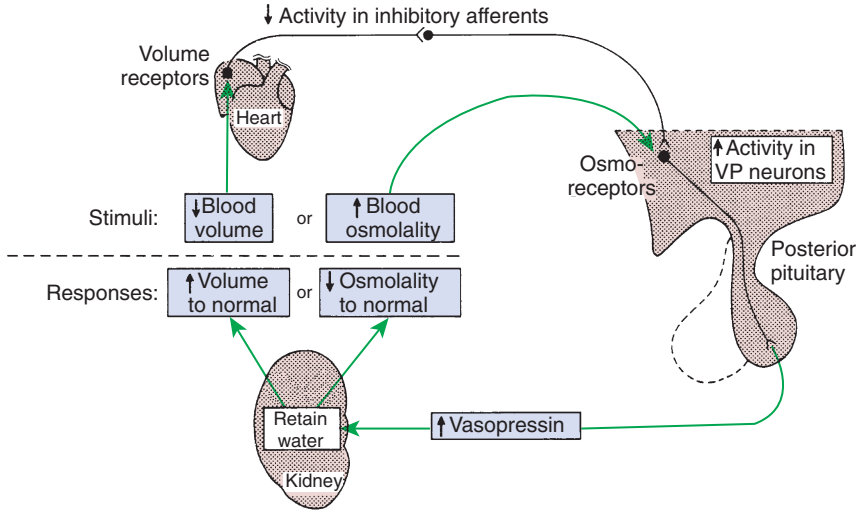
Diabetes insipidus (DI) is a disorder of water metabolism characterized by polyuria (passing abnormally large volumes of urine), urine of low specific gravity or osmolality, and polydipsia (excessive thirst and fluid intake). It is caused by defective secretion of ADH (central DI) or by the inability of the renal tubule to respond to ADH (nephrogenic DI). Deficiency of ADH can be partial or complete. *Central* DI is characterized by an absolute or relative lack of circulating ADH and is classified as *primary* (idiopathic or congenital) or *secondary.* Secondary central DI usually results from head trauma or neoplasia.

Central DI may appear at any age, in any breed of dog or cat, and in either gender; however, young adults (6 months of age) are most often affected. The major clinical signs of DI are profound polyuria and polydipsia (>100 mL/kg/day; normal range, 40–70 mL/kg/day), nocturia, and incontinence, usually of several months' duration. The severity of the clinical signs varies because DI may result from a partial or complete defect in ADH secretion or action. Other less consistent signs include weight loss (because these animals are constantly seeking water) and dehydration.

Routine complete blood cell count and serum biochemical and electrolyte profiles are usually normal in animals with DI. Plasma osmolality is often high (>310 mOsm/L) in central or nephrogenic DI as a result of *dehydration.* Animals with primary polydipsia (not of central DI or nephrogenic DI origin, and another cause of polyuria) often exhibit low plasma osmolality (<290 mOsm/L) as a result of *overhydration.* When present on initial evaluation, abnormalities (e.g., slightly increased hematocrit, hypernatremia) are usually secondary to dehydration from water restriction by the pet owner. In DI, the urinalysis is unremarkable except for the finding of a persistently dilute urine (urine specific gravity, 1.004–1.012).



• **Fig. 33.13** Antidiuretic mechanism of action of vasopressin (*VP*) on cells of the distal tubule and collecting ducts. *AC,* Adenyl cyclase; *AMP,* adenosine monophosphate; *ATP,* adenosine triphosphate; *cAMP,* cyclic AMP; *I* and *C,* inhibitory and catalytic subunits of the kinase, respectively; *PDE,* phosphodiesterase; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)



• **Fig. 33.14** Major mechanisms regulating vasopressin (*VP*) secretion. A perturbation in either blood volume or osmolality modifies vasopressin secretion to restore these parameters to their normal values. However, this restoration requires appropriate water intake adjustments by thirst as well as by the modulation of water retention depicted. Also, the two responses indicated may be affected by simultaneous changes in sodium balance. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

Diagnostic tests to confirm and differentiate central DI, nephrogenic DI, and primary polydipsia include the *modified water deprivation test* or response to ADH supplementation. The modified water deprivation test is designed to determine whether endogenous ADH is released in response to dehydration and whether the kidneys can respond to ADH. The more common causes of polyuria and polydipsia should be ruled out before this procedure is performed. Failure to recognize renal failure before water deprivation may lead to an incorrect or inconclusive diagnosis or may cause significant morbidity in the patient.

Hypersecretion of vasopressin in the absence of osmotic or volumetric stimulation is rare and called the *syndrome of inappropriate antidiuretic hormone secretion* (SIADH). Neoplastic processes are often involved in SIADH; ectopic tumors, often located in the lung, are the neoplasms most frequently involved.

The Anterior Pituitary Produces Growth Hormone, Prolactin, Thyroid-Stimulating Hormone, Follicle-Stimulating Hormone, Luteinizing Hormone, and Corticotropin

The adenohypophysis includes the pars distalis and the pars intermedia. The major hormones produced by the anterior pituitary are growth hormone (GH, also called *somatotropin*), PRL, thyroid-stimulating hormone (TSH), FSH, LH, and corticotropin (ACTH) (Table 33.1). GH is produced by acidophilic somatotropes, and PRL is produced by lactotropes; both are classified as *somatomammotropins.* GH and PRL are single-chain proteins that contain two and three disulfide bonds, respectively. There is overlap of activity between GH and PRL; this overlap is based on the approximately 50% homology of their amino acid sequences. Of these two major somatomammotropins, GH is uniquely species specific as to its activity.

TSH, produced by thyrotropes, and FSH and LH, produced by gonadotropes, are classified as *glycoproteins* because all three molecules have carbohydrate moieties. These hormones have α and β subunits that are linked by noncovalent binding. The α subunits are identical (and interchangeable) among the three glycoproteins. The β subunits, unique for each hormone, impart the specific action of each hormone. Other members of this family of hormones that are not of anterior pituitary origin include *equine chorionic gonadotropin* (also called *pregnant mare's serum gonadotropin*) and *primate chorionic gonadotropin,* which are produced by cells of the placental.

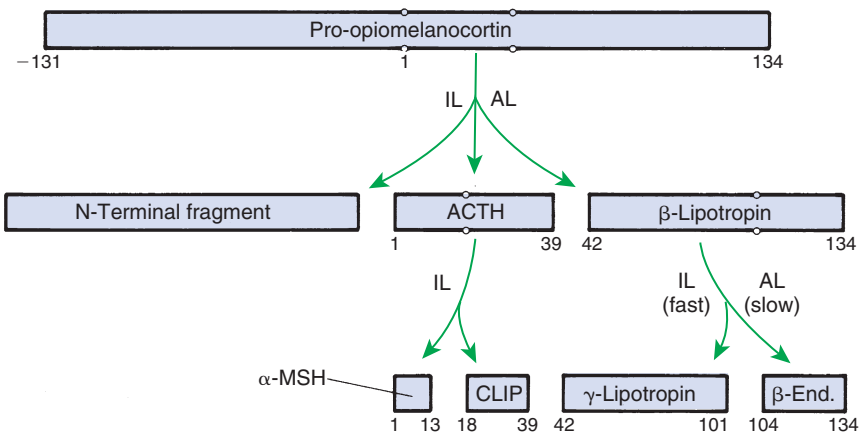
ACTH and β-lipotropin belong to the *proopiomelanocortin* family in that they originate from a common prohormone (Fig. 33.15). Cells in both the *pars distalis* and the *pars intermedia* synthesize proopiomelanocortin molecules. The emphasis on the type of hormone produced is different in the end product; ACTH is produced by *pars distalis* corticotropes. In the *pars intermedia*, ACTH is cleaved by corticotropes to form *α-melanocyte–stimulating hormone* (α-MSH), the predominant hormone of this lobe. The remaining peptide fragment is known as *corticotropin-like intermediate lobe peptide;* the physiological activity of this peptide fragment is not known. In both the *pars distalis* and the *pars intermedia,* β-lipotropin is cleaved to form β-endorphins and γ-lipotropin. Endorphins have opioid activity and appear to modulate gonadotropin secretion.

Control of adenohypophyseal activity was not understood for many years, primarily because the functional connection between the brain and the anterior pituitary gland was unknown. In the 1930s, Popa and Fielding (Budapest medical student and university

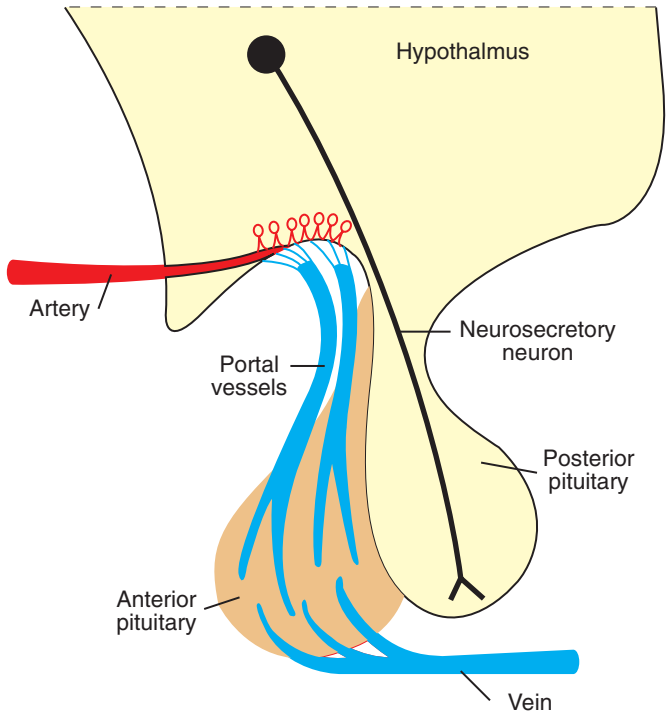
TABLE 33.1 Six Major Hormones Secreted by the Anterior Pituitary Gland

Hormone	Abbreviations
Glycoproteins	
Follicle-stimulating hormone	FSH
Luteinizing hormone (interstitial cell–stimulating hormone)	LH (ICSH)
Thyroid-stimulating hormone (thyrotropin)	TSH
Somatotropins	
Growth hormone (somatotropin)	GH
Prolactin	PRL
Proopiomelanocortin	
Corticotropin (adrenocorticotropic hormone)	ACTH

Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.



• **Fig. 33.15** Cleavage of proopiomelanocortin to yield ACTH and related peptides. By convention, the numbering of the amino acids begins with the first one of ACTH and then increases positively toward the carboxy terminal and negatively toward the amino terminal. Cleavage occurs at pairs of basic amino acids indicated by the circles. *ACTH,* Adrenocorticotropic hormone; *AL,* anterior lobe; *α-MSH,* alpha-melanocyte–stimulating hormone; *β-End,* beta-endorphin; *CLIP,* corticotropin-like intermediate lobe peptide; *IL,* intermediate lobe. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

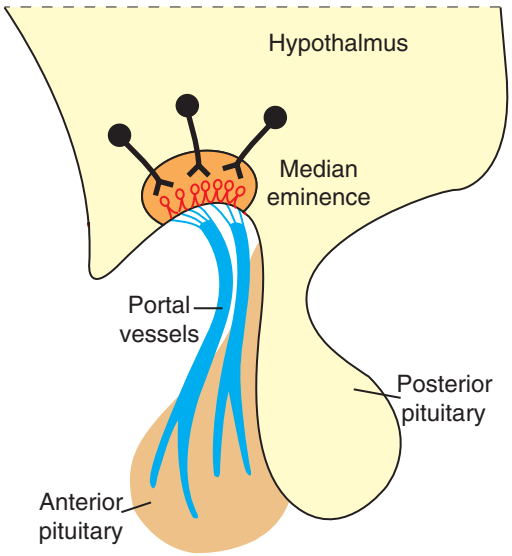


• **Fig. 33.16** Diagram of the hypothalamopituitary unit, contrasting the vascular connection between the brain and the anterior pituitary gland with the neuronal connection between the brain and the posterior pituitary gland. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

professor, respectively) described the vascular system that connects the hypothalamus with the pituitary gland but were unable to determine the direction in which blood flowed. In about 1950, Geoffrey Harris drew the important conclusion that the linkage involved blood passage from the hypothalamus to the anterior pituitary gland through the portal blood system previously described by Popa and Fielding (Fig. 33.16). The dorsal hypophyseal artery, which supplies nutrients and oxygen to the adenohypophysis (the ventral hypophyseal artery supplies the neurohypophysis), terminates in the median eminence as a capillary plexus. Blood from this plexus is drained by two veins that empty into sinusoidal capillaries of the *pars distalis,* completing the portal venous system (one vein supplies the ventral, central part of the *pars distalis;* the other supplies the dorsal, peripheral areas).

Adenohypophyseal Activity Is Controlled by Hypothalamic Releasing Hormones, Which Are Released Into the Portal System, Which in Turn Connects the Median Eminence of the Hypothalamus and the Anterior Pituitary Gland

Whereas neurons that compose the neurohypophysis are influenced directly by neural input within the hypothalamus, the imposition of a vascular system between the hypothalamus and the adenohypophysis requires a different type of control system. The hypothalamus produces regulatory or *hypophysiotropic* hormones, which are transported to and released within the median eminence (comparable to posterior lobe hormones) (Fig. 33.17). These regulatory hormones then pass via the portal venous system to the adenohypophysis, where they stimulate the release of the various anterior



• **Fig. 33.17** Hypothalamic neurosecretory neurons and hypothalamohypophyseal portal vessels. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

pituitary hormones. The synthesis of adenohypophyseal regulatory hormones is controlled by both neural and hormonal inputs at the level of the hypothalamus. Some of the hypophyseal hormones have been found in other areas of the brain and extraneural sites, including the gastrointestinal tract and the pancreas.

The initial isolation and identification of the hypothalamic hormones required large amounts of tissue as well as expertise in biochemistry. The first identified hypothalamic hormone, which controls ACTH release, was originally called *corticotropin-releasing factor* (now changed from *factor* to *hormone*). The initial work, done by Guillemin's group at the University of Houston in the early 1960s, required the collection, freezing, and transport of several hundred thousand sheep brains from abattoirs located in the western United States, as well as the subsequent dissection of the hypothalami. The hypothalamic hormones that have been characterized and the hormones they release include the following (Table 33.2):

- *Corticotropin-releasing hormone* (CRH). A 41–amino acid polypeptide that stimulates corticotropes to release all components of the proopiomelanocortin family of molecules including ACTH.
- *Gonadotropin-releasing hormone* (GnRH). A decapeptide that stimulates gonadotrope secretion of both FSH and LH.
- *Thyrotropin-releasing hormone* (TRH). A tripeptide that stimulates thyrotrope secretion of TSH.
- *Dopamine.* A catecholamine precursor of norepinephrine that inhibits lactotrope secretion of PRL and thyrotrope secretion of TSH.
- *Somatostatin.* A tetradecapeptide that inhibits somatotrope secretion of GH.
- *Growth hormone–releasing hormone* (GHRH). A 44–amino acid polypeptide that stimulates somatotrope secretion of GH.

Except for dopamine, all these hypophysiotropic hormones are peptides.

Previously, only four of the anterior pituitary hormones (FSH, LH, TSH, and ACTH) were considered to be *tropic;* that is, their main effect was stimulation of hormone secretion by specific endocrine organs located peripheral to the pituitary gland. More

recently, GH has been added to this list because GH stimulates the liver to produce *somatomedins, especially insulin-like growth factor-1* (IGF-1), which have a negative-feedback effect on GH secretion. PRL remains as the only *pars distalis* hormone for which negative-feedback inhibition has not been demonstrated through hormones produced by PRL target tissues.

The most important regulation of secretion of the protein hormones by the *pars distalis* is by feedback inhibition. One feedback system involves negative-feedback inhibition of the tropic pituitary hormone by interaction of the target organ hormone with the hypothalamus, as well as with the pituitary gland; this system is called a *long-loop feedback system* (Fig. 33.18). For example, cortisol is produced by the adrenal cortex as a result of ACTH stimulation, and cortisol in turn has a negative-feedback effect on ACTH production at the level of the hypothalamus and the anterior pituitary gland. *Short-loop feedback systems* have also been described; an anterior pituitary hormone such as ACTH has a direct negative-feedback inhibition of hormone secretion, in this case CRH, within the hypothalamus.

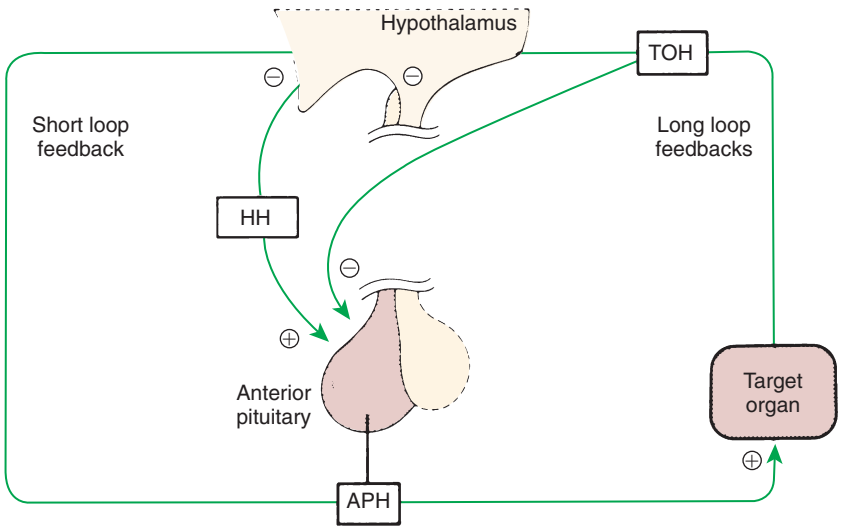
Even under conditions of negative-feedback inhibition, the secretion of anterior pituitary hormones is not constant. For example, even though estrogens exert a continuous, potent negative-feedback inhibition on gonadotropin secretion, gonadotropin secretion alternates between secretion and no secretion, with pulses of gonadotropins released into the vascular system. In the case of gonadotropins, the ovarian endocrine status influences the pulse rate and the amplitude of each pulse. Progesterone domination is associated with a decreased pulse rate and an increased pulse amplitude, whereas estrogen causes the opposite effect. The work of Irvine and Alexander has provided the best documentation of the precise relationship between secretory activity of hypothalamic regulatory and anterior pituitary hormones. Their data were obtained through analysis of hormones obtained from the intercavernous sinus, which collects venous blood from the pituitary gland of the horse.

Clinical syndromes of somatotropin deficiency and excess include pituitary dwarfism in the dog and acromegaly in the cat, respectively. Pituitary dwarfism results from destruction of the pituitary gland through a neoplastic, degenerative, or anomalous process. It may be associated with decreased production of other pituitary hormones, including TSH, ACTH, LH, FSH, and GH. *Pituitary dwarfism* is most common in German shepherd dogs aged 2 to 6 months. Other affected breeds include Carnelian bear dogs, Spitz-type dogs, toy pinschers, and Weimaraners. In German shepherd dogs, the disease is inherited as a simple autosomal recessive trait and occurs as a result of cystic Rathke's pouch. The first observable clinical signs of pituitary dwarfism are slow growth, noticed in the first 2 to 3 months of life, and mental retardation, usually manifested as difficulty in house training. Physical examination findings may include proportionate dwarfism, retained puppy hair coat, hypotonic skin, truncal alopecia, cutaneous hyperpigmentation, infantile genitalia, and delayed dental eruption. Clinicopathological

TABLE 33.2 Major Hypophysiotrophic Hormones

Hormone	Abbreviation	Site of Origin
Thyrotropin-releasing hormone	TRH	Paraventricular nucleus
Gonadotropin-releasing hormone	GnRH	Preoptic area of hypothalamus
Growth hormone–inhibiting hormone (somatostatin)	GHIH	Anterior hypothalamic area
Growth hormone–releasing hormone	GHRH	Arcuate nucleus
Corticotropin-releasing hormone	CRH	Paraventricular nucleus
Prolactin-releasing factor	PRF	Unknown
Prolactin-inhibiting hormone (or dopamine)	PIH	Arcuate nucleus

Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.



• **Fig. 33.18** Regulation of anterior pituitary hormone (*APH*) secretion by hypophysiotropic hormones (*HH*), short-loop negative feedback, and long-loop negative feedback by target organ hormones (*TOH*). *Plus signs* indicate stimulation, and *minus signs* indicate inhibition. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.)

features include eosinophilia, lymphocytosis, mild normocytic-normochromic anemia, hypophosphatemia, and occasionally hypoglycemia resulting from secondary adrenal insufficiency. Differential diagnoses include other causes of stunted growth, such as hypothyroid dwarfism, portosystemic shunt, diabetes mellitus, hyperadrenocorticism, malnutrition, and parasitism. Diagnosis is made by measuring serum GH concentrations (feline and canine assay no longer commercially available) or serum somatomedin C (IGF-1). The advantage of IGF-1 is that it is stable and not species-specific. There is usually a subnormal response to exogenous TSH and ACTH stimulation tests; furthermore, endogenous TSH and ACTH are decreased in affected dogs as a result of panhypopituitarism.

Acromegaly, or *hypersomatotropism,* is the condition resulting from chronic excessive GH secretion in the adult animal. Canine acromegaly is an extremely rare disorder observed after the administration of progestational compounds for suppression of estrus in intact female dogs. The disease is caused by excessive secretion of GH from mammary cells under the influence of exogenous progesterone. Acromegaly in cats, as in humans, is caused by a GH-secreting tumor of the anterior pituitary gland. Such tumors in cats grow slowly and may be present for a long time before the onset of clinical signs. Feline acromegaly occurs in older (8- to 14-year-old) cats and occurs more frequently in males. Canine acromegaly occurs in intact female dogs given progestational compounds for estrus prevention.

Clinical signs of uncontrolled diabetes mellitus (DM) are often observed as the first manifestations of acromegaly; therefore polydipsia, polyuria, and polyphagia are the most common presenting signs. Net weight gain of lean body mass in animals with uncontrolled DM is a key sign of acromegaly. Organomegaly, including renomegaly (observed in both cats and humans with acromegaly), hepatomegaly, and enlargement of endocrine organs, is also observed. Some dogs and cats show the classic enlargement of extremities, body size, jaw, tongue, and forehead that is characteristic of acromegaly in humans. Some of the most striking manifestations of acromegaly occur in the musculoskeletal system, such as an increase in muscle mass and growth of the acral segments of the body, including the paws, chin, and skull. Cardiovascular abnormalities, such as cardiomegaly (as determined radiographically and echocardiographically), systolic murmurs, and congestive heart failure, develop late in the course of the disease. Azotemia develops late in the disease in approximately 50% of acromegalic cats. Neurological signs of acromegaly in humans, such as peripheral neuropathies (paresthesias, carpal tunnel syndrome, sensory and motor defects), and parasellar manifestations, such as headache and visual field defects, are not generally detected in acromegalic small animals.

Impairments in glucose tolerance and insulin resistance that result in DM are observed in all cats and most dogs with acromegaly. Measurement of endogenous insulin reveals dramatically increased serum insulin concentrations. Despite severe insulin resistance and hyperglycemia, ketosis is rare in acromegalic animals. Feline acromegaly should be suspected in any diabetic cat (especially males) that has severe insulin resistance. A definitive diagnosis of acromegaly requires documentation of increased plasma GH or IGF-1 concentrations. Unfortunately, feline and canine GH assays are no longer commercially available. However, IGF-1 is a non–species-specific assay that can be used to identify suspected cases of acromegaly. In cats, it is estimated that almost 30% of cats with diabetes who do not go into remission on an ultra-low carbohydrate diet and insulin therapy are acromegalic.

At this time the most definitive test for the diagnosis of acromegaly in cats is computed tomography (CT) or magnetic resonance imaging (MRI) of the pituitary region coupled with increased serum IGF-1 concentrations. CT findings, coupled with the exclusion of other disorders that cause insulin resistance (hyperthyroidism, hyperadrenocorticism), in cats that exhibit clinical signs of acromegaly should lead the clinician to a diagnosis of acromegaly.

Acknowledgement

The authors would like to acknowledge and thank the late Dr. George H. Stabenfeldt as a previous contributor to this chapter. Some of his original work has been incorporated into this edition.

CLINICAL CORRELATIONS

Equine Pituitary Pars Intermedia Dysfunction (PPID)

History

You are called to examine a 20-year-old mare whose owner complains that the mare has had stiffness in her legs for the past 9 months. The mare has been used as a broodmare and delivered a foal the past spring (and for the previous 7 years). She failed to conceive last spring, and now, in the early summer of the next year, she has yet to exhibit normal estrous cycles.

Clinical Examination

As you gain a general perspective on the mare, you notice that she appears to have been clipped recently. She is not a show mare, and because it is early summer, you inquire why she has been clipped. The owner indicates that the mare has been slow to shed this spring, and she is tired of seeing the mare with a rough hair coat. The finding of a long hair coat out of season prompts you to ask about the water consumption of the mare; the owner indicates that the mare drank more water (and urinated accordingly) than would be expected. You examine the feet and find that the soles appear slightly "dropped"; you find a small abscess in the sole of one of the feet.

Comment

The main clue regarding the nature of the disease is the presence of a long hair coat out of season; this is the *sine qua non* of the disease. The usual complaints of owners of horses with PPID (formerly known as equine Cushing's disease) are related to chronic processes, such as pneumonia, laminitis, or weight loss, the latter often associated with parasitism and an inability to masticate properly because of tooth problems. Broodmares progressing into PPID often have a recent history of infertility after a successful broodmare career. Although the cause of infertility is not known, a disturbance of gonadotropin secretion is likely, along with a perturbation of the proopiomelanocortin system.

The disease represents a classic case of loss of control of the intermediate lobe of the pituitary gland by the hypothalamus, in this case the loss of dopaminergic control. Under normal conditions, melanotropes of the intermediate lobe process proopiomelanocortin to alpha-melanocyte–stimulating hormone (α-MSH) and acetylated β-endorphin residues 1 to 31, and nonopiate active carboxy-terminally shortened 1 to 26, or 1 to 27, endorphin. In the absence of dopamine, the melanotropes produce α-MSH, as well as β-endorphin, 1 to 31 (the active form), and small amounts of corticotropin (ACTH); the latter stimulates glucocorticoid production by the adrenal cortex. The negative-feedback control system fails in this situation because the melanotropes do not have glucocorticoid receptors, even under normal conditions. The result is unchecked synthesis and secretion of melanotrope products, including ACTH, and unchecked glucocorticoid

CLINICAL CORRELATIONS—cont'd

secretion. Activity of the corticotropes in the *pars distalis* is decreased because of negative-feedback inhibition by the glucocorticoids. One of the long-term effects of excess glucocorticoid secretion is muscle wasting, a common finding in these animals. Also, typical manifestations of the disease include polydipsia and polyuria, which result from compression of the *pars nervosa* by the enlarging pars intermedia and reduction in antidiuretic hormone (ADH) synthesis.

Although there is hyperplasia of the intermediate lobe in PPID, it has not been established whether this disease occurs because of autonomous hyperplasia of the intermediate lobe or, conversely, whether the hyperplasia occurs because of gradual loss of dopaminergic control by the hypothalamus.

Treatment

Standard treatment is to provide the best possible husbandry for animals with PPID disease. This care includes parasite control, floating of teeth, providing good nutrition, and taking proper care of the feet. Pergolide (a dopamine agonist) is also commonly used in an effort to control ACTH oversecretion in this disease. Cyproheptadine, which is a serotonin antagonist, has also been used extensively.

Agalactic Mare

History

A client calls you with a mare that has just foaled, and the mare does not appear to have any milk.

Clinical Examination

The mare and foal otherwise appear to be normal. However, the mare has little milk for the foal. On examining the placenta, the placenta appears to be a little bit thickened.

Comment

This mare likely has agalactia due to ingestion of fescue, likely contaminated with the fungus *Claviceps* sp., contained within the hay. The fungus produces an ergot alkaloid that is a dopamine agonist and thus inhibits prolactin secretion. Ergot alkaloids can also produce thickened fetal membranes. Domperidone can be given to inhibit dopamine, which will block the inhibition of prolactin. The increase in prolactin should increase milk production. Additionally, milk production could be increased by giving oxytocin several times daily. If the foal is nursing, this will stimulate oxytocin release which should stimulate milk let down.

Treatment

The mare can be given a combination of domperidone as well as oxytocin to stimulate milk production. Depending on when the domperidone is started, it can take up to 10 to 14 days to see the full effects. The foal may likely have to be supplemented with a milk replacement as well to be sure the foal is getting sufficient calories.

Practice Questions

1. In general, hormones are classified as proteins, peptides, and steroids. Which one of the following hormones is a peptide?
- **a.** Growth hormone
- **b.** Insulin
- **c.** Vasopressin
- **d.** Dopamine
- **e.** Epinephrine
- **f.** Melatonin

2. In general, steroid hormones are classified as mineralocorticoid, glucocorticoid, and sex steroids. Which one of the following hormones is a glucocorticoid?
- **a.** Aldosterone
- **b.** Growth hormone
- **c.** Cortisol
- **d.** Testosterone
- **e.** Estrone

3. Direct feedback control of corticotropin-releasing hormone by ACTH is termed:
- **a.** Negative feedback.
- **b.** Positive feedback.
- **c.** Short-loop feedback.
- **d.** Long-loop feedback.

4. Hormones from the proopiomelanocortin family are synthesized from precursor hormones produced in either the *pars distalis* or the *pars intermedia*. The two main hormones produced by these two lobes (in respective order) are:
- **a.** α-MSH and endorphin
- **b.** ACTH and endorphin
- **c.** α-MSH and ACTH
- **d.** ACTH and α-MSH
- **e.** ACTH and α-lipotropin

5. Increased hormonal activity that occurs during daylight hours is termed _____ rhythm.
- **a.** Circadian
- **b.** Diurnal
- **c.** Nocturnal
- **d.** Ultradian

Bibliography

Eiler H. Endocrine glands. In: Reece WO, eds. *Dukes' Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Ettinger SJ, Feldman EC, Cote E, eds. *Textbook of Veterinary Internal Medicine*. 8th ed. St. Louis: Elsevier; 2017.

Feldman EC, Nelson RW, Reusch CE, Scott-Moncrieff JCR, Behrend EN, eds. *Canine and Feline Endocrinology*. 4th ed. Philadelphia: Saunders; 2015.

Frazer GS. The pregnant mare. In: Reed SM, Bayly WM, Sellon DC, eds. *Equine Internal Medicine*. 3th ed. St Louis: Elsevier Saunders; 2010.

Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology*. Philadelphia: Saunders; 1987.

Melmed S, Polonsky KS, Larsen PR, Kronenberg HM. *Williams Textbook of Endocrinology*. 13th ed. Philadelphia: Elsevier Saunders; 2016.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames, Iowa: Iowa State University Press; 2003.

Troedsson MG, Christensen BW. Alerations in sexual function. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby Elsevier; 2015.

34

Endocrine Glands and Their Function

BRIAN K. PETROFF AND DEBORAH S. GRECO

KEY POINTS

The Thyroid Gland

1. The thyroid hormones are synthesized from two connected tyrosine molecules that contain three or four iodine molecules.
2. Thyroid hormones are stored outside the cell and attached to thyroglobulin in colloid.
3. The release of thyroid hormones involves transport of thyroglobulin with attached thyroid hormones into the cell, cleavage of the thyroid hormones from thyroglobulin, and release of hormones into the interstitial tissues.
4. Thyroid hormones are transported in the plasma attached to plasma proteins.
5. The main routes of metabolism of thyroid hormones are through deiodination or the formation of glucuronides and sulfates via hepatic mechanisms.
6. Thyroid hormones are the primary factors for the control of basal metabolism.
7. The ingestion of compounds that inhibit the uptake or organic binding of iodine blocks the thyroid's ability to secrete thyroid hormones and causes goiter.

The Adrenal Glands

1. The adrenal glands are composed of two organs: the outer gland (cortex) and the inner gland (medulla).

The Adrenal Cortex

1. The adrenal cortex has three zones: the zona glomerulosa, which secretes mineralocorticoids, and the zona fasciculata and the zona reticularis, which secrete glucocorticoids and sex steroids.
2. Adrenal corticosteroids are synthesized from cholesterol; the critical difference in the activity of these corticosteroids is related to the hydroxyl group on C-17 of glucocorticoids.
3. Adrenocortical hormones are carried in plasma in association with specific binding globulins (corticosteroid-binding globulin).
4. The metabolism of adrenocortical hormones involves the reduction of double bonds and conjugation of the steroids to glucuronides and sulfates.
5. One of the most important functions of glucocorticoids is control of metabolism, in particular the stimulation of hepatic gluconeogenesis.
6. Corticotropin (ACTH) is the pituitary hormone that regulates glucocorticoid synthesis by the adrenal cortex.
7. One of the most important clinical uses of glucocorticoids is the suppression of the inflammatory response.

The Adrenal Medulla

1. The synthesis of catecholamines is from tyrosine; the main catecholamine synthesized by the adrenal medulla is epinephrine.
2. The primary actions of catecholamines are on metabolism, especially those effects that increase the concentration of glucose.
3. The main factors that stimulate catecholamine secretion are hypoglycemia and conditions that produce stress.

Hormones of the Pancreas

1. The synthesis of insulin is biphasic; an acute phase involves the release of preformed insulin, and a chronic phase involves the synthesis of protein.
2. The metabolism of insulin involves splitting the A and B chains and reducing the chains to amino acids and peptides.
3. The main metabolic functions of insulin are anabolic.
4. Insulin deficiency (concentration or action) produces diabetes mellitus, which can culminate in diabetic ketoacidosis.
5. Dietary management is an important consideration in therapy for feline type 2 diabetes.
6. The most important functions of glucagon are to decrease glycogen synthesis, increase glycogenolysis, and increase gluconeogenesis.
7. Glucagon synthesis is stimulated by decreased glucose concentrations in the blood.
8. The main functions of somatostatin are to inhibit the secretion of hormones produced by the pancreas (insulin, glucagon, pancreatic polypeptide) and regulate growth hormone.

Calcium and Phosphate Metabolism

1. Calcium is important for many intracellular reactions, including muscle contraction, nerve cell activity, release of hormones through exocytosis, and activation of enzymes.
2. Phosphate is important for the structure of bone and teeth, and organic phosphate serves as part of the cell membrane and several intracellular components.
3. The most important body pool of calcium involved in homeostasis is the extracellular fluid component, but calcium in bone serves as a larger reservoir.
4. Parathyroid hormone (PTH) and calcitonin act together to provide calcium homeostasis with PTH increasing, and calcitonin decreasing, ionized calcium.

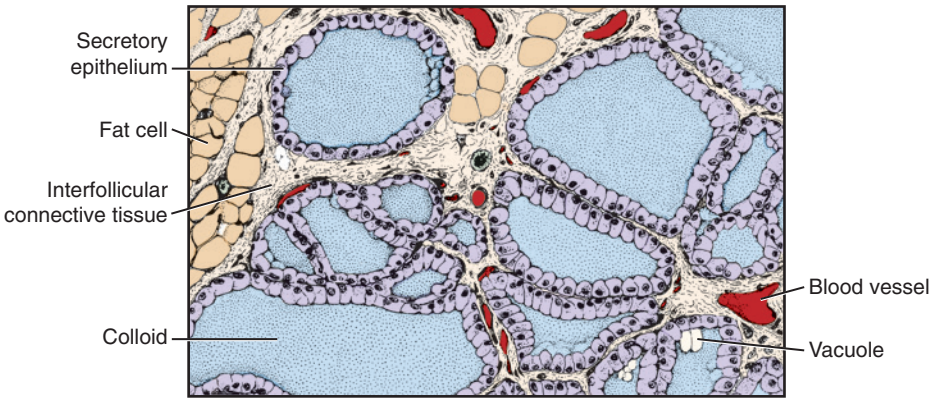
The Thyroid Gland

In most mammals, the thyroid glands are located caudal to the trachea at the level of the first or second tracheal ring. The thyroid gland is the most important endocrine gland for metabolic regulation. The glandular tissue has cells formed in a circular arrangement called a thyroid *follicle* (Fig. 34.1). The follicles are filled with a homogeneous-staining substance called *colloid,* which is the main storage site of the thyroid hormones. The follicular cells are cuboidal when the secretion is basal and are elongated when the cells are stimulated to release hormone. Another important endocrine cell, the *parafollicular cell,* or *C cell,* is located outside the follicles. This cell secretes *calcitonin,* a hormone important for the regulation of calcium. The activity of this hormone is discussed in the section on calcium metabolism.

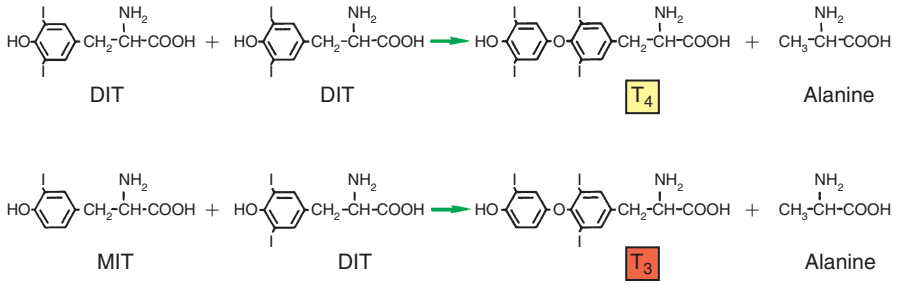
The Thyroid Hormones Are Synthesized From Two Connected Tyrosine Molecules That Contain Three or Four Iodine Molecules

The synthesis of thyroid hormone is unusual because a large amount of the active hormone is stored in the colloid outside the follicle cells within the lumen (or acinus) created by the circular arrangement of glandular cells. Two molecules are important for thyroid hormone synthesis: tyrosine and iodine. *Tyrosine* is a part of a large molecule (molecular weight, 660,000 D) called *thyroglobulin* that is formed within the follicle cell and secreted into the lumen of the follicle. *Iodine* is converted to iodide in the intestinal tract and then is transported to the thyroid, where the follicle cells effectively trap the iodide through an active transport process. This allows intracellular iodide concentrations to be 25 to 200 times higher than extracellular concentrations.

As iodide passes through the apical wall of the cell, it attaches to the ring structures of the tyrosine molecules, which are part of the thyroglobulin amino acid sequence. The tyrosyl ring can accommodate two iodide molecules; if one iodide molecule attaches, it is called *monoiodotyrosine,* and if two attach, it is called *diiodotyrosine.* The coupling of two iodinated tyrosine molecules results in the formation of the main thyroid hormones; two diiodotyrosine molecules form *tetraiodothyronine,* or *thyronine* (T_4), and one monoiodotyrosine and one diiodotyrosine molecule form *triiodothyronine* (T_3) (Fig. 34.2). A key enzyme in the biosynthesis of thyroid hormones is *thyroperoxidase* (which works in concert with an oxidant, hydrogen peroxide). Thyroperoxidase catalyzes the iodination of the tyrosyl residues of *thyroglobulin* and the formation of T_3 and T_4. In addition to the unusual molecular storage form of the hormone, thyroid hormones are also unique in that they are the only hormones that contain a halide (i.e., iodine).



• **Fig. 34.1** Histological features of the normal thyroid gland of the rat. All normal thyroid glands are structurally similar, although slight variations occur with age, diet, habitation, and sexual status (neutered or intact). The normal animals of the colony to which this rat belonged were maintained on a high-protein ration, which probably accounts for the slight hypertrophic condition of the secretory epithelium. (Modified from Turner CD, Bagnara JT. *General Endocrinology.* 6th ed. Philadelphia: Saunders; 1976.)



• **Fig. 34.2** Production of tetraiodothyronine (thyronine, T_4) and triiodothyronine (T_3) by the coupling of iodinated tyrosyl residues with thyroglobulin molecule. *DIT,* Diiodotyrosine; *MIT,* monoiodotyrosine. (Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

Thyroid Hormones Are Stored Outside the Cell and Attached to Thyroglobulin in Colloid

When thyroid hormones are synthesized, they remain in the extracellular acinar lumen until release. This extracellular storage of hormone within an endocrine gland is a unique storage arrangement. It allows the thyroid gland to have a large reserve of hormone. From a teleological standpoint, thyroid hormone is the most important hormone of metabolism; it allows mammals to withstand periods of iodine deprivation without an immediate effect on the production of thyroid hormones.

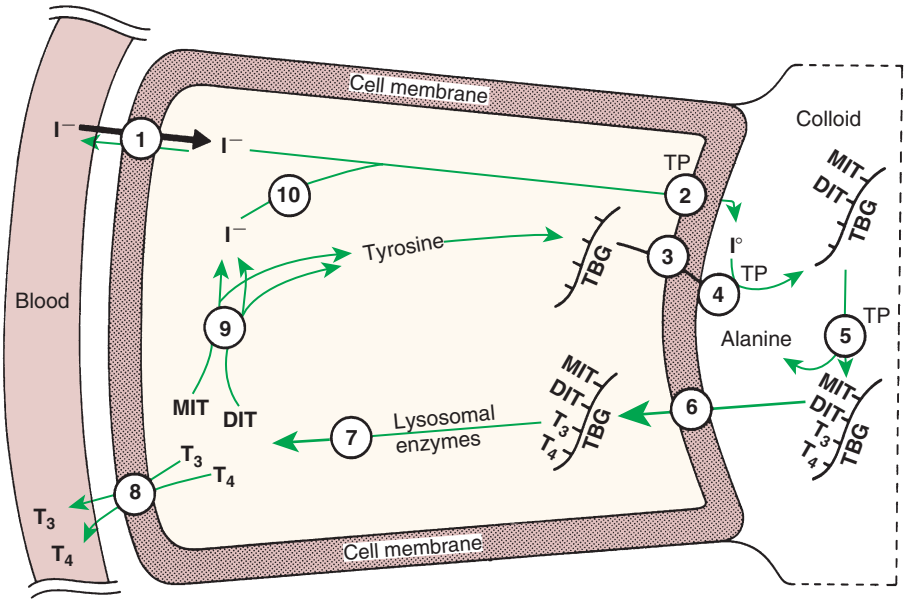
The Release of Thyroid Hormones Involves Transport of Thyroglobulin With Attached Thyroid Hormones Into the Cell, Cleavage of the Thyroid Hormones From Thyroglobulin, and Release of Hormones Into the Interstitial Tissues

For thyroid hormones to be released from the thyroid gland, thyroglobulin—with its attached monoiodotyrosine, diiodotyrosine, T_4 and some T_3 molecules—must be translocated into the follicle cell, and the hormones must be cleaved from thyroglobulin (Fig. 34.3). Key enzymes in this transfer are found in the lysosomes. On entering the cell, the thyroglobulin molecules fuse with lysosomes, and lysosomal enzymes cleave both the iodinated tyrosine molecules and the iodinated thyronines from the thyroglobulin molecule. The thyronines are released through the basal cell membrane (they freely pass through the cell membrane); monoiodotyrosine and diiodotyrosine are deiodinated by an enzyme called *iodotyrosine dehalogenase;* and both the iodide and the remaining tyrosine molecules are recycled to form new hormone in association with thyroglobulin.

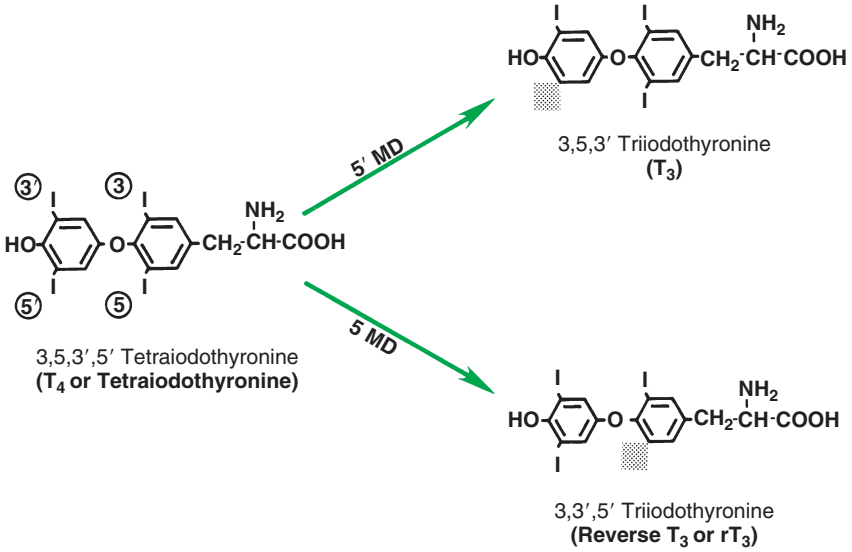
The thyroid gland itself produces mostly T_4 with small amounts of T_3. The majority of T_3 formation occurs outside the thyroid gland by deiodination of T_4. Tissues that have the highest concentration of deiodinating enzymes are those of the liver and kidneys, although muscle tissue produces more T_3 on the basis of relative size. The enzyme involved in the removal of iodide from the outer phenolic ring of T_4 in the formation of T_3 is called *5′-monodeiodinase* (Fig. 34.4). Another type of T_3 in which an iodide molecule is removed from the inner phenolic ring of T_4, a compound called *reverse T_3,* is also formed. Reverse T_3 has little of the biological effects of thyroid hormones and is formed only by the action of extrathyroidal deiodinating enzymes and not by activity of the thyroid gland.

Thyroid Hormones Are Transported in the Plasma Attached to Plasma Proteins

As indicated in Chapter 33, lipid-soluble hormones are transported in the vascular system through association with specific binding plasma proteins. There is considerable species variation in the proteins that bind thyroid hormones. The most important carrier protein is thyroxine-binding globulin (TBG), which has high affinity for T_4, although it also has low capacity because of its low concentration. TBG also is an important carrier protein for T_3. TBG has been reported in all domestic animals except the cat. *Albumin* is also involved in the transport of thyroid (and many other) hormones; however, albumin has low affinity for T_3 and T_4 but high capacity because of its high concentration in plasma. In the absence of TBG, albumin is the most important carrier of thyroid hormones. All species have a third plasma protein, thyroxine-binding prealbumin, which is specific for T_4 and has a specificity and capacity that are intermediate between those of TBG and albumin. The term *prealbumin* refers to the migration of the protein during electrophoresis, not to synthesis of the molecule.



• **Fig. 34.3** Depiction of follicular cell showing steps in the synthesis and release of triiodothyronine (T_3) and thyronine (T_4). The numbers identify the major steps: *1,* trapping of iodide; *2,* oxidation of iodide; *3,* exocytosis of thyroglobulin; *4,* iodination of thyroglobulin; *5,* coupling of iodotyrosines; *6,* endocytosis of thyroglobulin; *7,* hydrolysis of thyroglobulin; *8,* release of T_3 and T_4; *9,* deiodination of monoiodotyrosine (*MIT*) and diiodotyrosine (*DIT*); and *10,* recycling of iodide. *TBG,* Thyroxine-binding globulin; *TP,* thyroperoxidase. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.4** Structure and nomenclature of thyroxine and its conversion to the two triiodothyronines by 5′-monodeiodinase and 5-monodeiodinase (*MD*). *Shaded squares* indicate the sites of deiodination. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

As with all lipid-soluble hormones transported in plasma, most of the T_3 and T_4 is bound; little is free to interact with receptors on the cells of the target tissues. The amount of thyroid hormone free in plasma is remarkably low (e.g., in humans, 0.03% of T_4 and 0.3% of T_3). In dogs, the amount of free hormone is somewhat greater (slightly less than 1.0% for T_4 and slightly more than 1.0% for T_3) because of less affinity between plasma-binding proteins and thyroid hormones in canine plasma than in human plasma. The equilibrium between free and bound hormone is easily shifted because of physiological or pharmacological situations, such as illness or the increase in estrogen concentrations that occurs during pregnancy. Adjustments to maintain a normal amount of free hormone occur rapidly with a decline in the rate of metabolism or with stimulation of thyroid hormone production through the release of *thyroid-stimulating hormone* (TSH).

The Main Routes of Metabolism of Thyroid Hormones Are Through Deiodination or the Formation of Glucuronides and Sulfates via Hepatic Mechanisms

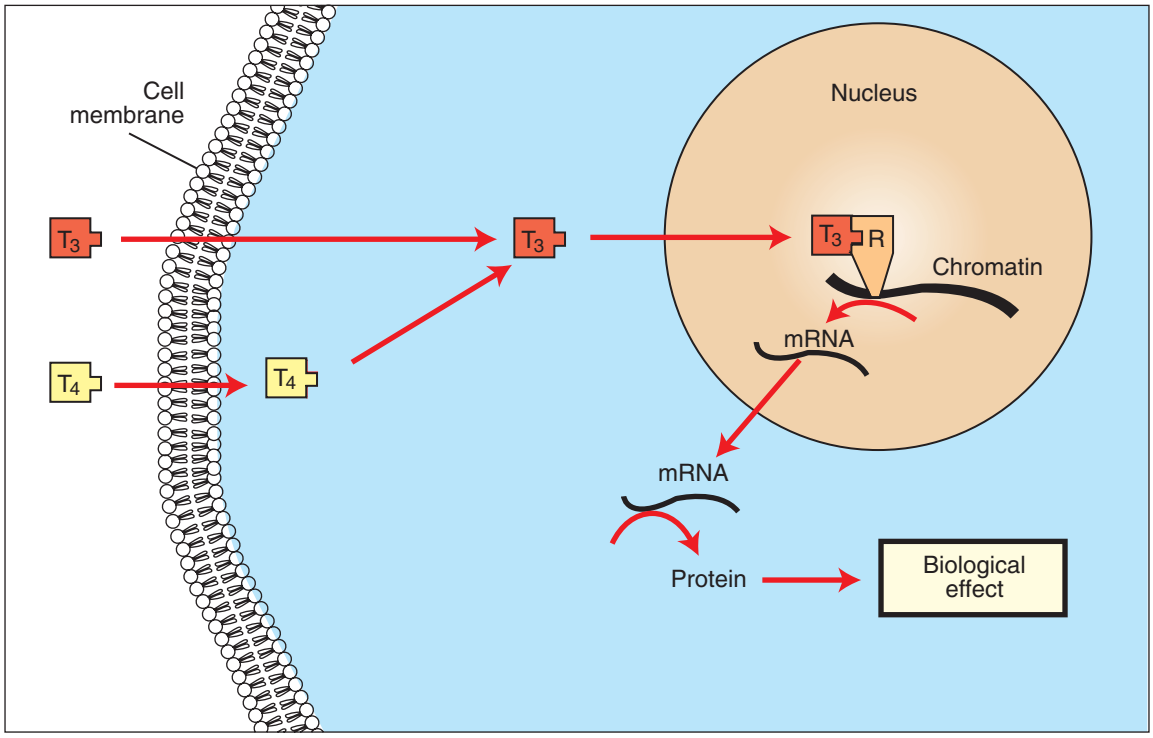
The main form of metabolism of thyroid hormones involves the removal of iodide molecules. Except for the T_3 formed from T_4, none of the deiodinated thyronine derivatives has any significant metabolic activity. The two enzymes involved in T_3 and reverse T_3 synthesis, 5′-deiodinase and 5-deiodinase, are also involved in the catabolism of thyroid hormones. Only these two enzymes are needed for catabolism because they do not differentiate between the 3 and 5 positions of the phenolic rings of the thyronines. Skeletal muscle, liver, and kidney tissues are important tissues involved in the catabolism of thyroid hormones through deiodination. The formation of thyroid hormone conjugates represents another form of inactivation; sulfates and glucuronides are formed mainly in the liver and kidneys. Conjugation is less common than deiodination as a means of metabolism of thyroid hormones. Another form of metabolism involves modification of the alanine moiety of the thyronines by either transamination or decarboxylation. The deiodinated and conjugated forms of the thyronines are eliminated primarily in the urine; unmetabolized thyronines are excreted with feces through bile secretion. Degradation of the conjugate forms in the feces results in the production of iodide molecules, which are reabsorbed as part of the *enterohepatic cycle.* Humans are more efficient than dogs in recovery of iodide both intrathyroidally and enterohepatically.

One of the striking aspects of thyroid hormones is their long half-lives in humans; T_3 has a half-life of 1 day and T_4 has a half-life of 6 to 7 days, whereas most other hormones have half-lives of minutes to hours. One reason for these long half-lives of thyroid hormones in humans is the large percentage of the circulating thyronines that are bound to the plasma proteins, which protects them from degradation. The difference in half-lives between T_3 and T_4 results from the tighter T_4 protein binding compared with T_3 and the resultant reduction in free circulating hormone. In contrast, the half-life for T_4 is relatively short in certain domestic species; dogs and cats exhibit a T_4 half-life of less than 24 hours.

Thyroid Hormones Are the Primary Factors for the Control of Basal Metabolism

The mechanism of action of thyroid hormones at the cellular level is based on their ability to penetrate the cell membrane even though they are amino acids; in essence, they are lipophilic. Although it is thought that thyroid hormones interact directly with the nucleus to initiate the transcription of *messenger ribonucleic acid* (mRNA) (Fig. 34.5), the presence of T_3 receptors has been reported on mitochondria.

Thyroid hormones are likely the primary determinants of basal metabolism. It is difficult to define their discrete physiological effects. However, many actions of thyroid hormones have been demonstrated through the creation of hypothyroid or hyperthyroid states. It has long been recognized that thyroid hormones increase oxygen consumption of tissues and, as a result, heat production. This effect is known as the calorigenic effect. One site of



• **Fig. 34.5** Proposed subcellular mechanism of thyroid hormone action. *mRNA,* Messenger ribonucleic acid; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

action of the *calorigenic effect* of thyroid hormones is within the mitochondrion.

Thyroid hormones affect carbohydrate metabolism in several ways, including increasing intestinal glucose absorption and facilitating the movement of glucose into both fat and muscle. Furthermore, thyroid hormones facilitate insulin-mediated glucose uptake by cells. Glycogen formation is facilitated by small amounts of thyroid hormones; however, glycogenolysis occurs after larger dosages.

Thyroid hormones, acting in concert with growth hormone, are essential for normal growth and development. This is accomplished in part by the enhancement of amino acid uptake by tissues and enzyme systems involved in protein synthesis.

Thyroid hormones affect many aspects of lipid metabolism, but the emphasis is placed on *lipolysis.* One particular effect of thyroid hormones is the tendency to reduce plasma cholesterol levels. This appears to involve both increased cell uptake of *low-density lipoproteins* (LDLs) with associated cholesterol molecules and a tendency for increased degradation of both cholesterol and LDL. These effects on lipid metabolism are usually seen in pathophysiological situations involving hypersecretion of thyroid hormone or in thyroid deficiency states, and hypercholesterolemia is a hallmark of thyroid deficiency. In this same context, the effects of thyroid hormones on metabolic processes, including carbohydrate, protein, and lipid metabolism, are often described as *catabolic.*

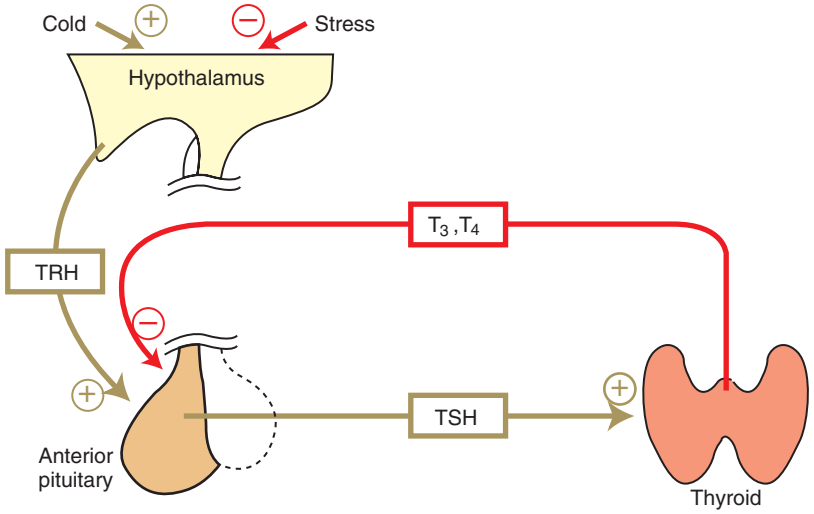
Thyroid hormones have noteworthy effects on the nervous and cardiovascular systems. The effects of the sympathetic nervous system are enhanced by the presence of thyroid hormones. This is thought to occur through thyroid stimulation of β-adrenergic receptors in tissues that are targets for the catecholamines, such as epinephrine and norepinephrine. In the central nervous system (CNS), thyroid hormones are important for normal development of tissues in the fetus and neonate; inhibition of mental activity occurs when thyroid hormone exposure is inadequate. In humans, persons with hypothyroid activity are mentally dull and lethargic, which suggests that normal CNS function in the adult depends on the presence of adequate amounts of thyroid hormone.

Thyroid hormones increase the heart rate and force of contraction, probably through their interaction with the catecholamines. This interaction is caused by an increase in tissue responsiveness through the induction of catecholaminergic β receptors by thyroid hormones. Blood pressure is elevated because of increased systolic pressure with no change in diastolic pressure; the end result is an increase in cardiac output. These responses are most easily observed in situations of increased thyroid activity. With regard to the effect of thyroid hormones on cardiovascular activity, it may be concluded that they are important for maintaining normal contractile activity of cardiac muscle, including the transmission of nerve impulses.

Thyroid hormone was used in classic experiments involving the metamorphosis of amphibian larvae. Thyroxine administration causes the differentiation of tadpoles into frogs, whereas thyroidectomy results in development into large tadpoles. Thyroid-induced metamorphosis is limited to amphibians, but thyroid hormones are important for many (subtle) aspects of differentiation in other classes of animals.

Thyroid hormone activity is usually defined in terms of tissue or organ responses to inadequate or excessive amounts of hormone. A more balanced view is that thyroid hormones are important for the *normal* metabolic activity of all tissues.

TSH, thyroid stimulating hormone or *thyrotropin,* is the most important regulator of thyroid activity. It acts through the initiation of cyclic adenosine *3′,5′-monophosphate* (cAMP) formation and the phosphorylation of protein kinases. Thyrotropin secretion is regulated by thyroid hormones through negative-feedback inhibition



• **Fig. 34.6** Hypothalamopituitary-thyroid axis. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. *T_3,* Triiodothyronine; *T_4,* thyronine; *TRH,* thyrotropin-releasing hormone; *TSH,* thyroid-stimulating hormone. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

of the synthesis of *thyrotropin-releasing hormone* (TRH) at the level of the hypothalamus and by inhibition of TSH activity at the level of the pituitary gland (Fig. 34.6).

The Ingestion of Compounds That Inhibit the Uptake or Organic Binding of Iodine Blocks the Thyroid's Ability to Secrete Thyroid Hormones and Causes Goiter

An inability to secrete adequate amounts of thyroid hormone often leads to the enlargement of the thyroid gland, a condition known as *goiter.* In many places in the world, this condition is, or has been, caused by a deficiency of iodine in the diet. This has largely been corrected through the use of iodized salt. Certain plants, such as cruciferous plants (e.g., cabbage, kale, rutabaga, turnip, rapeseed), contain a potent antithyroid compound called *progoitrin,* which is converted into goitrin within the digestive tract. *Goitrin* interferes with the organic binding of iodine. Many of the goitrogenic feeds also contain thiocyanates, which interfere with the trapping of iodine by the thyroid gland. The feeding of excess iodine can sometimes overcome the effects of thiocyanate but has less influence on overcoming the effects of goitrin. Studies of these phenomena have led to the development of compounds for the treatment of hyperthyroidism, the most potent being the thiocarbamides, *thiourea* and *thiouracil.* Other antithyroid drugs include sulfonamides, *p*-aminosalicylic acid, phenylbutazone, and chlorpromazine.

Hypothyroidism in Dogs

Hypothyroidism is most common in the dog, and the usual etiology of primary hypothyroidism is lymphocytic thyroiditis. Congenital hypothyroidism may be caused by thyroid dysgenesis, dyshormonogenesis, T_4 transport defects, goitrogens, or, in rare cases, iodine deficiency. Secondary hypothyroidism may be a secondary effect of pituitary tumors, radiation therapy, or ingestion of endogenous or exogenous glucocorticoids. Tertiary hypothyroidism can be acquired, as in the case of hypothalamic tumors, or can be congenital as a result of defective TRH or TRH receptor defects.

The signalment of hypothyroid dogs carries a distinct breed predisposition; high-risk breeds manifest symptoms as early as 2 to 3 years of age, and low-risk breeds manifest symptoms at a slightly older age (4–6 years of age). There are a number of breeds with elevated risk of autoimmune thyroiditis and hypothyroidism including English Setters, Eurasiers, Spanish Water dogs, Irish Red and White Setters, Dalmatians, Boxers, Shetland Sheepdogs, Tibetan Terriers, Kuvasz, German Wirehaired Pointers, Rhodesian Ridgebacks, and others.

Clinical signs of hypothyroidism are gradual and subtle in onset; lethargy and obesity are most common. Dermatological evidence of hypothyroidism is the next most common clinical finding. Symmetrical truncal or tail head alopecia is a classic finding in hypothyroid dogs. The skin is often thickened because of myxedematous accumulations in the dermis. Common hair coat changes seen in the hypothyroid dog include dull dry hair, poor hair regrowth after clipping, and presence or retention of puppy hair.

Cardiovascular signs of hypothyroidism include bradycardia, decreased cardiac contractility, and atherosclerosis, but these are uncommon presenting complaints. Neuromuscular signs such as myopathies and megaesophagus are also uncommon manifestations of canine hypothyroidism. Neuropathies, including bilateral or unilateral facial nerve paralysis, vestibular disease, and lower motor neuron disorders, are occasionally seen in hypothyroid dogs. Myxedema coma is an unusual finding in hypothyroid dogs and is secondary to myxedematous fluid accumulations in the brain and severe hyponatremia. Less common signs of hypothyroidism include reproductive disorders in female dogs, such as prolonged interestrous intervals, silent heat, and delivery of weak or stillborn puppies. Corneal lipid deposits and gastrointestinal (GI) problems such as constipation are occasionally observed in hypothyroid dogs.

Clinicopathological findings, such as anemia resulting from erythropoietin deficiency, decreased bone marrow activity, and decreased serum iron and iron-binding capacity, are observed in about 25% to 30% of hypothyroid dogs. Hypercholesterolemia is seen in approximately 75% of hypothyroid dogs because of altered lipid metabolism, decreased fecal excretion of cholesterol, and decreased conversion of lipids to bile acids. Hyponatremia, a common finding in humans with hypothyroidism, is observed as

a mild decrease in serum sodium in about 30% of hypothyroid dogs in one study. Hyponatremia is caused by an increase in total body water as a result of impaired renal excretion of water and by retention of water by hydrophilic deposits in tissues. An unusual clinicopathological feature of hypothyroidism is increased serum creatine phosphokinase levels, possibly as a result of hypothyroid myopathy.

Diagnosis is typically based on measurement of serum basal total thyroxine (TT_4) and triiodothyronine (TT_3) concentrations, serum free T_4 (FT_4) concentrations, and endogenous canine serum TSH levels (Table 34.1) in combination with clinical signs. The many variables that affect T_4 include age, breed, environmental and body temperature, diurnal rhythm, obesity, and malnutrition. For example, sighthound breeds appear to have lower normal TT_4 and FT_4 concentrations. Obese dogs have mild increases in serum TT_4 concentrations. In puppies, the serum TT_4 concentration is two to five times higher than in adult dogs. Furthermore, there is an age-related decline in serum TT_4 concentrations and response to TSH stimulation in dogs. *Euthyroid sick syndrome* is characterized by a decrease in serum TT_4 and an increase in reverse T_3. Concurrent illnesses such as diabetes mellitus (DM), chronic renal failure, hepatic insufficiency, and infections can cause euthyroid sick syndrome, resulting in decreases in serum TT_4 concentrations. Drugs such as anesthetics, phenobarbital, primidone, diazepam, trimethoprim-sulfa, quinidine, phenylbutazone, salicylates, and glucocorticoids can also decrease serum basal TT_4 concentrations.

TABLE 34.1 Serum T_4 and T_3 Values by Radioimmunoassay

Species[a]	T_4 (mg/dL)	T_3 (ng/dL)
Equine		
M ± SD	1.63 ± 0.51	77.1 ± 45.75
Range	0.95–2.38	31–153
Bovine		
M ± SD	6.22 ± 2.03	92.50 ± 53.61
Range	3.60–8.9	41–170
Caprine		
M ± SD	3.45 ± 0.47	145.9 ± 29.32
Range	3.0–4.23	88–190
Ovine		
M ± SD	4.41 ± 1.13	99.6 ± 27.34
Range	2.95–6.15	63–150
Porcine		
M ± SD	3.32 ± 0.80	89.8 ± 36.7
Range	1.70–4.68	43–140
Canine		
M ± SD	1.15 ± 0.38	96.2 ± 21.39
Range	0.70–2.18	63–130
Feline		
M ± SD	2.02 ± 0.61	64.7 ± 20.62
Range	1.18–2.95	39–112

[a]*N* = 10, for all species listed.
T_3, Triiodothyronines; T_4, thyronine; *M ± SD,* median plus/minus standard deviation.
From McDonald LE. *Veterinary Endocrinology and Reproduction.* 4th ed. Philadelphia: Lea & Febiger; 1989.

Free thyroid hormone concentrations, or unbound T_4 and T_3 (FT_4 and FT_3), are used in human medicine to differentiate between euthyroid sick syndrome and true hypothyroidism. In humans, the diagnostic accuracy of a single FT_4 measurement is approximately 90%. Measurement of FT_4 concentrations is achieved by equilibrium dialysis (*gold standard*) or analog immunoassays. Theoretically, FT_4 is not subject to spontaneous or drug-induced changes that occur with TT_4. Results of early studies, classifying dogs as hypothyroid on the basis of TSH stimulation tests, indicated that FT_4 measurements by equilibrium dialysis were 90% accurate, whereas other FT_4 assays (analog assays) were no better than TT_4. Glucocorticoids decrease both FT_4 fraction and TT_4 in dogs.

The measurement of endogenous TSH allows assessment of neuroendocrine feedback from thyroid hormones. With thyroid gland failure, decreases in serum FT_4 and TT_4 are sensed by the pituitary gland, resulting in an increase in serum endogenous TSH concentration. In humans, when endogenous TSH concentrations are increased and FT_4 concentrations are decreased, diagnostic accuracy for primary hypothyroidism approaches 100%. As FT_4 concentration falls, there is a logarithmic increase in serum endogenous TSH concentration, which makes the TSH assay the most sensitive test for the detection of early hypothyroidism. The use of endogenous TSH concentration *alone* is not recommended as a method of assessing thyroid function in animal species. Additionally, 20% to 30% of hypothyroid dogs have normal concentrations of TSH despite loss of T_3 and T_4.

The antithyroglobulin autoantibody test (TgAA) is intended to detect thyroid pathology before the onset of hypothyroidism in dogs with autoimmune thyroiditis. This test is used to identify dogs with hereditary thyroid disease before breeding. For many years, the TSH stimulation test was considered the gold standard for diagnosis of hypothyroidism in dogs. Unfortunately, this test does not differentiate between early hypothyroidism and euthyroid sick syndrome, and does not identify dogs with secondary or tertiary hypothyroidism. Furthermore, exogenous bovine TSH is no longer commercially available. Other thyroid function tests include the TRH stimulation test, thyroid scan, and thyroid biopsy. However, each test has drawbacks (expense, inaccuracy, or invasiveness).

In summary, diagnosis of canine hypothyroidism is based on signalment, historical findings, physical examination findings, clinicopathological features, and confirmation with a battery of thyroid function tests. The author examines TT_4 and TSH initially, followed by FT_4 by dialysis. If all measurements are abnormal, the dog is hypothyroid. If two of the three are abnormal, secondary hypothyroidism (low FT_4, low TSH) or early primary hypothyroidism (high TSH, low FT_4) is possible. If only one of the three thyroid measurements is abnormal, the dog should be reevaluated in 3 to 6 months.

Hyperthyroidism in Cats

Hyperthyroidism is the most common endocrinopathy of cats and is caused by adenomatous hyperplasia of the thyroid gland. Middle-aged to older cats are typically affected, and there is no predilection for breed or sex, although incidence is higher in purebred than in mixed breed cats.

Balanced pet foods normally provide sufficient iodine but vary widely in iodine content. The effects of this feed variation have been theorized to be important in cats, but there is no data to support or refute the theory. Although acute changes in dietary iodine have been associated with changes in FT4 in cats, most chronic changes in dietary iodine are associated with "adaptation" of the thyroid gland and therefore are unlikely to be the cause of feline hyperthyroidism.

As noted earlier, goitrogens can result in hypothyroidism. However, some have theorized that chronic exposure to goitrogens can lead to toxic nodular goiter resulting in hyperthyroidism. Flavinoids from soy proteins have been suggested to play a role in the pathogenesis of hyperthyroidism in cats. Polyphenolic soy isoflavones, such as genistein and daidzein, were identified in almost 60% of dry cat foods tested. This contradicts the epidemiological data which shows that hyperthyroidism is less common in cats that were fed dry foods. Some dry foods contain isoflavone contents consistent with levels shown to interfere with thyroid function by inhibiting thyroperoxidase in rats and 5′-deiodinase activity in cats. In a prospective study of 18 clinically normal cats eating a soy diet (400 mg isoflavones/kg diet), total T_4 and FT_4 concentrations were modestly, but significantly, increased, whereas T_3 concentrations were unchanged. However, many studies in humans have shown no detrimental effect of soy isoflavones on thyroid function particularly when incorporated into a balanced diet with adequate iodine intake.

Canned cat food has been implicated as a cause of feline hyperthyroidism in multiple epidemiological studies. The suspected goitrogen is bisphenol-A-diglycidyl ether (BADGE), a substance used in making the liner of easy-open "pop-top" cans. It is suspected that this compound can leach into the foods and be consumed by cats. Although this BADGE-based lining is generally considered safe and is used for foods for human consumption, cats may be more susceptible to toxic effects of this compound because they have a greatly reduced ability to detoxify it via hepatic glucuronidation. Bisphenol A also reduces triiodothyronine binding and causes increased TSH secretion resulting in hyperthyroidism and goiter in rats and some humans. Although cat studies may not be available, rodent studies show a very high safety margin. It should be noted that epidemiological studies showing associations are not the same as cause and effect. More than 90% of cats in the United States consume commercial pet foods as their primary nutritional source, and relatively few develop hyperthyroidism.

The disease in cats is similar to toxic nodular goiter in humans and is characterized by autonomous growth of thyroid follicles. The pathogenesis of toxic nodular goiter is an abnormality in the signal transduction of the thyroid cell. The TSH receptor on the thyroid cells activates receptor-coupled guanosine triphosphate-binding proteins (G proteins; see Chapter 1). Uniquely, the thyroid cell proliferation and hormone production are both controlled by the TSH receptor-G-protein-cAMP signaling. Overexpression of stimulatory G proteins and underexpression of inhibitory G proteins have been demonstrated in some humans with toxic nodular goiter. Mutations of the TSH receptor that result in the receptor remaining activated without ligand (i.e., TSH) have also been reported in humans with toxic nodular goiter.

In hyperthyroid cats, the same abnormalities have been investigated, and it appears that activation mutation (activation without ligand) of the TSH receptor may be part of the pathogenesis of feline hyperthyroidism in some cats. Furthermore, abnormalities of G proteins, specifically significantly decreased G inhibitory protein expression, have been described in tissues from hyperthyroid cats. Feline hyperthyroidism does not appear to involve activating antibodies against the TSH receptor—the mechanism of Graves' disease in humans.

Although the use of cat litter has been associated with an increased risk of hyperthyroidism, this may simply be a marker of cats that are kept indoors. Indoor cats are likely to live longer and hence have a higher risk of developing hyperthyroidism. Exposure to pesticides and herbicides has been associated with thyroid abnormalities in other species. In particular, the use of flea control products was associated with an increased risk of developing hyperthyroidism; however, no specific product or ingredient could be identified.

One report implicated brominated flame retardants (BFRs) as carcinogens/goitrogens possibly associated with feline hyperthyroidism. Coincidently BFRs were introduced 30 years ago at the same time that feline hyperthyroidism emerged. Bromide, a halide, is an intriguing agent to implicate in feline hyperthyroidism because of the unique composition of thyroid hormones which contain the halide iodide. In this report, serum levels of lipid-adjusted serum polybrominated diphenyl ethers (PBDE) were 10- to 400-fold higher than those found in humans. It has been theorized that findings of higher than human PBDE serum levels in cats is in accord with the most consistently identified risk factor, which is indoor living. The authors also propose that cats are at increased risk because of meticulous grooming behavior and increased exposure to flame retardants in furniture and carpets. The smaller size of cats, relative to humans, is also a possible risk factor for increased serum levels of PBDEs.

Hyperthyroidism is characterized by hypermetabolism; therefore, polyphagia, weight loss, vomiting, diarrhea, polydipsia, and polyuria are the most prominent features of the disease. Activation of the sympathetic nervous system is also seen; hyperactivity, tachycardia, pupillary dilation, and behavioral changes are characteristic of the disease in cats. Long-standing hyperthyroidism leads to hypertrophic cardiomyopathy, high-output heart failure, and cachexia, which may lead to death.

Clinicopathological features of hyperthyroidism include erythrocytosis and a stress leukogram (neutrophilia, lymphocytosis) caused by increased circulating catecholamine concentrations. Increased catabolism of muscle tissue in hyperthyroid cats may result in increased levels of blood urea nitrogen (BUN) but not creatinine. In fact, glomerular filtration rate (GFR) is increased in hyperthyroid cats, and this increase may mask underlying renal insufficiency. Although hyperthyroidism increases GFR, the effect of thyroid hormone excess on the urinalysis is variable. Most cats, however, have decreased urine specific gravity, particularly if they are exhibiting polyuria as a clinical sign. Increased metabolic rate results in liver hypermetabolism; therefore, serum activities of liver enzymes (alanine aminotransferase, aspartate aminotransferase) increase in 80% to 90% of hyperthyroid cats. Serum cholesterol decreases, not because of decreased synthesis, but rather because of increased hepatic clearance mediated by thyroid hormone excess.

Feline hyperthyroidism is diagnosed through measurement of TT_4. Because the disease has become more common and recognized in its early stages, FT_4 concentrations have been shown to be more diagnostic of early or "occult" hyperthyroidism. T_3 suppression testing can be used to confirm hyperthyroidism in borderline cases. In this test, exogenous T_3 administration results in at least a 50% decrease in T_4 in normal cats, but hyperthyroid cats fail to respond to the negative-feedback effects of increased T_3. Thyroid scintigraphy can be used to visualize hyperfunctional thyroid glands in affected cats as well.

The Adrenal Glands

The Adrenal Glands Are Composed of Two Organs: the Outer Gland (Cortex) and the Inner Gland (Medulla)

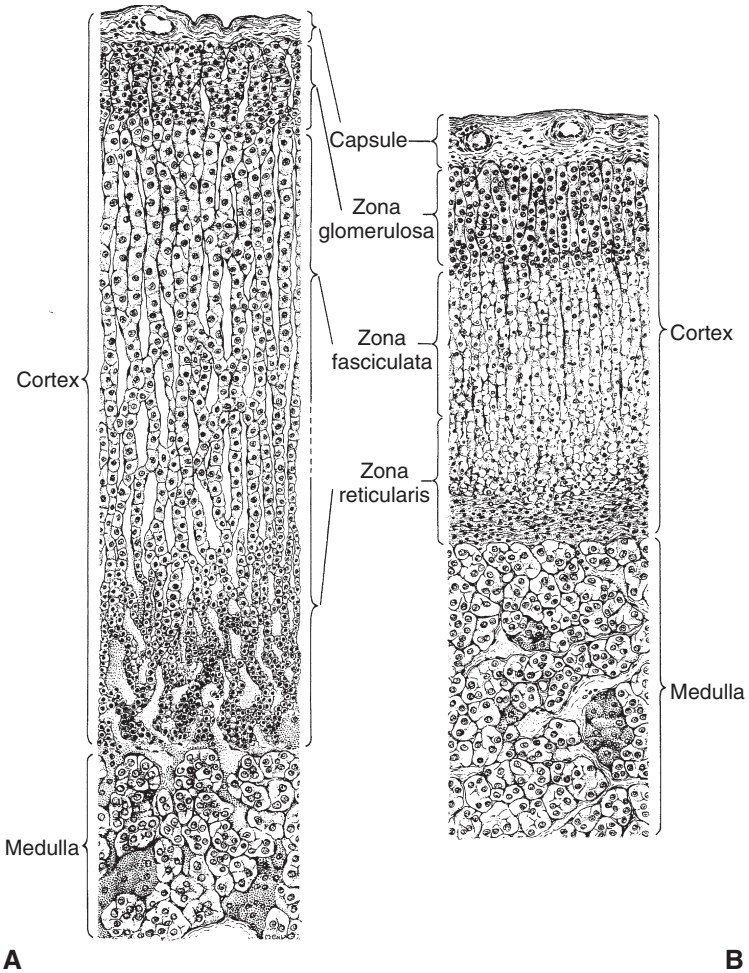
The adrenal glands are two bilaterally symmetrical endocrine organs located just anterior to the kidneys. Each gland is divided into two separate entities: a medulla and a cortex (Fig. 34.7), each of which produces different types of hormones. These adrenal tissues have different embryonic origins. The *medulla* arises from the neuroectoderm and produces amines such as norepinephrine and epinephrine. The *cortex* arises from the mesodermal coelomic epithelium and produces steroid hormones such as cortisol, corticosterone, sex steroids, and aldosterone. The utility of placing two such disparate tissues together is not apparent. The one common factor is that both sets of hormones are important for adaptation to adverse environmental conditions (i.e., stress).

Interest in the function of the adrenal cortex was heightened in the 1930s because of the research of Hans Selye. He published a series of papers on the effects of adrenalectomy and the ability of the surgically treated animal to defend itself against injury. Selye's hypothesis was termed the *general adaptation syndrome,* which he divided into three parts: the alarm reaction, the stage of resistance, and the stage of exhaustion. The critical aspect of this theory was that, in addition to specific responses to injury, animals responded in *nonspecific* ways to combat injury, and the adrenal cortex was the most important organ in leading the nonspecific response to stress. One example of the beneficial effects of glucocorticoids in a situation of injury is the mobilization of glucose, a readily usable source of energy for running away or healing injury. The adaptation of animals to stressful environments is often accompanied by enlargement of the adrenal cortex, such as in domestic chickens raised in crowded conditions and wild animals living in relatively high density.

The Adrenal Cortex

The Adrenal Cortex Has Three Zones: the Zona Glomerulosa, Which Secretes Mineralocorticoids, and the Zona Fasciculata and the Zona Reticularis, Which Secrete Glucocorticoids and Sex Steroids

The adrenal cortex is organized into three zones in mammals (see Fig. 34.7). The outer zone, the *zona glomerulosa,* is relatively narrow, and its cells are organized in a whorl-type arrangement. The middle zone, the *zona fasciculata,* is relatively wide, and its cells are organized in columns. In the cow and sheep, the zona fasciculata is further divided into inner and outer layers. The inner zone of the adrenal



• **Fig. 34.7** Depiction of comparable sections through the adrenal glands of (A), normal rats, and (B), hypophysectomized rats. Because the functional capacity of the adrenal cortex is conditioned by the release of adrenocorticotropic hormone (corticotropin), hypophysectomy results in tremendous shrinkage of the cortex. The medulla is not influenced by hypophysectomy. Both sections are drawn to scale. (From Turner CD, Bagnara JT. *General Endocrinology.* 6th ed. Philadelphia: Saunders; 1976.)

cortex, the *zona reticularis,* which is adjacent to the adrenal medulla, is intermediate in size, and cells are more randomly organized.

All the cells of the adrenal cortex have intracellular features characteristic of steroid hormone synthesis: an abundance of lipid droplets (containing cholesterol esters), mitochondria, and smooth endoplasmic reticulum. Human adrenal glands have an additional zone, the fetal zone, that is present during fetal life and for the first year of life. The fetal zone participates with the placenta in the production of estrogen during gestation. Immature mice and rabbits have an inner X zone that becomes the zona reticularis at puberty.

The adrenal cortex produces two major types of steroid hormones: the mineralocorticoids and the glucocorticoids. These hormones have distinctly different functions. The *mineralocorticoids,* produced by the zona glomerulosa, play an important role in electrolyte balance and therefore are important in the regulation of blood pressure (see later discussion). The major mineralocorticoid is *aldosterone.* The *glucocorticoids,* produced by the zona fasciculata (which accounts for the majority of glucocorticoid production) and zona reticularis, are important in the regulation of all aspects of metabolism, either directly or through an interaction with other hormones. The major glucocorticoid is *cortisol.*

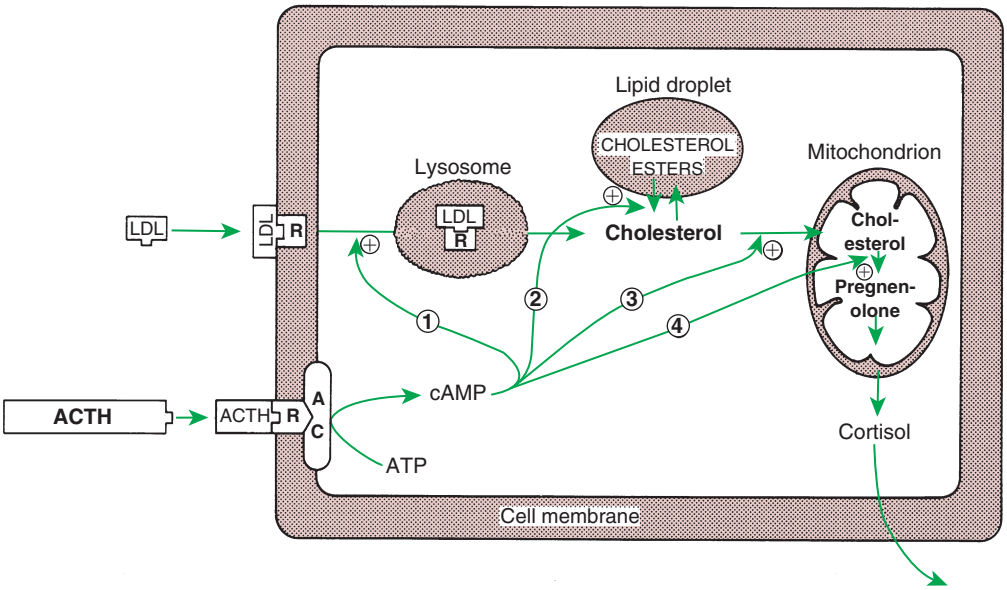
Adrenal Corticosteroids Are Synthesized From Cholesterol; the Critical Difference in the Activity of These Corticosteroids Is Related to the Hydroxyl Group on C-17 of Glucocorticoids

The synthesis of adrenal steroids involves the classic pathways for steroid biosynthesis. As indicated previously, cholesterol is the major starting material for the synthesis of steroid hormones. Cholesterol is readily available to the steroid-synthesizing cells because it is stored in large quantities in ester form within lipid droplets in these cells. One of the initial steps in steroid formation is the hydrolysis of the ester. The first step in steroid synthesis involves an enzyme that cleaves the carbon side chain from the steroid molecule, leaving a C-21 steroid known as *pregnenolone.* This step occurs within the mitochondrion (Fig. 34.8). The synthesis of all steroid hormones, regardless of their form, utilizes pregnenolone in the synthetic pathway (see Fig. 33.5).

The critical aspect of adrenal corticoid synthesis, which differentiates adrenal corticoids from the progesterone family of steroids, is a hydroxylation step at C-21 (directed by a C-21 hydroxylase). The difference between the mineralocorticoids (aldosterone) and the glucocorticoids (cortisol) is a hydroxyl group on C-17, which is part of the glucocorticoid molecule. As expected, cells of the zona fasciculata and zona reticularis have the hydroxylating enzyme for C-17 (17α-hydroxylase), whereas cells of the zona glomerulosa do not have this enzyme. Both aldosterone and cortisol have hydroxyl groups on C-11. Because of the marked difference in biological activity of the mineralocorticoids and glucocorticoids, it is useful to view the zona glomerulosa as an endocrine organ distinct from the zona fasciculata and zona reticularis.

Two intermediate compounds in the synthesis of aldosterone have significant adrenocortical activity. *11-Deoxycorticosterone* has significant mineralocorticoid activity, although it is secreted in relatively small amounts. *Corticosterone,* the immediate precursor to aldosterone, is a relatively important glucocorticoid in animals, although its potency is less than that of cortisol.

In adrenal cortical cells, biosynthetic pathways allow some synthesis of androgens and estrogens. Although the amount of sex



• **Fig. 34.8** Mechanism of action of ACTH (adrenocorticotropic hormone; corticotropin) on adrenocortical steroidogenesis. The numbers indicate the processes stimulated (indicated by *plus signs*) by ACTH as follows: *1,* stimulation of the uptake of low-density lipoproteins (*LDL*), which are further processed to free cholesterol; *2,* stimulation of the hydrolysis of stored cholesterol esters to generate free cholesterol; *3,* stimulation of the transport of cholesterol into mitochondria, where cleavage of the cholesterol side chain occurs; and *4,* promotion of the binding of cholesterol to the enzyme. *AC,* Adenyl cyclase; *ATP,* adenosine triphosphate; *cAMP,* cyclic adenosine monophosphate; *R,* receptor. (Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

steroids produced by the adrenal cortex under normal conditions is low, significant amounts can be synthesized under pathological conditions.

Adrenocortical Hormones Are Carried in Plasma in Association With Specific Binding Globulins (Corticosteroid-Binding Globulin)

Steroid hormones, as indicated previously, are lipids and depend on binding to plasma proteins for transport in the blood. A specific globulin that has a high affinity for cortisol has been identified: *corticosteroid-binding globulin,* or *transcortin.* Of the cortisol carried in plasma, 75% is bound to transcortin and 15% to albumin, leaving 10% in the unbound, or free, state. This amount of free hormone is large compared with thyroid hormones: less than 0.1%–1% of T_4 is free. The transport of aldosterone is mainly associated with albumin (50%), and only 10% is associated with transcortin, leaving a very large amount (40%) in the free state.

Changes in physiological or pathophysiological states can influence the amount of binding proteins present in plasma. Estrogen produced in increasing amounts by the fetoplacental unit during pregnancy results in an increase in hepatic synthesis of transcortin, whereas liver dysfunction can result in lower concentrations of transcortin. The large pool of hormone present in the bound state during pregnancy gives animals a good reserve from which to make appropriate adjustments in the amount of free hormone available for influencing biological activity. Because the total amount of glucocorticoid is determined in the assay of plasma concentrations, the veterinary clinician needs to be aware that total concentrations not only reflect secretion rate but also can be influenced by the amount of glucocorticoid-binding plasma proteins.

The Metabolism of Adrenocortical Hormones Involves the Reduction of Double Bonds and Conjugation of the Steroids to Glucuronides and Sulfates

The clearance half-life of cortisol is about 60 minutes, and that of aldosterone is about 20 minutes. This difference is attributable to the observed difference in protein binding of these hormones within the plasma. In general, metabolism of mineralocorticoid and glucocorticoid hormones involves the reduction of double bonds and ketone configurations, which reduces the biological activity of the molecules. The liver, an organ important for modification of these hormones, is also an important site for the conjugation of these steroids with sulfates and glucuronides. This reduces their biological potency and renders them water soluble for passage in the urine.

One of the Most Important Functions of Glucocorticoids Is Control of Metabolism, in Particular the Stimulation of Hepatic Gluconeogenesis

The mechanism of action of adrenal hormones is similar to that of other lipophilic hormones; they are able to penetrate the cell membrane and interact in the cytoplasm with specific cytosolic receptors. This complex is transferred to the nucleus, resulting in transcription of certain genes and the synthesis of specific proteins that affect the biological action of the adrenal hormones.

TABLE 34.2 Relative Glucocorticoid and Mineralocorticoid Potencies of Various Steroids

Steroid	Glucocorticoid Potency (Relative to Cortisol)	Mineralocorticoid Potency
Cortisol	1	1
Aldosterone	0.1	400
Corticosterone	0.2	2
11-Deoxycorticosterone	<0.1	20
Dexamethasone	30	2
Fludrocortisone	10	400
Prednisone	4	0.7
Triamcinolone	5	<0.1

From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.

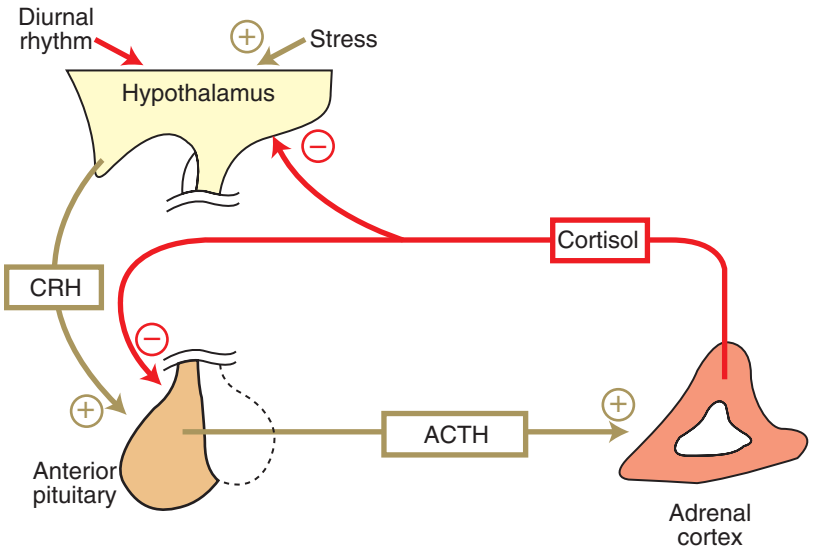
As emphasized previously, adrenocortical hormones are classified as either glucocorticoid or mineralocorticoid in their activity. Before the biological actions of each class are discussed, it is important to realize that there is overlap of activity (Table 34.2). For example, whereas cortisol is the dominant glucocorticoid hormone, it also has mineralocorticoid effects, although at a reduced potency.

The glucocorticoid hormones are important mediators of intermediary metabolism. One of the important specific effects of glucocorticoids is the stimulation of hepatic gluconeogenesis, which involves the conversion of amino acids to carbohydrates. The net result is an increase in hepatic glycogen and a tendency to increase blood glucose levels. These effects on glycogen metabolism are observed mainly in animals that have excessive glucocorticoid secretion (hyperadrenocorticism) or an insulin deficiency. The effect of glucocorticoids on carbohydrate metabolism is "permissive"; that is, their presence is required for the gluconeogenic and glycogenolytic actions of glucagon and epinephrine, respectively.

Glucocorticoids and insulin have similar effects on liver glycogen metabolism, but their effects on the peripheral use of glucose are different. Glucocorticoids inhibit glucose uptake and metabolism in the peripheral tissues, particularly in muscle and adipose cells. This effect has been termed an *antiinsulin effect.* The chronic administration of glucocorticoids can lead to the development of a syndrome called *steroid diabetes* because of the hyperglycemic effect produced at the level of the liver; use of glucose decreases in the peripheral tissues because of insulin antagonism.

Whereas the actions of glucocorticoids on fat metabolism tend to be complex, the direct effect on adipose tissue is to increase the rate of lipolysis and to redistribute fat into the liver and abdomen. This fat redistribution leads to the classic "potbelly" appearance of animals and humans with hyperadrenocorticism.

Protein synthesis is inhibited by glucocorticoids; in fact, protein catabolism is enhanced, with an accompanying release of amino acids. This process supports hepatic gluconeogenesis. Two tissues, cardiac and brain, are spared from the effect of glucocorticoids on protein catabolism. Chronic administration of glucocorticoids results

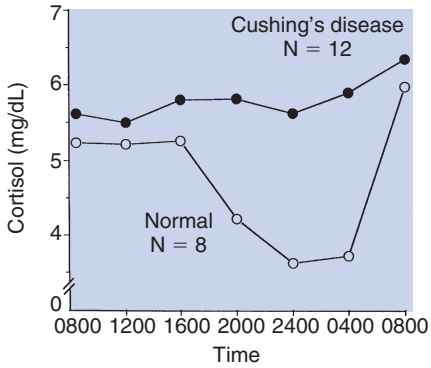


• **Fig. 34.9** Regulation of cortisol secretion by the hypothalamopituitary axis. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. *ACTH,* Adrenocorticotropic hormone (corticotropin); *CRH,* corticotropin-releasing hormone. (Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

TABLE 34.3 Glucocorticoid Effects and Target Tissues

Effect	Site of Action
Stimulates gluconeogenesis	Liver
Increases hepatic glycogen	Liver
Increases blood glucose	Liver
Facilitates lipolysis	Adipose tissue
Is catabolic (negative nitrogen balance)	Muscle, liver
Inhibits ACTH secretion	Hypothalamus, anterior pituitary gland
Facilitates water excretion	Kidney
Blocks inflammatory response	Multiple sites
Suppresses immune system	Macrophages, lymphocytes
Stimulates gastric acid secretion	Stomach

ACTH, Adrenocorticotropic hormone (corticotropin).
Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.



• **Fig. 34.10** Circadian changes in cortisol secretion in normal horses (*open circles*), in comparison with no circadian change in horses with equine Cushing's disease (*solid circles*). (From Dybdal N. *The Pathophysiology of Pituitary Pars Intermedia Dysfunction in the Horse.* Davis: University of California–Davis; 1989 [PhD thesis].)

in muscle wasting and the weakening of bone. The mobilization and incorporation of amino acids into glycogen result in an increase in urinary excretion of nitrogen and a negative nitrogen balance.

Glucocorticoids play a role in water diuresis (i.e., the enhancement of water excretion). Whereas glucocorticoids inhibit vasopressin activity in the distal tubule, the most important effect is to increase the GFR. Table 34.3 summarizes the effects of glucocorticoids.

Corticotropin (ACTH) Is the Pituitary Hormone That Regulates Glucocorticoid Synthesis by the Adrenal Cortex

The control of the secretion of the glucocorticoids by the zona fasciculata and zona reticularis is by the tropic hormone (corticotropin or adrenocorticotropic hormone [ACTH]) (Fig. 34.9). A negative-feedback system exists, whereby glucocorticoids inhibit the release of hypothalamic corticotropin–releasing hormone, which in turn results in decreased ACTH secretion by the pituitary gland. Some evidence indicates that glucocorticoids also have a negative-feedback effect at the level of the pituitary gland. The potency of a glucocorticoid in negative-feedback inhibition of ACTH is directly related to its glucocorticoid potency; for example, cortisol has more potent negative-feedback effects than corticosterone and has more potent glucocorticoid effects.

The negative-feedback control system that exists for the secretion of glucocorticoids does not result in the maintenance of uniform hormone concentrations in blood throughout the day. Sleep and activity patterns are superimposed on the negative-feedback system, so a predictable circadian rhythm occurs in which concentrations of glucocorticoids are lowest late at night and highest in the early-morning hours (Fig. 34.10).

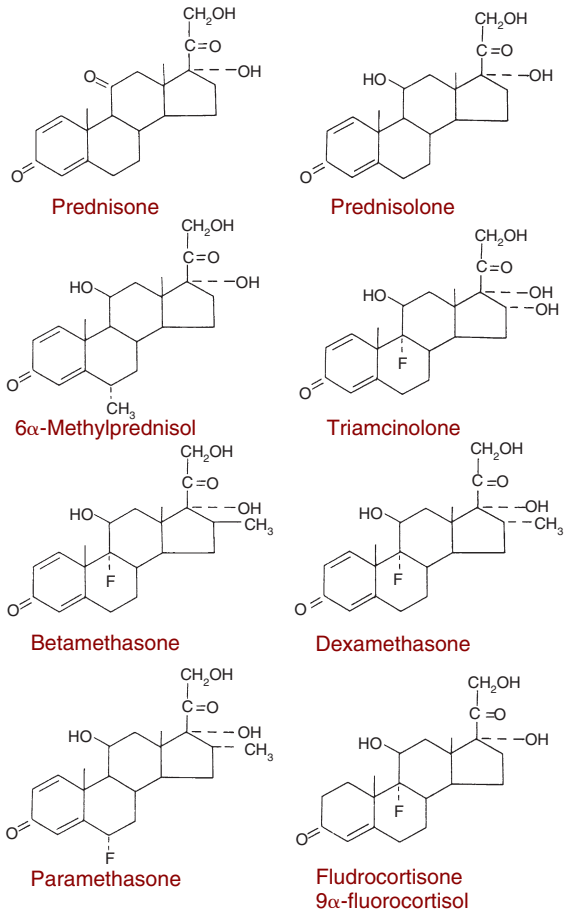
Another factor that can modify the negative-feedback control of glucocorticoids is stress. Stress can result from physical or psychological stimuli that are harmful to the individual. The effects of stress, as with the factors that influence circadian rhythms of glucocorticoid secretion, are mediated through the CNS. The glucocorticoid response to stress is immediate; concentrations of cortisol increase rapidly to reach, within minutes, values that are

severalfold higher than normal. The glucocorticoid response is proportional to the severity of the stress; that is, lower levels of stress result in less cortisol production than do higher levels of stress.

One of the Most Important Clinical Uses of Glucocorticoids Is the Suppression of the Inflammatory Response

Glucocorticoids have valuable clinical effects, particularly the inhibition of the inflammatory response, including the prevention of capillary dilation, extravasation of fluid into tissue spaces, leukocyte migration, fibrin deposition, and connective tissue synthesis. The process of inflammation is important for the walling off and destruction of systemic noxious agents, but the end response is often the replacement of functional tissue with fibrous connective tissue, with a resultant loss of function. For example, inflammatory processes in the mammary gland often result in the isolation of the injurious agent by the laying down of connective tissue as a part of the defense mechanism; however, the gland may lose much of its functional capacity as a result. Administration of glucocorticoids, in conjunction with antibiotic therapy, can help reduce the loss of functional tissue by inhibiting the development of connective tissue. Fig. 34.11 shows the chemical structures of some of the synthetic glucocorticoids used in clinical practice.

One of the ways in which glucocorticoids inhibit the inflammatory response is through the inhibition of the formation of substances that promote inflammation. Glucocorticoids inhibit the synthesis of inflammatory mediators, such as prostaglandins, thromboxanes, and leukotrienes, that arise as a result of arachidonic acid metabolism. This effect is mediated through the stabilization of lysosomal membranes and the prevention of phospholipase A_2 activation. Glucocorticoids are also used to inhibit allergic reactions. This action occurs through the inhibition of the release of certain biogenic amines, such as histamine, from the granules of mast cells.



• **Fig. 34.11** Chemical structures of some clinically useful glucocorticoid analogs. (Modified from Martin CR. *Endocrine Physiology.* New York: Oxford University Press; 1985.)

Hyperadrenocorticism

Hyperadrenocorticism (Cushing's syndrome) in the dog may be caused by a pituitary tumor, pituitary hyperplasia, adrenal tumors, adrenal hyperplasia, or nonendocrine tumors (usually of the lung), or it may be iatrogenic. Approximately 85% of dogs with hyperadrenocorticism have pituitary gland–dependent disease, whereas 15% exhibit adrenal tumors. Hyperadrenocorticism is a disease of middle-aged and older dogs (7–12 years of age). Breeds typically affected by pituitary-dependent hyperadrenocorticism include miniature poodles, dachshunds, boxers, Boston terriers, and beagles. Adrenal tumors are seen more frequently in large-breed dogs, and there is a predilection for females (3:1 ratio with males). Hyperadrenocorticism is a rare endocrine disorder of cats and is usually pituitary in origin in that species.

The most common clinical signs associated with canine hyperadrenocorticism are polydipsia, polyuria, polyphagia, heat intolerance, lethargy, abdominal enlargement or "potbelly," panting, obesity, muscle weakness, and recurrent urinary tract infections (UTIs). Dermatological manifestations of canine hyperadrenocorticism can include alopecia (especially truncal), thin skin, phlebectasias, comedones, bruising, cutaneous hyperpigmentation, calcinosis cutis, pyoderma, dermal atrophy (especially around scars), seborrhea, and secondary demodicosis. Thin skin is the hallmark of feline hyperadrenocorticism. Cats with Cushing's syndrome develop such severe thinning of the epidermis that they may incur open wounds just by grooming themselves.

Attempts to diagnose hyperadrenocorticism can be challenging. Uncommon clinical manifestations of hyperadrenocorticism in dogs can include signs such as hypertension, congestive heart failure, bronchial calcification, pulmonary thromboembolism, polyneuropathy, polymyopathy, pseudomyotonia, behavioral changes, and blindness. Evidence of increased collagenase activity caused by hypercortisolemia may result in nonhealing corneal ulceration and bilateral cranial cruciate rupture (in small dogs). Unusual reproductive signs may include testicular atrophy, prostatomegaly in castrated male dogs, clitoral hypertrophy, and perianal adenoma in females or castrated males.

Serum chemistry abnormalities associated with hypercortisolemia in dogs include increased serum activities of alkaline phosphatase and alanine aminotransferase, hypercholesterolemia, hyperglycemia, and decreased BUN. The hemogram is often characterized by evidence of erythroid regeneration (nucleated red blood cells) and a classic *stress leukogram.* Basophilia is occasionally observed. Many dogs with hyperadrenocorticism have evidence of UTI without pyuria. Proteinuria resulting from glomerulosclerosis is also common. Urine specific gravity is usually decreased and may be hyposthenuric. Thyroid status is often affected in animals with hyperadrenocorticism, as evidenced by (1) decreases in TT_4 and TT_3 caused by euthyroid sick syndrome and (2) a response to TSH stimulation that is attenuated as a result of overcrowding of pituitary thyrotrophs by adrenocorticotrophs. Overt DM may result from the insulin antagonism caused by hypercortisolemia in about 15% of dogs

with hyperadrenocorticism and 85% of cats with hyperadrenocorticism. Conversely, hyperadrenocorticism can be a cause of insulin resistance and poor glycemic control in diabetic animals.

The diagnosis of hyperadrenocorticism should be based on suggestive clinical signs and supporting minimal database abnormalities (e.g., high serum cholesterol, increased serum alkaline phosphatase activity) and confirmed by an appropriate screening test. The use of screening tests should not be based solely on laboratory abnormalities (i.e. elevated alkaline phosphatase) without suggestive clinical signs. If screening test results are inconclusive, the dog should be retested at a later date (3–6 months) rather than subjected to treatment without a definitive diagnosis.

Screening tests for hyperadrenocorticism, such as the low-dose dexamethasone suppression (LDDS) test and the ACTH stimulation test, work on the principle of suppression or stimulation of the pituitary-adrenal axis. In the case of the LDDS test, dexamethasone is administered at a low dosage to cause negative feedback to the pituitary gland. In a normal animal, this negative feedback results in a decrease in endogenous ACTH secretion and a resultant decrease in circulating cortisol concentrations. Dexamethasone is the only synthetic corticosteroid that does not cross-react with the cortisol assay. The ACTH stimulation is used to determine the extent of adrenal enlargement. Adrenal glands that are enlarged because of chronic pituitary stimulation by ACTH or that are neoplastic show an exaggerated response to exogenous ACTH.

The LDDS test has traditionally been the screening test of choice for canine hyperadrenocorticism. It is sensitive (92%–95%); only 5% to 8% of dogs with pituitary-dependent hyperadrenocorticism exhibit suppressed cortisol concentrations at 8 hours (i.e., 5%–10% false-negative results). In addition, 30% of dogs with pituitary-dependent hyperadrenocorticism exhibit suppression at 3 or 4 hours, followed by "escape" of suppression at 8 hours; this pattern is diagnostic for pituitary-dependent disease and makes further testing unnecessary. The major disadvantage of the LDDS test is the lack of specificity in dogs with nonadrenal illness. It is recommended that a dog be allowed to recover from the nonadrenal illness before being assessed for hyperadrenocorticism with the LDDS test.

The urine cortisol-to-creatinine ratio (UC:CR) is sensitive (useful for its negative predictive value, i.e., if a normal UC:CR is obtained, hyperadrenocorticism is unlikely), inexpensive, easy to perform, and useful as a screening test. Home collection (nonstressed) of urine is preferred. An at-home LDDS test using the UC:CR is used in some areas as well.

Mineralocorticoids

The mineralocorticoids, produced in the outer zone (zona glomerulosa) of the adrenal cortex, have surprisingly different functions compared with glucocorticoids. As indicated previously, electrolyte balance and blood pressure homeostasis are the principal physiological effects of mineralocorticoids (Table 34.4). These actions are carried out at the level of the distal tubules in the kidney. The effect of the mineralocorticoids is to promote retention of sodium and secretion of potassium and hydrogen. The cellular response to mineralocorticoids is to synthesize a protein that increases the permeability of the luminal cell surface to sodium influx from the renal filtrate and increases sodium/potassium-adenosinetriphosphatase (Na^+, K^+-ATPase) activity in the contraluminal cell surface, which allows movement of Na^+ out of the cell and into the interstitial tissue (Fig. 34.12).

The control of secretion of K^+ by mineralocorticoids is passive in the sense that K^+ is retained in the renal filtrate to maintain the osmolality of urine. However, evidence suggests that mineralocorticoids have an effect on Na^+ secretion independent of Na^+ retention. The secretion of K^+ continues to be influenced by mineralocorticoids after mineralocorticoid administration, whereas Na^+ retention decreases within a few days.

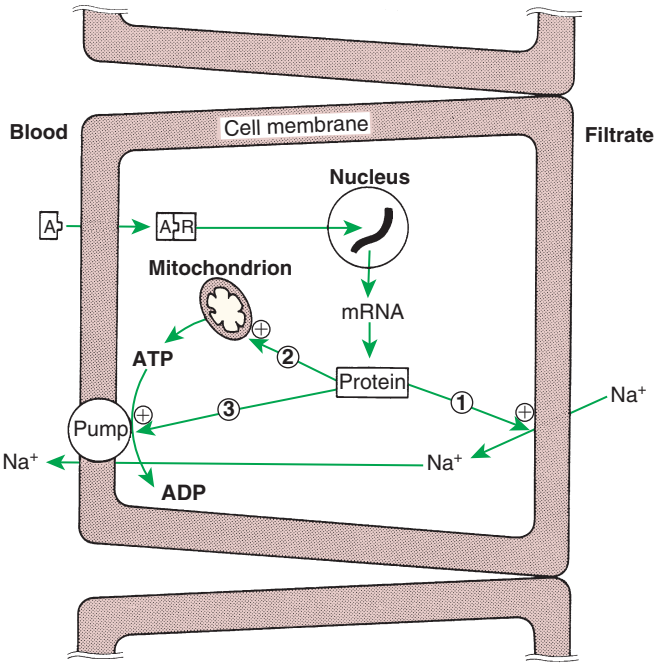
In situations of excessive mineralocorticoid production (e.g., hyperaldosteronism), the effects of increased Na^+ retention are to increase the extracellular fluid (ECF) volume and to cause hypertension; conversely, low blood pressure (hypotension) occurs as a result of inadequate secretion of mineralocorticoids. Hypersecretion of mineralocorticoids can also lead to excessive hydrogen ion (H^+) loss and metabolic alkalosis, whereas hyposecretion (e.g., Addison's disease) can result in increased retention of H^+ and metabolic acidosis.

The regulation of mineralocorticoid secretion, in contrast to glucocorticoid secretion, is not controlled by tropic hormones

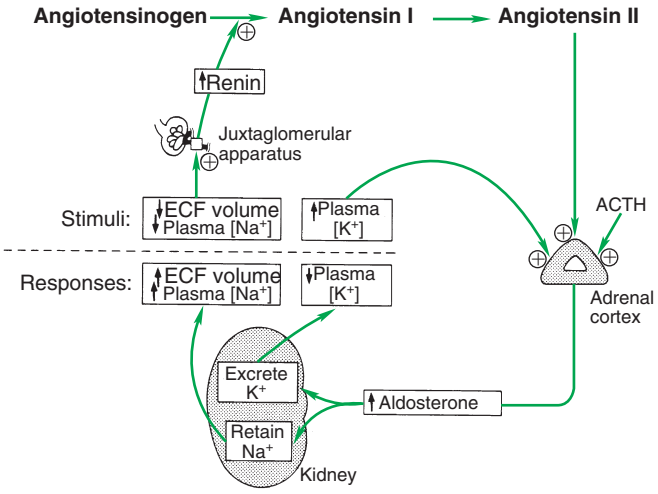
TABLE 34.4 Mineralocorticoid Effects and Target Tissues

Effect	Site of Action
Stimulates Na^+ reabsorption	Kidney, salivary glands, sweat glands
Stimulates K^+ excretion	Kidney, salivary glands, sweat glands
Stimulates H^+ excretion	Kidney

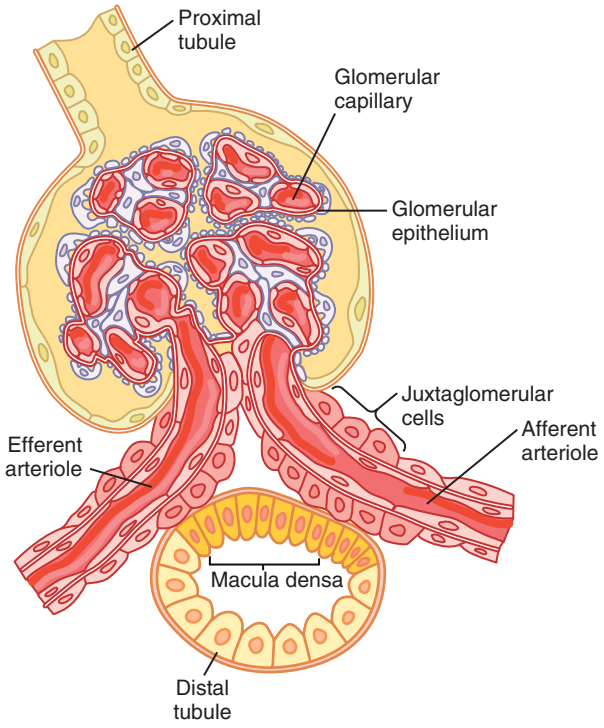
From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.



• **Fig. 34.12** Mechanisms of action of aldosterone on sodium transport in the renal tubular cell. The *numbered arrows* indicate the three putative sites of action of aldosterone: *1,* increasing the permeability of the luminal membrane to sodium; *2,* increasing mitochondrial adenosine triphosphate (*ATP*) production; *3,* increasing Na^+,K^+-ATPase activity in the contraluminal membrane. *Plus signs* indicate stimulation. *A,* Aldosterone; *ADP,* adenosine diphosphate; *mRNA,* messenger ribonucleic acid; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.13** Regulation of aldosterone secretion by the zona glomerulosa of the adrenal cortex. *Plus signs* indicate stimulation. *ACTH,* Adrenocorticotropic hormone (corticotropin); *ECF,* extracellular fluid. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.14** Diagrammatic representation of juxtaglomerular apparatus. (Modified from Hall JE. *Guyton and Hall Textbook of Medical Physiology.* 12th ed. Philadelphia: Saunders Elsevier; 2011.)

from the pituitary gland (Fig. 34.13). In the case of mineralocorticoids, the main controlling factors are produced in the target organ, the kidney. Cells in the juxtaglomerular apparatus of the kidney produce an enzyme, *renin,* in response to decreases in blood pressure. Renin acts on *angiotensinogen,* an α_2 globulin produced by the liver and present in the circulation, and this results in the production of *angiotensin I,* a decapeptide. Angiotensin I is further hydrolyzed to *angiotensin II,* an octapeptide, by angiotensin-converting enzyme. Angiotensin II stimulates the zona glomerulosa to produce mineralocorticoids such as aldosterone. Angiotensin II also increases peripheral resistance of the blood vascular system by causing vasoconstriction of smooth muscle of the blood vessels. Angiotensin II, if present on a long-term basis, also increases the size of the zona glomerulosa.

Evidence indicates that cells of the macula densa, groups of specialized cells located at the origin of the kidney's distal tubule (Fig. 34.14), exert control on the *renin-angiotensin system.* This is done through the sensing of changes in Na^+ concentrations in tissue fluids; increased Na^+ results in decreased renin release, and decreased Na^+ results in increased renin release. In either case, the change produced tends to restore the mineralocorticoid concentrations to normal. In addition to the effect of sodium, the macula densa may control changes in the renin-angiotensin system through the sensing of changes in chloride ion (Cl^-) concentrations in tissue fluids.

Another major regulatory factor in the control of mineralocorticoid secretion is the blood potassium concentration. An increase in K^+ concentration stimulates the zona glomerulosa to secrete mineralocorticoids, whereas a decline in K^+ has the opposite effect. This stimulation is independent of the renin-angiotensin system.

It has been thought that ACTH has minimal effect on control of the zona glomerulosa, because experimental studies showed that hypophysectomy has little effect on the zona glomerulosa. More studies have shown that cells of the zona glomerulosa will respond to ACTH, which may play some role, although minor, in the physiological control of mineralocorticoid secretion.

In contrast to the sodium-conserving effect of mineralocorticoids, the 28–amino acid *atrial natriuretic peptide* (ANP) reduces Na^+ retention by the kidneys. ANP also causes peripheral vasodilation and thus a lowering of blood pressure. ANP may inhibit the production of mineralocorticoids and renin as well. ANP is produced by cells of the cardiac atria, but it is also produced in other sites, including the brain.

Hypoadrenocorticism

Hypoadrenocorticism, caused by lack of mineralocorticoids and glucocorticoids, is most often diagnosed in young female dogs and usually has an immune-mediated etiology. Certain breeds, such as Leonbergers, standard poodles, and Portuguese water dogs, are at increased risk for the disease; however, hypoadrenocorticism may be diagnosed in any breed. Addison's disease is a rare condition of young to middle-aged cats. Historical findings compatible with hypoadrenocorticism include intermittent vomiting, diarrhea, weight loss, lethargy, anorexia, and weakness. These symptoms often resolve with fluid therapy and corticosteroid treatment. Physical examination of animals in an acute hypoadrenal crisis reveals weak pulse, bradycardia, prolonged capillary refill time, severe mental slowness, and profound muscle weakness. Clinical features of hypoadrenocorticism that should raise suspicion include a normal or slow heart rate in the presence of circulatory shock and the "waxing and waning" course of disease before collapse. Cats with hypoadrenocorticism may exhibit similar clinical signs; however, bradycardia is much less common.

Electrolyte abnormalities consisting of severe hyponatremia and hypochloremia associated with hyperkalemia are the hallmarks of hypoadrenocorticism. Azotemia and hyperphosphatemia can also accompany primary hypoadrenocorticism, which makes it difficult to differentiate it from acute renal failure. Azotemia may be prerenal as a result of dehydration and hypovolemia, or increase in BUN may be caused by GI hemorrhage. Hematological abnormalities

consist of eosinophilia and lymphocytosis, or eosinophil and lymphocyte counts may be normal in the presence of severe metabolic stress. The anemia of hypoadrenocorticism has classically been attributed to lack of glucocorticoid effects on the bone marrow. However, hemorrhagic gastroenteritis may contribute significantly to the anemia as well. Although hypoglycemia is more common with secondary or atypical hypoadrenocorticism, it is rarely seen with typical hypoadrenocorticism.

Urine specific gravity is frequently low, attributable to medullary washout (inadequate medullary gradient caused by sodium depletion) and decreased medullary blood flow. Dilute urine in the presence of azotemia and hyperkalemia may easily be mistaken for acute renal failure. Hormonal assays are necessary to confirm the presence or absence of adrenal disease and to differentiate between hypoadrenocorticism and renal failure.

Diagnosis of primary hypoadrenocorticism is based on clinical signs, classic electrolyte imbalances, and confirmation with an ACTH response test. The baseline cortisol sample should be collected with the initial blood work, and synthetic ACTH should be administered intravenously during the initial fluid therapy. A 1-hour, post-ACTH sample may then be drawn, and glucocorticoids may be administered after the 1-hour sample is taken. Intramuscular injection of ACTH (gel or synthetic) may not be absorbed in animals in circulatory shock; therefore, intravenous administration of synthetic ACTH is preferred. If glucocorticoids must be administered before the measurement of cortisol, dexamethasone sodium phosphate is preferred because it does not cross-react with the cortisol assay. Plasma ACTH may be measured to determine whether the hypoadrenocorticism is primary or secondary.

Dogs and cats with primary hypoadrenocorticism exhibit a subnormal response to ACTH administration. Both the baseline and the post-ACTH cortisol concentrations are usually low or undetectable. Plasma ACTH concentrations are dramatically increased in animals with primary hypoadrenocorticism as a result of loss of negative feedback to the pituitary gland, caused by decreased serum cortisol concentrations. In the rare cases of secondary hypoadrenocorticism, caused by a pituitary deficiency of ACTH, the plasma ACTH concentrations are typically decreased (<20 pg/mL). The response to exogenous ACTH is diminished but not as dramatically as for primary hypoadrenocorticism.

Hyperaldosteronism (Conn's Syndrome)

Hyperaldosteronism is caused by either an adrenal tumor hypersecreting aldosterone or bilateral adrenal hyperplasia; the condition is usually diagnosed in older cats and occasionally dogs. Oversecretion of aldosterone results in the classic electrolyte changes of hypokalemia (increased plasma K^+), hypernatremia (elevated plasma Na^+), and metabolic alkalosis (increased pH; opposite of Addison's disease). Tumors of the adrenal gland are usually benign. Clinical signs consist of muscle weakness and cervical ventroflexion in cats with aldosterone-secreting adrenal tumors. Less commonly, hypertension and blindness occur. In cats with bilateral adrenal hyperplasia, hypertension, blindness, and renal failure are more common than signs of hypokalemia (i.e., muscle weakness). Laboratory signs include hypokalemia, elevated creatine kinase (CK), and metabolic alkalosis in all adrenal tumors and less commonly with bilateral adrenal hyperplasia. Hypernatremia is seen in less than 30% of cases. Diagnosis is achieved by documentation of increased plasma aldosterone (N = 150–430 pmol/L), low to undetectable plasma renin, and/or increased plasma aldosterone concentration-to-plasma renin activity ratios (PAC:PRA; normal = 0.3–3.8), and visualization of bilateral adrenal hyperplasia or unilateral enlargement of the adrenal gland on ultrasound. Treatment of hyperaldosteronism consists of potassium supplementation (oral and intravenous), spironolactone, and calcium channel blockers for hypertension. Surgical removal of the adrenal tumor is usually curative.

The Adrenal Medulla

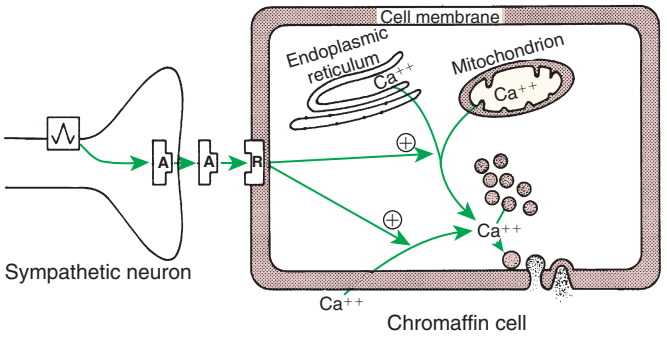
The adrenal medulla, as its name indicates, occupies the central portion of the adrenal gland (see Fig. 34.7). A stimulatory effect of adrenal medullary extracts on cardiac activity was first recognized by Oliver and Schafer in 1894. Thereafter, the main hormone of the adrenal medulla, epinephrine, became the first hormone to be isolated (by Abel in 1898), crystallized (by Takamine and Aldrich in 1901), and synthesized (by Stolz in 1904). Theories on the importance of the adrenal medulla include that of Cannon, who in 1932 proposed the "fight or flight" hypothesis, in which the adrenal medulla is activated to aid in combating situations of extreme stress. Others advocated the "tonus" theory, which stated that cells of the adrenal medulla are constantly in a state of readiness. In fact, the adrenal medulla has a constant output of catecholamines that can be accentuated dramatically if the need arises.

It was recognized early in this research that cells of the adrenal medulla are the equivalent of postganglionic cells of the sympathetic nervous system. Therefore it was assumed that *epinephrine* was the mediator of postganglionic activity of the sympathetic nervous system. It was later recognized that another catecholamine, *norepinephrine,* is the neurotransmitter of the sympathetic nervous system. Both epinephrine and norepinephrine are released when preganglionic nerve fibers to the adrenal medulla are stimulated; in fact, most of the norepinephrine found in plasma originates from the adrenal medulla. However, epinephrine is the major catecholamine secreted by the adrenal medulla of most mammals. Exceptions to this generalization include the dominance of norepinephrine over epinephrine in whales and chickens and in the fetal tissues of all species.

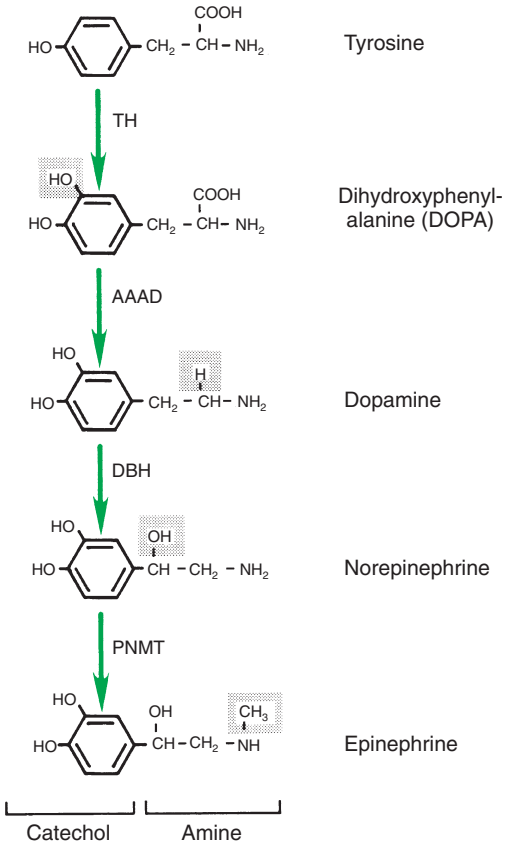
The Synthesis of Catecholamines Is From Tyrosine; the Main Catecholamine Synthesized by the Adrenal Medulla Is Epinephrine

The cells of the adrenal medulla that synthesize catecholamines are classified as *chromaffin cells.* This classification is based on the histochemical reaction of the cells when exposed to potassium dichromate, that is, a darkening of the cells as a result of the formation of colored pigments in conjunction with the oxidation of the catecholamines. The cells that produce epinephrine are different from those that synthesize norepinephrine; accordingly, the type of chromaffin granule present is different for each cell type. In cattle, the epinephrine-secreting cells tend to be on the outer edge of the medulla. *Acetylcholine* release from the preganglionic nerve fibers initiates the synthesis of the catecholamines by the medullary cells (Fig. 34.15). Acetylcholine also stimulates the release of catecholamines from chromaffin granules, a phenomenon called *stimulus-secretion coupling.*

The synthesis of the catecholamines begins with either of the amino acids phenylalanine or *tyrosine.* However, tyrosine is a naturally occurring amino acid, and most synthesis of catecholamines begins with it (Fig. 34.16). The initial step in the biosynthetic pathway begins with the conversion of tyrosine to 3,4-dihydroxyphenylalanine, or *DOPA.* Tyrosine hydroxylase, the



• **Fig. 34.15** Stimulus-secretion coupling in the adrenal chromaffin cell. Note that cytosolic calcium may be derived from intracellular or extracellular sources. *Circled plus signs* indicate stimulation. *A,* Acetylcholine; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.16** Pathway of catecholamine synthesis in the adrenal medulla. *Shaded areas* denote the structural changes occurring at each step. *AAAD,* Aromatic-L-amino acid decarboxylase; *DBH,* dopamine-β-hydroxylase; *PNMT,* phenylethanolamine-*N*-methyltransferase; *TH,* tyrosine hydroxylase. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

enzyme responsible for the conversion of tyrosine, is the rate-limiting enzyme in the formation of catecholamines. The end products of tyrosine metabolism, including DOPA, dopamine, norepinephrine, and epinephrine, inhibit the activity of tyrosine hydroxylase. DOPA is converted to dopamine through the enzymatic activity of aromatic-L-amino acid decarboxylase (DOPA decarboxylase). To this point, the biochemical transformations have occurred in the cytosol. The conversion of dopamine to norepinephrine occurs within the chromaffin granule because the key enzyme, dopamine-β-hydroxylase, is localized within the granule (Fig. 34.17).

If the cell secretes norepinephrine, the biochemical pathway is ended, and the hormone remains in the norepinephrine granule, ready for secretion. If the cell secretes epinephrine, norepinephrine moves back into the cytosol, where it is converted to epinephrine through the activity of phenylethanolamine-*N*-methyltransferase (PMNT). Epinephrine then moves into an epinephrine granule for storage before its release. The metabolism of catecholamines is rapid (2 minutes for norepinephrine, less for epinephrine) and is accomplished mainly by the liver and kidneys.

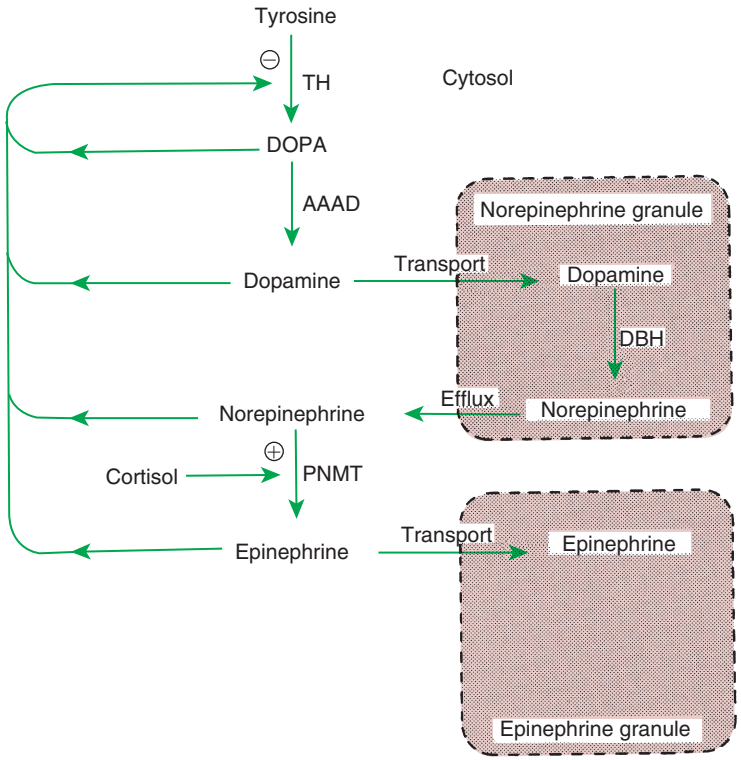
The importance of the anatomical association of the adrenal cortex and medulla may be related to the fact that cortisol is important for the activity of the enzyme PMNT. The chromaffin cells are located close to the venous sinuses that drain the adrenal cortex and therefore are exposed to venous effluent that contains high concentrations of cortisol.

The Primary Actions of Catecholamines Are on Metabolism, Especially Effects That Increase the Concentration of Glucose

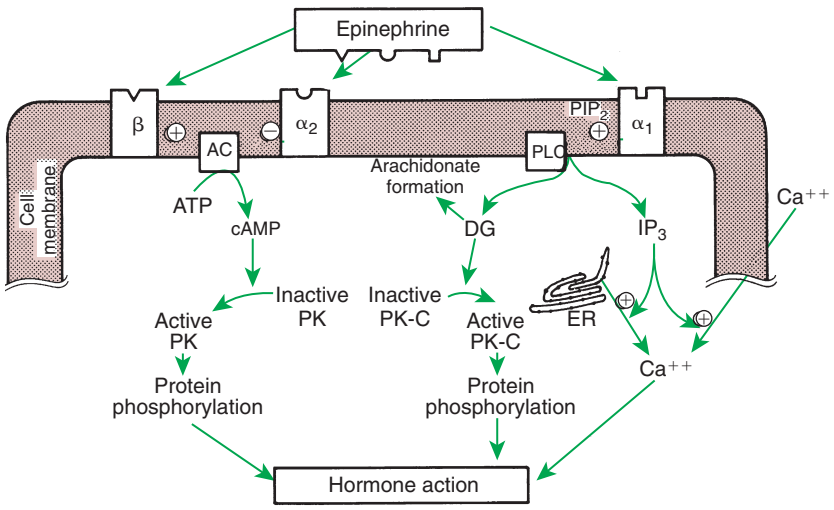
The actions of the catecholamines involve the regulation of intermediary metabolism, as well as responses that allow animals to adjust to situations involving acute stress. The actions of catecholamines are mediated through adrenergic receptors located on target tissues (Fig. 34.18). There are two major types of adrenergic receptors, alpha (α) and beta (β), which are subdivided into α_1, α_2, β_1, and β_2. The α-adrenergic receptors control catecholamine release from sympathetic nerve endings, with α_1 affecting postsynaptic nerve endings and α_2 affecting presynaptic terminals. The β_1 receptors affect mainly the heart, and β_2 receptors affect smooth muscle contraction and intermediary metabolism. Whereas all adrenergic receptors are responsive to both epinephrine and norepinephrine, the responses to the two catecholamines are different. In addition, the receptor types on various tissues vary in number, which, together with the different responses of adrenergic receptors on tissues, results in variable adrenergic responses being produced by a particular catecholamine.

The metabolic effects of catecholamines are mediated mainly by β_2 receptors. Because epinephrine is 10 times more potent than norepinephrine with β_2 receptors, epinephrine plays a much more important role in the control of intermediary metabolism than norepinephrine. The effects of epinephrine on glucose metabolism are similar to those of glucagon and opposite to those of insulin. Epinephrine increases blood glucose concentrations with the effect mainly in the liver; that is, epinephrine promotes both hepatic glycogenolysis and gluconeogenesis. Epinephrine also stimulates glycogenolysis in skeletal muscle, which in this situation is in contrast with the action of glucagon. Because glucose-6-phosphatase is not present in skeletal muscle, lactate is produced instead of glucose; the liver takes up lactate and converts it to glucose. Additional effects on glucose metabolism include the inhibition of insulin secretion (through α receptors) and stimulation of glucagon secretion by the pancreas; both actions increase blood glucose concentrations.

Epinephrine promotes lipolysis through interaction with two receptors on adipose cells. Activation of a lipase enzyme results in an increase in free fatty acids in the blood. Glucocorticoids potentiate the effect of epinephrine on lipolysis during the stress response.



• **Fig. 34.17** Regulation of catecholamine biosynthesis in the adrenal medulla. *Plus sign* indicates stimulation; minus sign indicates inhibition. *AAAD,* Aromatic-L-amino acid decarboxylase; *DBH,* dopamine-β-hydroxylase; *DOPA,* dihydroxyphenylalanine; *PNMT,* phenylethanolamine-*N*-methyltransferase; *TH,* tyrosine hydroxylase. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.18** Mechanisms of action of epinephrine in target cells mediated by β-, α_2-, and α_1-adrenergic receptors. *Plus signs* indicate stimulation; *minus sign* indicates inhibition. *AC,* Adenyl cyclase; *ATP,* adenosine triphosphate; *cAMP,* cyclic adenosine monophosphate; *DG,* diacylglycerol; *ER,* endoplasmic reticulum; *IP_3,* inositol 1,4,5-triphosphate; *PIP_2,* phosphatidylinositol 4,5-bisphosphate; *PK,* protein kinase; *PK-C,* protein kinase C; *PLC,* phospholipase C. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

Catecholamines stimulate cardiac function. Both epinephrine and norepinephrine interact with β_1 receptors to increase both the force of contraction and the heart rate, the latter resulting from the promotion of a shorter period of diastolic depolarization. Whereas both catecholamines promote arteriolar constriction through interaction with α receptors, epinephrine, through its high affinity for β_2 receptors, causes the dilation of blood vessels both in the heart and in skeletal muscle. The end result is that total peripheral resistance is decreased by the action of epinephrine with a concomitant decline in diastolic pressure; however, blood

pressure is minimally changed, and cardiac output increases because of the increase in heart rate. The action of epinephrine to increase cardiac output is an obvious beneficial effect in situations that are described as "fight or flight."

Catecholamines affect smooth muscle. Epinephrine causes relaxation of bronchial smooth muscle, particularly when the muscle is in a contracted state. Because the action is mediated through β_2 receptors, norepinephrine has little effect on bronchial smooth muscle. Epinephrine causes relaxation of the smooth muscle of the GI tract through interaction with β_2 receptors. Catecholamine stimulation of β-adrenergic receptors results in contraction of uterine smooth muscle, and stimulation of β_2 receptors results in relaxation. Because of its dominant effect on β_2 receptors, epinephrine causes relaxation of the uterus, whereas both epinephrine and norepinephrine interact with α receptors to cause contraction.

The effects of the catecholamines on bladder smooth muscle depend on different locations of α and β receptors; α-adrenergic receptors are located within the neck of the bladder, and β-adrenergic receptors are located within the body of the bladder. Epinephrine relaxes the body and contracts the neck of the bladder; norepinephrine contracts the neck of the bladder. The net effect is retention of urine.

Although the parasympathetic nervous system is the principal system involved in penile erection, the sympathetic nervous system may also play a role. Epinephrine promotes erection through vasodilation of the vasculature mediated by β receptors. Higher concentrations of epinephrine (and norepinephrine) can cause ejaculation through α-receptor interaction and vasoconstriction.

In the eye, epinephrine causes relaxation of the lens through stimulation of β receptors on the ciliary muscles. It causes dilation of the pupil through stimulation of α receptors, with resultant contraction of the radial muscle of the iris.

The effects of epinephrine on the CNS are excitatory. Drugs that affect the CNS probably do so by modulation of catecholamine concentrations, whereby sedation is associated with lower concentrations of epinephrine. Other effects of catecholamine include the promotion of sweating and piloerection. Epinephrine also increases renin production by the renal juxtaglomerular cells. Table 34.5 summarizes the effects of catecholamines.

The Main Factors That Stimulate Catecholamine Secretion Are Hypoglycemia and Conditions That Produce Stress

Any factor that increases sympathetic nervous system stimulation of the adrenal medulla results in the immediate secretion of catecholamines. The main physiological factor that influences catecholamine secretion is hypoglycemia. In this situation, epinephrine secretion is stimulated by decreases in blood glucose concentrations that are within normal physiological limits. In contrast, other parts of the sympathetic nervous system are depressed by decreases in blood glucose levels. Factors that elicit a massive release of catecholamines are categorized as "stressful," particularly those that are acute. Catecholamines are especially important for the maintenance of blood pressure in conjunction with severe blood loss; decreased blood pressure stimulates epinephrine secretion. Catecholamines are also important for adaptation to cold exposure in terms of increased heat production; decreased temperature increases epinephrine secretion. The response to acute stress can be particularly marked, because each preganglionic sympathetic neuron that supplies the adrenal medulla affects a number of chromaffin cells and the signal is greatly amplified.

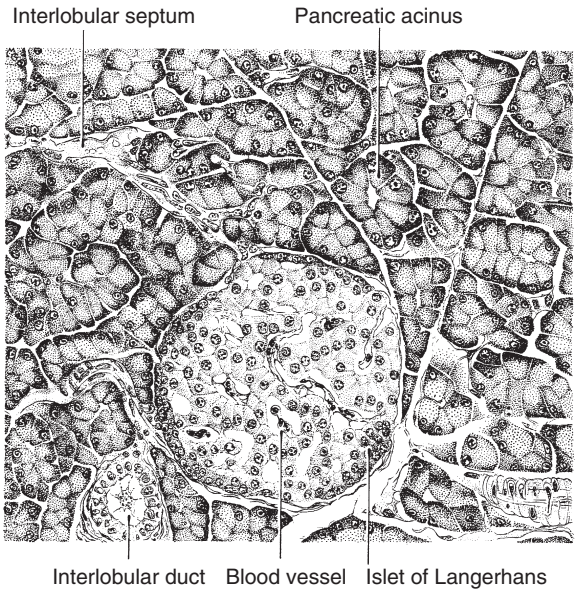
TABLE 34.5 Responses of Target Tissues to Catecholamines

Target Tissue	Receptor Type	Responses
Liver	β_2	Glycogenolysis, lipolysis, gluconeogenesis
Adipose tissue	β_2	Lipolysis
Skeletal muscle	β_2	Glycogenolysis
Pancreas	α_2	Decreased insulin secretion
	β_2	Increased insulin secretion
Cardiovascular system	β_1	Increased heart rate, increased contractility, increased conduction velocity
	α_2	Vasoconstriction
	β_2	Vasodilation in skeletal muscle arterioles, coronary arteries, and all veins
Bronchial muscle	β_2	Relaxation
Gastrointestinal tract	β_2	Decreased contractility
Urinary bladder	α_2	Sphincter contraction
	β_2	Detrusor relaxation
Uterus	α_2	Contraction
	β_2	Relaxation
Male sex organs	α_2	Ejaculation, detumescence
	β_2	Erection?
Eye	α_1	Radial muscle contraction
	β_2	Ciliary muscle relaxation
Central nervous system	α_2	Stimulation
Skin	α_2	Piloerection, sweat production
Renin secretion	β_1	Stimulation

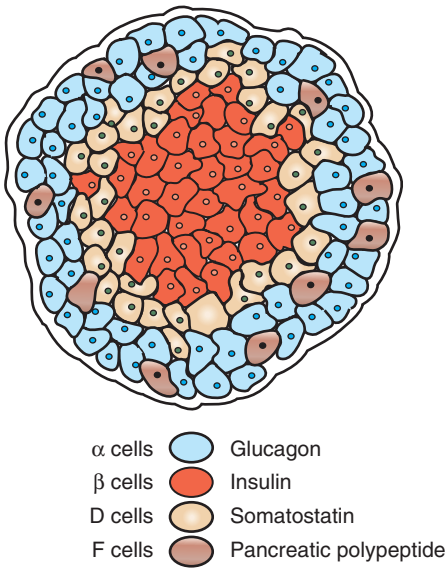
From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.

Hormones of the Pancreas

The pancreas has important endocrine and nonendocrine functions. The nonendocrine functions result from activity of the exocrine part of the pancreas and are involved in GI function. The endocrine portion of the pancreas is organized as discrete islets (islets of Langerhans) that contain four cell types, each of which produces a different hormone (Fig. 34.19). The most numerous of the islet cells are β cells, which produce insulin; α cells produce glucagon; D cells produce somatostatin; and F or PP cells produce pancreatic polypeptide (Fig. 34.20). Although these hormones have different functions, all are involved in the control of metabolism and, more specifically, in glucose homeostasis.



• **Fig. 34.19** Depiction of section through the pancreas of the rat. The islet of Langerhans is a gland of internal secretion, whereas the surrounding acinar tissue forms an exocrine gland. (From Turner CD, Bagnara JT. *General Endocrinology.* 6th ed. Philadelphia: Saunders; 1976.)



• **Fig. 34.20** Depiction of the pancreatic islet. (From McDonald LE. *Veterinary Endocrinology and Reproduction.* 4th ed. Philadelphia: Lea & Febiger; 1989.)

Insulin

The first studies that associated the pancreas with carbohydrate metabolism were done by von Mering and Minkowski in 1889 when they showed that pancreatectomy of dogs resulted in signs similar to those characteristic of DM. Later, Banting and Best were able to show that injection of pancreatic extracts could alleviate the signs of DM in dogs and humans. Able was the first to crystallize insulin, and its structure was elucidated by Sanger in 1960.

Insulin is a protein consisting of two chains, designated A (21 amino acids) and B (30 amino acids), that are connected by two disulfide bridges. The monomer form of the hormone is thought to be the active form; insulin also exists in dimer and hexamer forms, the latter complexed with two zinc molecules. Although there are some differences in amino acid composition among species, the differences are small; for example, cattle, sheep, horses, dogs, and whales differ only in positions 8, 9, and 10 of the A chain. As a result, the biological activities of insulin are not highly species-specific. Of the domestic species, feline insulin is most similar to bovine insulin, and canine insulin is identical to porcine insulin in its amino acid structure.

The Synthesis of Insulin Is Biphasic; an Acute Phase Involves the Release of Preformed Insulin, and a Chronic Phase Involves the Synthesis of Protein

The synthesis of insulin, similar to that of other peptide hormones, begins with the formation of a linear polypeptide preproinsulin within the rough endoplasmic reticulum. A small peptide fragment is removed to form proinsulin. Proinsulin is coiled, and the end fragments are joined by disulfide bonds. Proinsulin is transferred to the Golgi apparatus, where it is further processed and packaged into granules that contain both insulin and the connecting or C peptide (33 amino acids in length).

The secretion of insulin follows biphasic kinetics in response to appropriate stimuli (Fig. 34.21). The initial, acute release of insulin involves the exocytosis of preformed insulin from secretion granules. After the acute phase, a chronic phase of secretion occurs that involves the synthesis of protein and thus probably the synthesis of insulin.

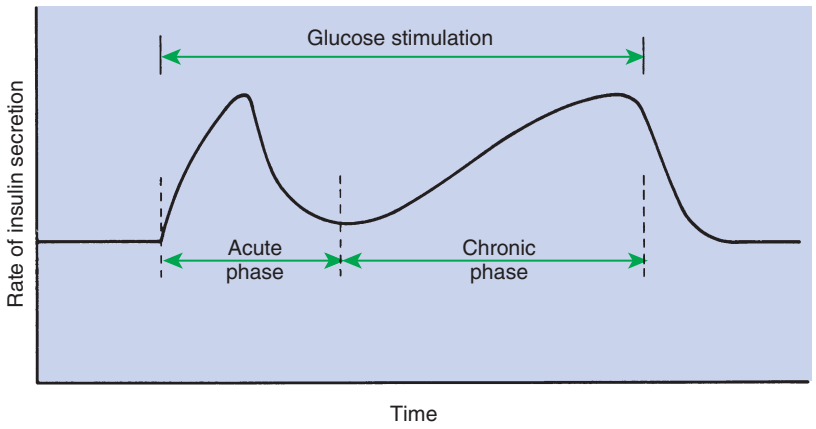
The Metabolism of Insulin Involves Splitting the A and B Chains and Reducing the Chains to Amino Acids and Peptides

Insulin is metabolized mainly by the liver and kidneys. Enzymes that are present reduce the disulfide bonds that link the A and B chains, and the chains are then subjected to protease activity, which reduces them to peptides and amino acids. The half-life of insulin is about 10 minutes.

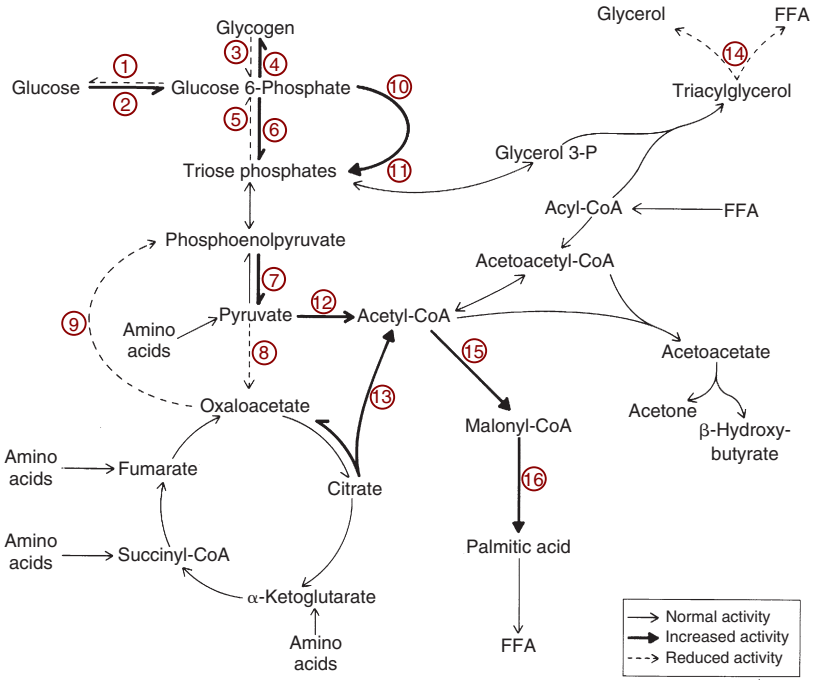
The Main Metabolic Functions of Insulin Are Anabolic

Insulin acts at a number of sites within the metabolic pathways of carbohydrates, fats, and proteins (Fig. 34.22). The liver is an especially important target organ, in part because the pancreatic venous effluent passes directly to the liver. The net effect of the actions of insulin is to lower blood concentrations of glucose, fatty acids, and amino acids, and to promote intracellular conversion of these compounds to their storage forms: glycogen, triglycerides, and protein, respectively (Table 34.6). Glucose does not readily penetrate cell membranes except in a few tissues, such as brain, liver, and red and white blood cells, all of which must have continual access to glucose. The presence of insulin is critical to the movement of glucose through the plasma membrane into the cell.

Insulin has profound effects on carbohydrate metabolism. Insulin facilitates the use of glucose, namely, glycolysis, which involves the oxidation of glucose to pyruvate and lactate through the induction of enzymes, such as glucokinase, phosphofructokinase, and pyruvate kinase. Insulin promotes glycogen production in the liver, in adipose tissue, and in skeletal muscle by increasing glycogen synthetase activity with a concomitant decrease in glycogen phosphorylase activity. Gluconeogenesis is decreased by insulin because of the promotion of protein synthesis in peripheral tissues, thereby decreasing the amount of amino acids available for gluconeogenesis. In addition, insulin decreases the activities of hepatic enzymes (fructose 1,6-bisphosphate aldolase, pyruvate carboxylase,



• **Fig. 34.21** Kinetics of insulin secretion by the β cell in response to a continued glucose stimulus. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.22** Metabolic pathways affected by insulin. The numbers correspond to each of the following enzymes: *1,* glucose-6-phosphatase; *2,* glucokinase; *3,* phosphorylase; *4,* glycogen synthase; *5,* fructose-1,6-bisphosphate aldolase; *6,* 6-phosphofructokinase; *7,* pyruvate kinase; *8,* pyruvate carboxylase; *9,* phosphoenolpyruvate carboxykinase; *10,* glucose-6-phosphate-dehydrogenase; *11,* 6-phosphogluconate dehydrogenase; *12,* pyruvate dehydrogenase; *13,* adenosine triphosphate (ATP)–citrate lyase; *14,* hormone-sensitive lipase; *15,* acetyl–coenzyme A (CoA) carboxylase; *16,* fatty-acid synthase. *FFA,* Free fatty acid. (Modified from Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

phosphoenolpyruvate carboxylase, and glucose-6-phosphatase) involved in the conversion of amino acids to glucose.

In adipose tissue, insulin promotes the synthesis of triglycerides. Insulin facilitates the intracellular use of glucose, which results in increased levels of pyruvate, a precursor of acetyl-coenzyme A (acetyl-CoA, in turn, a precursor of fatty acids) and increased glycerol-3-phosphate for the esterification of fatty acids. Insulin activates the enzymes pyruvate dehydrogenase and acetyl-CoA carboxylase, which promote the synthesis of fatty acids from acetyl-CoA. Insulin also increases the activity of lipoprotein lipase located in the endothelium of capillaries of extrahepatic tissues, which promotes the movement of fatty acids into adipose tissue. Finally, insulin decreases lipolysis in adipose tissue.

With protein metabolism, insulin promotes uptake of amino acids by most tissues, including skeletal muscle, but not liver. Insulin promotes protein synthesis and inhibits protein degradation. Therefore insulin promotes the maintenance of a positive nitrogen balance. With insulin deficiency, protein catabolism increases with increased amounts of amino acids available for hepatic gluconeogenesis and a resultant increase in the blood glucose concentrations.

The most important factor in the control of insulin secretion is the concentration of blood glucose. Increased concentrations of

blood glucose initiate the synthesis and release of insulin by the β cells of the pancreatic islets (Fig. 34.23). Two theories explain the mechanism of cellular induction of insulin synthesis and release. In the first, glucose interacts with a membrane receptor protein that directs intracellular events toward the synthesis and release of insulin. In the second theory, insulin regulation occurs at the intracellular level, whereby the metabolism of glucose produces the signal for insulin synthesis and release. Glucose control of insulin secretion is a positive-feedback system in which increased concentrations of glucose lead to increased concentrations of insulin.

Because the oral administration of glucose produces a greater insulin response than systemic administration, factors from the intestinal tract were thought to affect insulin secretion. It is now known that a number of GI hormones stimulate insulin secretion, including gastrin, cholecystokinin, secretin, and gastric inhibitory peptide. The presence of amino acids and fatty acids in the intestinal tract also stimulates the release of insulin, although with less potency than that of glucose (Box 34.1).

Hormones other than those from the GI tract are important for the control of insulin secretion. Glucagon from the α cells of the pancreas has a direct stimulatory effect on the β cells to secrete insulin. Conversely, somatostatin inhibits the secretion of insulin. Both hormones work through the adenyl cyclase system, glucagon being stimulatory and somatostatin being inhibitory. Catecholamines tend to decrease insulin secretion through an interaction with the α-adrenergic receptors on the β cells. Epinephrine is the main circulating catecholamine that affects insulin secretion, but

TABLE 34.6 Sites of Action and Effects of Insulin on Carbohydrate, Lipid, and Protein Metabolism

Process Affected	SITE OF ACTION		
	Liver	Muscle	Adipose
Carbohydrate Metabolism			
↑Glucose transport		X	X
↑Glycogen synthesis	X	X	X
↓Glycogenolysis	X	X	X
↓Gluconeogenesis	X		
Lipid Metabolism			
↑Lipogenesis	X		X
↓Lipolysis	X		X
Protein Metabolism			
↑Amino acid uptake		X	
↑Protein synthesis		X	
↓Protein degradation		X	
↓Gluconeogenesis	X		

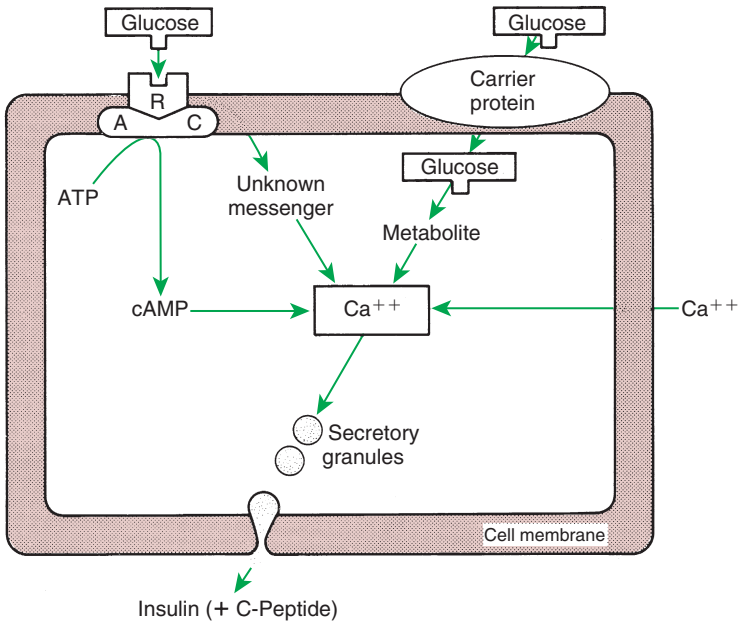
From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.

• BOX 34.1 Factors Affecting Insulin Secretion

Stimuli
- Glucose
- Amino acids
- Fatty acids
- Gastrin
- Pancreozymin-cholecystokinin
- Secretin
- Gastric inhibitory polypeptide
- Glucagon
- Acetylcholine

Inhibitors
- Somatostatin
- Epinephrine
- Norepinephrine

From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.



• Fig. 34.23 Proposed mechanisms of action of glucose on insulin secretion by the β cell. *AC,* Adenyl cyclase; *ATP,* adenosine triphosphate; *R,* receptor; *cAMP,* cyclic adenosine monophosphate. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

norepinephrine also influences insulin secretion because the pancreas has adrenergic innervation by the autonomic nervous system. The pancreas has also cholinergic innervation by the autonomic nervous system, and in contrast to adrenergic stimulation, cholinergic activity increases insulin secretion through the release of acetylcholine.

Insulin Deficiency (Concentration or Action) Produces Diabetes Mellitus, Which Can Culminate in Diabetic Ketoacidosis

A lack or deficiency of insulin produces a syndrome called *diabetes mellitus* (DM). DM may be *type 1,* which is most common in dogs, or *type 2,* which is most common in cats. Type 1 DM is caused by an autoimmune destruction of the β cells of the pancreas; it leads to absolute insulin deficiency and a propensity to develop ketosis. *Diabetic ketoacidosis* (DKA) is the culmination of DM that results in unrestrained ketone body formation in the liver, metabolic acidosis, severe dehydration, shock, and possibly death. Hepatic lipid metabolism becomes deranged with insulin deficiency, and nonesterified fatty acids are converted to acetyl-CoA rather than incorporated into triglycerides. Acetyl-CoA accumulates in the liver and is converted into acetoacetyl-CoA, then ultimately to ketones, including acetoacetic acid, β-hydroxybutyrate (the primary ketone in dogs and cats), and acetone.

When insulin deficiency culminates in DKA, accumulation of ketones and lactic acid in the blood and loss of electrolytes and water in the urine result in profound dehydration, hypovolemia, metabolic acidosis, and shock. Ketonuria and osmotic diuresis, resulting from glycosuria, cause sodium and potassium loss in the urine, exacerbating hypovolemia and dehydration. Nausea, anorexia, and vomiting, caused by stimulation of the chemoreceptor trigger zone through ketonemia and hyperglycemia, contribute to the dehydration caused by osmotic diuresis. Dehydration leads to further accumulation of glucose and ketones in the blood. Stress hormones such as cortisol and epinephrine contribute to the hyperglycemia in a vicious cycle. Eventually, severe dehydration may result in hyperviscosity, thromboembolism, severe metabolic acidosis, renal failure, and finally death.

Historical Findings. Most dogs and cats with DKA present with a previous history of uncomplicated diabetes, including polyuria, polydipsia, and dramatic rapid weight loss in the presence of a good or even ravenous appetite. Other historical findings may include anorexia, weakness, depression, vomiting, and diarrhea. Occasionally, owners fail to notice the significance of the classic signs of DM, and the animals are presented solely with an acute history of DKA. DKA may also develop in previously well-controlled, treated diabetic patients.

Physical Examination Findings. The most common examination findings in DKA are lethargy and depression, dehydration, unkempt hair coat, and muscle wasting. Hepatomegaly is common in both diabetic cats and diabetic dogs. Cataracts are often observed in diabetic dogs. A plantigrade rear limb stance resulting from diabetic neuropathy is often observed in diabetic cats. Other findings include tachypnea, dehydration, weakness, vomiting, and occasionally a strong acetone breath odor. Cats may present recumbent or comatose, which may be a manifestation of mixed ketotic-hyperosmolar syndrome. Icterus can develop from the complicating factors of hemolysis, hepatic lipidosis, and acute pancreatitis.

Laboratory Findings. Glucose level is greatly elevated. The average blood glucose concentration in patients with DKA is 25 mmol/L (455 mg/dL) and can range from 10 mmol/L (182 mg/dL) to greater than 50 mmol/L (909 mg/dL), but the latter is more characteristic of hyperosmolar coma. Although portable glucose meters are typically used to monitor glucose concentrations in DKA, caution is advised in relying on these monitors for baseline glucose concentrations because of inaccuracies in animals with severe hyperglycemia. All DKA patients have a relative or absolute deficiency of insulin and excessive hepatic production of glucose, resulting in hyperglycemia. Hyperglycemia is further exacerbated by dehydration and the corresponding reduction in GFR, and these factors are important determinants of its severity. This is supported by the findings that (1) glucose concentrations should exceed 25 mmol/L (455 mg/dL) only when dehydration is severe enough to reduce GFR, and thus the ability of the kidneys to excrete glucose, and (2) fluid administration alone can significantly reduce blood glucose concentrations.

Osmolality usually is mildly to extremely increased in the DKA patient as a result of hyperglycemia, but this may not be detected, in part, because of concurrent hyponatremia. Sodium (and to a lesser extent potassium), glucose, and urea concentrations are the determinants of the calculated serum osmolality. Reference values for serum osmolality in dogs and cats are approximately 290 to 310 mOsm/kg. Hyperosmolality is generally mild enough to resolve with intravenous fluid and insulin therapies.

Nonketotic hyperosmolar diabetes is defined by extreme hyperglycemia (>30 mmol/L), hyperosmolality (>350 mOsm/L), severe dehydration, and CNS depression without ketone formation and with absent or mild metabolic acidosis. Affected patients are more likely to have underlying renal or cardiovascular disease and to be non–insulin dependent. Although this specific syndrome, as defined in humans, is infrequently encountered in veterinary medicine, ketotic or nonketotic diabetic cats may be seen with significant hyperosmolality and CNS alterations.

Most patients with DKA have a total body K^+ deficit caused by urinary (osmotic diuresis), anorexia, and GI losses (vomiting, anorexia). The metabolic acidosis, relative or absolute insulin deficiency, and serum hypertonicity combine to cause a K^+ shift from the intracellular to the extracellular compartment. This is capable of masking the severity of total body hypokalemia when plasma concentrations are measured. Insulin therapy, as well as correction of the acid-base disturbance with fluids and bicarbonate, will drive serum K^+ intracellularly, potentially causing marked circulating hypokalemia. Polyuric patients are predisposed to severe hypokalemia, whereas oliguric or anuric patients are predisposed to severe hyperkalemia.

In general, DKA causes significant total body Na^+ deficits. Excessive urinary loss of Na^+ results from the osmotic diuresis induced by high glucose and ketone concentrations and the lack of insulin, which usually aids in reabsorption of Na^+ from the distal nephron. Hyperglucagonemia, vomiting, and diarrhea also contribute to the total body Na^+ loss. Hyperosmolality may contribute to a low Na^+ concentration because, as osmolality increases, water is drawn from the interstitium into the vascular space, thus diluting plasma Na^+ and Cl^-.

Phosphorus is the major intracellular anion and is important for energy production and maintenance of cell membranes. Concentrations are regulated by dietary intake, renal elimination, factors promoting its movement into and out of cells, and vitamin D and parathyroid interactions (covered later in this chapter). In DKA, circulating concentrations are usually within reference range or increased initially because of dehydration and renal disease. Phosphorus may also be low at presentation because of urinary loss from osmotic diuresis. As long as renal function is not compromised, a significant decrease in phosphorus should be anticipated with therapy. After insulin administration, phosphorus shifts to

the intracellular compartment with glucose. Clinical signs of hypophosphatemia, such as hemolytic anemia (also seen with Heinz bodies in DKA), lethargy, depression, and diarrhea, may develop once concentrations reach 0.32 mmol/L. Oversupplementation of phosphorus should be avoided because hypocalcemia or metastatic calcification may result.

Magnesium (total serum) is not usually measured routinely, but concentrations may be abnormal in DKA. A study in cats demonstrated high total serum magnesium concentrations at presentation in those with DKA, which improved with treatment of DM. Magnesium deficiency may be caused by poor oral intake, decreased intestinal absorption, increased renal loss, or changes in distribution because it is the second most abundant intracellular cation. Clinical signs of hypomagnesemia include neuromuscular weakness and cardiac arrhythmias, signs that can be seen with other electrolyte alterations. Hypomagnesemia can also cause decreases in other electrolytes, such as potassium and calcium. Correction of deficits may resolve electrolyte disturbances and improve clinical outcome in the severely deficient patient.

Liver enzyme elevations are common in DM patients. Further increases potentially occur in DKA. Alanine aminotransferase and aspartate aminotransferase are most often affected, increasing secondary to hypovolemia and poor hepatic blood flow, with subsequent hepatocellular damage. Further increases in serum alkaline phosphatase concentration may occur if pancreatitis and secondary cholestasis ensue. Cholesterol and triglycerides may be elevated secondary to derangements of lipid metabolism resulting from decreased insulin.

Metabolic acidosis is one of the most prominent features of DKA. As ketone bodies accumulate in the blood and overwhelm the body's buffering capabilities, there is an increase in hydrogen ions and a decrease in bicarbonate. As dehydration worsens, blood flow to peripheral tissues decreases, and the resulting lactic acidosis may contribute to the acid-base disturbance. Acidosis may manifest as lethargy, vomiting, hyperventilation, decreased myocardial contractility, peripheral vasodilation, stupor, and coma. Initiation of insulin therapy (to stop ketogenesis) and fluid therapy (to correct dehydration) will help improve the metabolic acidosis in most patients. Bicarbonate supplementation should be pursued with caution and is generally not recommended unless the patient's blood pH is less than 7.1 or serum bicarbonate is less than 12 mmol/L.

Anion gap may be normal or elevated. An elevated value further characterizes the metabolic acidosis caused by DKA. The anion gap is a representation of the circulating anions not routinely measured on biochemical analyses. The normal anion gap ranges from 10 to 20 and is calculated by the following equation:

$$\text{Osmolality (mOsm)} = 2(Na^+ + K^+ \text{ [mEq/L]}) + \text{Glucose (mmol/L)} + \text{BUN (mmol/L)}$$

In DKA, the ketones become unmeasured anions as they dissociate from ketoacids. However, if significant dehydration is present secondary to the osmotic diuresis and vomiting, lactic acidosis secondary to tissue hypoxia may contribute to the unmeasured anions, thus increasing the anion gap.

Circulating urea and creatinine concentrations may be within reference range or high. These values are high in most patients because of severe dehydration, but renal insufficiency or failure is also a possible cause. Increases in urea and creatinine must be evaluated in light of the urine specific gravity. A low urine specific gravity at presentation does not always guarantee a diagnosis of renal insufficiency or failure, because osmotic diuresis and chronic hypokalemia can contribute to low specific gravities in DM patients. Therefore reevaluation of urea, creatinine, and urine specific gravity must be done after treatment of the crisis. If urea and creatinine are initially elevated and remain static or increase with appropriate therapy, concurrent renal disease is strongly suspected.

The most important part of urinalysis is measurement of glucose and ketones. A strongly positive glucose result confirms DM, and a positive result for ketones confirms DKA. However, a negative ketone result does not definitively rule out ketosis. The nitroprusside reagent used in urine sticks detects only acetoacetate and acetone. It is not as sensitive for β-hydroxybutyrate, the most prevalent ketone body, and therefore may be negative in the presence of ketosis. Using a cutoff value of 3.8 mmol/L for β-hydroxybutyrate was associated with the best combination of specificity (95%) and sensitivity (72%) for DKA.

The presence of pyuria and hematuria on urinalysis, along with confirmation by examination of urine sediment, supports the presence of UTI—a common comorbidity with glucosuria and DM. However, urine culture should be performed regardless of urine sediment activity.

The hemogram may be normal at presentation but usually reveals a leukocytosis with a mature neutrophilia (common in cats) or a stress leukogram. There may be a regenerative or degenerative left shift suggestive of a severe inflammatory and infectious process. The red blood cell count and hematocrit may be elevated as a result of dehydration. Heinz bodies, with or without anemia, may be noted in cats because feline hemoglobin is uniquely susceptible to oxidative damage.

Concurrent Disease. Frequently, an underlying stressful event precipitates the shift from DM to DKA or nonketotic hyperosmolar DM. Impaired immune function secondary to DM increases the risk of infections. The precipitating event may be a UTI or other viral or bacterial infection, or inflammatory disorder, such as pancreatitis, pyelonephritis, cholangiohepatitis, inflammatory bowel disease (IBD), eosinophilic granuloma complex, prostatitis, pyometra, upper respiratory infection, or pneumonia. Other concurrent diseases may include renal insufficiency or failure, hepatic lipidosis, neoplasia, and congestive heart failure. Recent drug therapy may also precipitate a crisis, especially administration of corticosteroids or progestagens. Therefore further diagnostic testing of the diabetic patient that presents in a crisis is essential, particularly abdominal radiography or ultrasonography, as well as thoracic radiography and echocardiography if indicated. Additional testing for concurrent endocrine diseases, such as hyperthyroidism in cats and hypothyroidism and hyperadrenocorticism in dogs, may also be indicated but should be postponed until some control of the DM is achieved, because uncontrolled disease may affect the results of the tests.

Ancillary tests for pancreatitis include abdominocentesis or diagnostic peritoneal lavage if pancreatitis is suspected. Serum amylase and lipase, if determined on presentation, may be elevated in the absence of pancreatitis, secondary to severe dehydration or renal insufficiency, and demonstration of an elevated circulating concentration of trypsin-like immunoreactivity (TLI) may be preferable. Cats and dogs with acute necrotizing pancreatitis usually present with vomiting, abdominal pain, and concurrent DKA. Physical examination findings include icterus, cranial abdominal pain, and abdominal effusion. Radiographs may reveal a "ground glass" appearance of the abdomen, and abdominal ultrasound usually shows enlargement and hypoechogenicity of the pancreas. Concurrent hepatopathies are often present in patients with DKA, but evaluation is complicated by the effect of both DM and DKA on liver enzymes and liver function tests. Ultrasonography may

be more useful in these cases. Diagnostic peritoneal lavage is usually necessary to demonstrate inflammatory nonseptic peritonitis, and abdominal lipase is usually increased dramatically in affected cats and dogs.

Dietary Management Is an Important Consideration in Therapy for Feline Type 2 Diabetes

Type 2 DM is caused by insulin resistance and secondary β-cell failure. Type 2 DM may be managed using oral hypoglycemic agents, diet, or insulin. DM is one of the most common feline endocrinopathies, affecting 1 in 300 cats. The pathogenesis of type 2 DM in cats is reviewed earlier. Diagnosis of DM can be challenging, particularly in the early stages when the cats are non–insulin dependent. However, when clinical signs of diabetes are observed (polyuria, polydipsia, neuropathy), many cats may still benefit from alternatives to insulin therapy. In general, the primary abnormalities associated with type 2 DM, such as obesity and insulin resistance, are reversible. Insulin secretory ability, however, may be reversible (glucose toxicity) or irreversible (pancreatic amyloid deposition). In cats, the differentiation of insulin-dependent (type 1) DM and non–insulin-dependent (type 2) DM is virtually impossible before treatment; therefore the clinician may have to rely on the response to oral hypoglycemic agents as a guide to whether the cat has sufficient β-cell function to be managed with oral hypoglycemic agents.

Goals of therapy for DM include restoration of normal fasting serum glucose concentrations, normalization of serum fructosamine, and reversal or attenuation of chronic complications, such as diabetic neuropathy and nephropathy. As in human patients with type 2 DM, the best approach to cats is a stepwise progression from dietary management to oral hypoglycemics and finally to insulin therapy when "islet burnout" occurs.

Exercise and diet is the cornerstone of therapy in human patients with type 2 DM. In most diabetic cats, exercise is not a reasonable option. One mechanism by which cats may be encouraged to exercise is by feeding the cat multiple small meals hidden in various places within the house. For example, an obese diabetic cat might be encouraged to jump up on the refrigerator or counter to find small amounts of food and then have to hunt for the rest of the food at the opposite end of the house.

In human diabetic patients, fiber supplementation is beneficial in the management of the disease. In humans and dogs, increased amounts of fiber slow the rate of glucose absorption from the intestine and minimize the postprandial fluctuations in blood glucose. This allows better glycemic control and correction of obesity; however, the data in cats are less compelling, and high-fiber diets are no longer advocated for diabetic cats. The cat is an obligate carnivore and, as such, is unique among mammals in its insulin response to dietary carbohydrates, protein, and fat. The feline liver exhibits normal hexokinase activity, but glucokinase activity is virtually absent. *Glucokinase* converts glucose to glycogen for storage in the liver and is important in "mopping up" excess postprandial glucose. Normal cats are similar to diabetic humans because glucokinase levels drop precipitously with persistent hyperglycemia in humans with type 2 DM. Amino acids, rather than glucose, are the signal for insulin release in cats. In fact, one study demonstrated more effective assessment of insulin reserve in cats using the arginine response test rather than a glucose tolerance test. Another unusual aspect of feline metabolism is the increase in hepatic gluconeogenesis seen after a normal meal. Normal cats maintain essential glucose requirements from gluconeogenic precursors (i.e., amino acids) rather than from dietary carbohydrates. As a result, cats can maintain normal blood glucose concentrations even when deprived of food for more than 72 hours; furthermore, feeding has a negligible effect on blood glucose concentrations in normal cats. In summary, the cat is uniquely adapted to a carnivorous diet (mice) and is not metabolically adapted to ingestion of excess carbohydrate.

When type 2 DM occurs in cats, the metabolic adaptations to a carnivorous diet become even more deleterious, leading to severe protein catabolism; feeding it a diet rich in carbohydrates may exacerbate hyperglycemia and protein wasting in these diabetic cats. In humans with type 2 DM, the first recommendation is to restrict excess dietary carbohydrates such as potatoes and bread and to control obesity by caloric restriction. Furthermore, human patients with type 2 DM have improved glycemic control and nitrogen turnover during weight loss when a low-energy diet (high protein) is combined with oral hypoglycemic therapy.

High-protein, low-carbohydrate diets have been found to be beneficial in increasing lean body mass and reducing postprandial hyperglycemia. Caution should be used when high-protein, low-carbohydrate diets are used in cats also treated with insulin because the insulin requirement may decrease. Usually the insulin dose is decreased by 25% in cats given high-protein diets and insulin. On the other hand, high-protein diets and oral hypoglycemic agents appear to be complementary treatments in cats that are underweight. In some studies, 60% to 90% of diabetic cats discontinued insulin injections after 4 to 8 weeks on an ultralow-carbohydrate (<5% dry matter), high-protein diet.

Glucagon

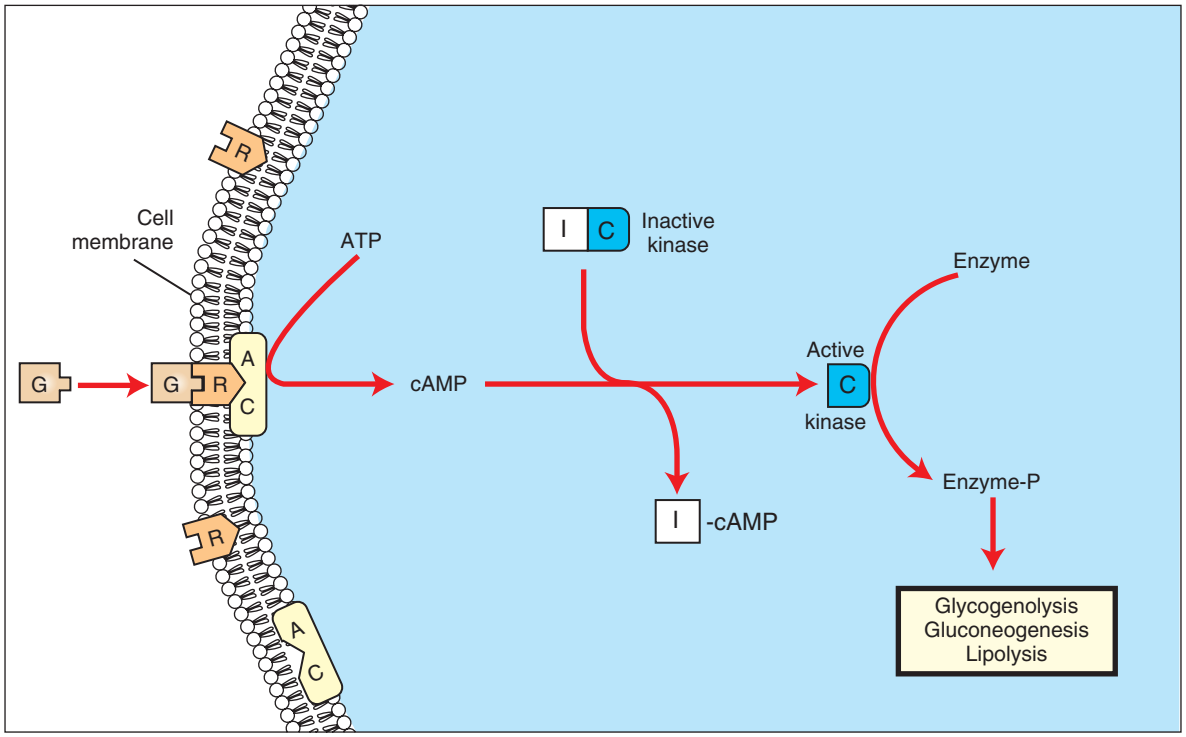
Glucagon is a protein hormone produced by the α cells of the islets of Langerhans. It has a close relationship with insulin in the control of glucose metabolism.

Glucagon is a polypeptide consisting of a single chain composed of 29 amino acids. There is considerable homology in amino acid composition among species. Glucagon is produced in other sites besides the pancreas; the stomach produces a molecule called *gut glucagon* that is identical to the pancreatic glucagon molecule, and the small intestine produces an immunologically similar molecule called *glicentin.* As with other polypeptide hormones, glucagon is first synthesized in the endoplasmic reticulum as part of a precursor molecule, packaged in the Golgi apparatus, and final processing occurs in the secretory granules. Glucagon is released by exocytosis. Glucagon is metabolized mainly by the liver and kidneys. It has a half-life in plasma of about 5 minutes.

The Most Important Functions of Glucagon Are to Decrease Glycogen Synthesis, Increase Glycogenolysis, and Increase Gluconeogenesis

The physiological actions of glucagon are the opposite of those of insulin; the main effect of glucagon is centered in the liver. Glucagon increases cAMP production in the liver, which leads to decreased glycogen synthesis, increased glycogenolysis, and increased gluconeogenesis, the last related to the effects of glucagon on protein metabolism (Fig. 34.24). The net result is an increase in glucose concentrations in the blood.

Changes in glucagon secretion counterbalance the effects of insulin in association with the daily ingestion of food and the intervals between food intake periods. After the consumption of food, the initial response of the metabolic system is increased insulin secretion, which results in conservation of energy through the formation of storage forms of carbohydrates, fats, and proteins.



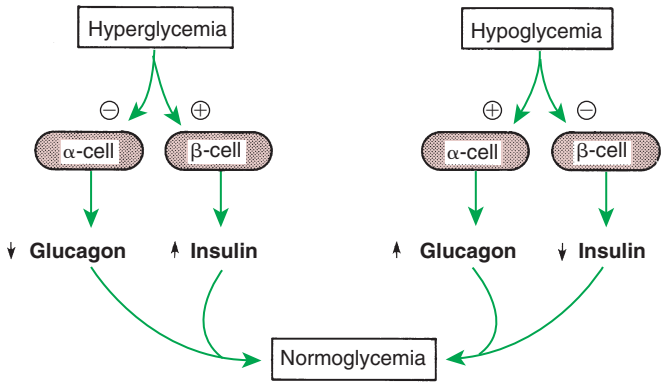
• **Fig. 34.24** Mechanism of action of glucagon (*G*) on its target cells. *AC,* Adenyl cyclase; *ATP,* adenosine triphosphate; *cAMP,* cyclic adenosine monophosphate; *I* and *C,* inhibitory and catalytic subunits of the kinase, respectively; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

Glucagon secretion, which begins with the ingestion of food, increases as the interval from food ingestion lengthens and blood glucose concentrations begin to decline. This secretion allows the individual to mobilize energy stores for the maintenance of glucose homeostasis (i.e., to prevent postprandial hypoglycemia) (Fig. 34.25).

Glucagon Synthesis Is Stimulated by Decreased Glucose Concentrations in the Blood

The main factor that regulates glucagon secretion is plasma glucose concentration. In contrast to insulin synthesis, decreased glucose concentrations stimulate glucagon synthesis and release, a relationship that represents a negative-feedback system. It must be remembered that glucagon regulation works in tandem with insulin regulation to maintain glucose concentrations within the physiological range. In fact, if glucagon were not secreted to maintain blood glucose concentrations, the individual would die of hypoglycemic shock. Because the α cells require insulin for glucose entry into the cells (as do most cells), in clinical syndromes involving insulin insufficiency (DM), glucose entry into the α cells is reduced and plasma glucagon concentrations are paradoxically elevated. Glucagon promotes lipolysis and an increase in fatty acids, which has a negative-feedback effect on glucagon secretion.

Protein ingestion represents an exception to the rule of opposite responses of glucagon and insulin. The release of both insulin and glucagon in response to protein ingestion appears logical; increased insulin secretion, in response to increased plasma amino acid levels, leads to lower glucose concentrations, and increased glucagon would counteract this through increased hepatic gluconeogenesis, resulting in maintenance of blood glucose within normal limits. The complementary responses of insulin and glucagon allow growth to occur in animals fed a diet only of protein and fat.



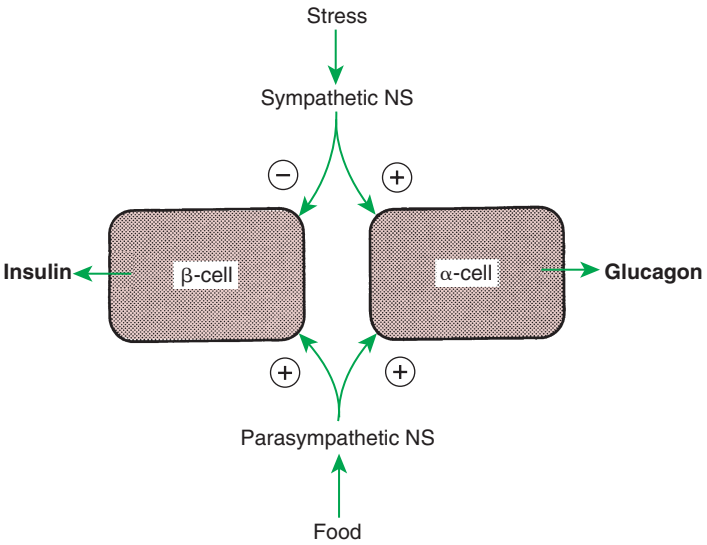
• **Fig. 34.25** Effects of hyperglycemia and hypoglycemia on the secretion of insulin and glucagon by the pancreatic β cells and α cells, respectively. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

Intestinal hormones, with the exception of secretin, stimulate both glucagon and insulin secretion. A similar (inhibitory) response to somatostatin is observed for both glucagon and insulin. Both sympathetic and parasympathetic stimulation of the autonomic nervous system induce secretion of glucagon (Fig. 34.26).

Some birds have a predominance of glucagon in the pancreas, which suggests that glucagon may have a more important role in carbohydrate metabolism of avian species than in mammals.

Somatostatin

As indicated in Chapter 33, somatostatin was first described in the brain as a 14–amino acid peptide that inhibits growth hormone secretion by the pars distalis. The molecule has since been identified in a number of tissues, including other areas of the brain, the GI tract, and the D cells of the pancreatic islets. Its synthesis and secretion are similar to those observed for other protein hormones. The metabolism of somatostatin is rapid (about 5 minutes) and occurs mainly in the liver and kidneys.



• **Fig. 34.26** Regulation of insulin and glucagon secretion by the autonomic nervous system. *Plus signs* indicate stimulation; *minus sign* indicates inhibition. *NS,* Nervous system. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

The Main Functions of Somatostatin Are to Inhibit the Secretion of Hormones Produced by the Pancreas (Insulin, Glucagon, Pancreatic Polypeptide) and Regulate Growth Hormone

The actions of somatostatin can be classified as inhibitory. Pancreatic somatostatin inhibits the digestive processes by decreasing nutritive absorption and digestion. The motility and secretory activity of the GI tract are decreased by somatostatin. One of the most important physiological functions of pancreatic somatostatin is the regulation of the endocrine cells of the pancreas (Fig. 34.27). Somatostatin inhibits secretion of all endocrine cell types of the islets of Langerhans, including the D cells. The α cells are more affected by the inhibitory action of somatostatin than are β cells; therefore, glucagon secretion is more affected by somatostatin than is insulin secretion.

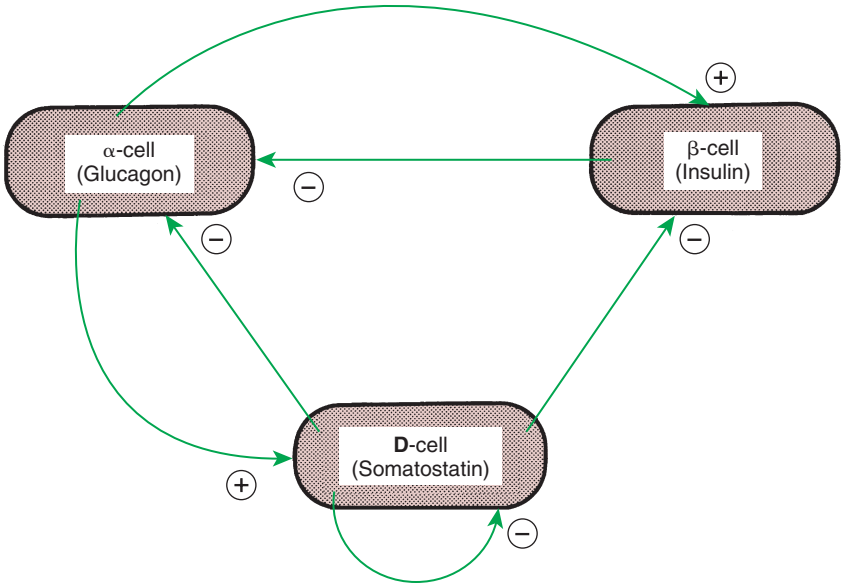
Somatostatin secretion is increased by nutrients (e.g., glucose, amino acids) and by the neurotransmitters of the autonomic nervous system (epinephrine, norepinephrine, acetylcholine). Of the hormones produced by the pancreas, only glucagon stimulates somatostatin secretion.

Pancreatic Polypeptide

Pancreatic polypeptide, a 36–amino acid polypeptide, is produced by the F cells of the pancreas (see Fig. 34.20). In contrast to somatostatin secretion, pancreatic polypeptide secretion is limited to the pancreas.

The effects of pancreatic polypeptide are directed toward the GI tract. The secretion of pancreatic enzymes and the contraction of the gallbladder are inhibited by the actions of this hormone. Both gut motility and gastric emptying are increased by the action of pancreatic polypeptide.

The secretion of pancreatic polypeptide is stimulated by intestinal hormones, including cholecystokinin, secretin, and gastrin. Stimulation of the vagus nerve also promotes pancreatic polypeptide secretion. The ingestion of protein is stimulatory for secretion, whereas carbohydrates and fats have little effect. As indicated previously, somatostatin inhibits pancreatic polypeptide secretion.



• **Fig. 34.27** Possible cell-to-cell interactions in the pancreatic islets. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

Calcium and Phosphate Metabolism

Calcium Is Important for Many Intracellular Reactions, Including Muscle Contraction, Nerve Cell Activity, Release of Hormones Through Exocytosis, and Activation of Enzymes

The control of calcium and phosphate metabolism is important because these ions play a major role in physiological processes. Calcium homeostasis is tightly controlled; adjustments are made within a range of 5% of normal. Calcium is important for a number of intracellular reactions, including muscle contraction, nerve cell activity, the release of hormones through the process of exocytosis, and the activation of several enzymes. Calcium is important for coagulation of blood and for maintaining the stability of cell membranes and the linkage between cells. On a less acute basis, calcium is important for the structural integrity of bone and teeth.

Phosphate Is Important for the Structure of Bone and Teeth, and Organic Phosphate Serves as Part of the Cell Membrane and Several Intracellular Components

Phosphate concentrations are controlled by the same systems that control calcium concentrations. *Inorganic phosphate* in blood serves as the source of phosphate, which is important for the structure of bone and teeth. Inorganic phosphate also functions as an important H^+ buffering system in blood. *Organic phosphate* is an important part of the cell, including the plasma membrane and intracellular components, such as nucleic acids, adenosine triphosphate, and adenosine monophosphate.

The Most Important Body Pool of Calcium Involved in Homeostasis Is the Extracellular Fluid Component, but Calcium in Bone Serves as a Larger Reservoir

Almost all calcium (99%) in the body is in bone in the form of *hydroxyapatite* crystals, which contain calcium, phosphate, and water. The next largest pool of calcium is intracellular calcium. As stated previously, calcium is important for the response of cells in carrying out their physiological activities, including the secretion of hormones. In the inactive cell state, calcium concentrations are relatively low in the cytosol; calcium is bound to proteins or contained within the mitochondria or granules of the endoplasmic reticulum. Increased intracellular calcium concentrations are indicative of increased cell activity.

The smallest pool of calcium, which resides in the ECF, is the most important pool for physiological control of calcium concentrations in the blood. This component comprises interstitial calcium, blood calcium, and a small (0.5%) but important part of the bone calcium pool, which exists as amorphous crystals or in solution. The soluble bone calcium pool allows access to the large reserve of calcium that resides in bone.

The regulation of calcium levels involves control of the movement of calcium between the ECF and three body organs: bone, GI tract, and kidneys. The exchange of calcium ions between the ECF and intracellular fluid occurs in conjunction with the control of intracellular metabolism, with little effect on plasma concentrations of calcium.

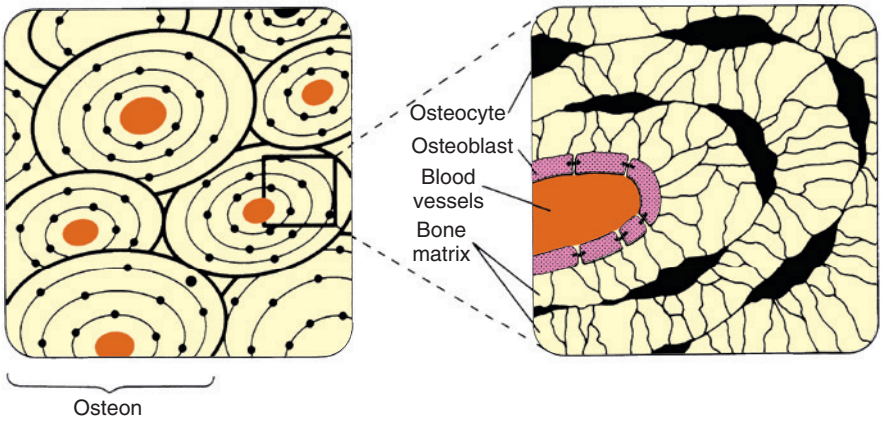
The absorption of calcium from the GI tract is by passive diffusion and active transport. The *passive diffusion* of calcium across the intestinal mucosa occurs in the presence of high concentrations and, as such, is not an important aspect of calcium absorption. *Active transport* involves the movement of calcium into the intestinal cell down a concentration gradient, which is facilitated by carrier proteins located on the luminal side of the mucosal cell. Calcium is moved through the serosal side of the mucosal cell into the interstitial fluid through a calcium pump system. The active transport system adjusts according to the amount of calcium in the diet, becoming more active when calcium concentrations in the diet are lower and less active when calcium concentrations are higher. Calcium excretion into the GI tract is not affected by calcium uptake, and this can exacerbate conditions involving hypocalcemia. The GI tract serves as the source of calcium for the body, even though both absorption and excretion of calcium occur through the tract. As discussed later, vitamin D plays an important role in the absorption of calcium from the GI tract.

The kidneys serve as the route of excretion of calcium. Most of the calcium that passes into the kidneys is reabsorbed with a net loss of only about 2%. This amount is matched by net absorption of calcium by the GI tract. Most of the calcium filtered by the kidneys is reabsorbed in the proximal tubules; the next largest amount is absorbed by the distal tubules, and a lesser amount by the ascending loop of Henle. The distal tubules are under hormonal control and therefore are the sites of regulation of calcium in the kidneys.

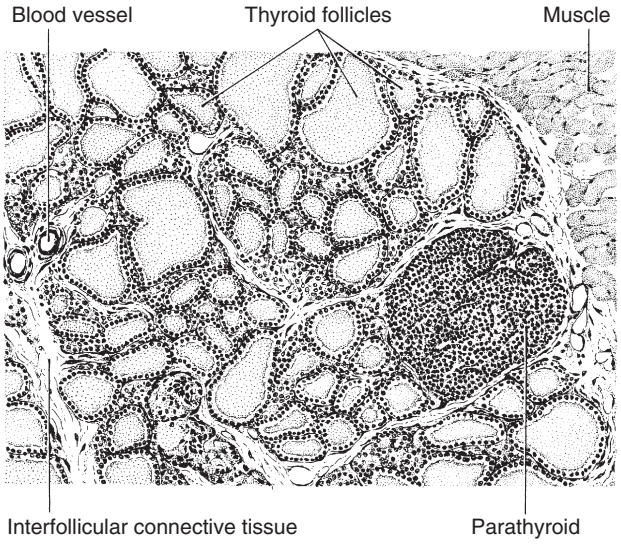
The most important regulation of calcium metabolism between bone and ECF involves the soluble portion of bone. Amorphous crystals and soluble calcium, which form the source of ready exchange of ions with the blood, are located between the *osteoblasts,* which line the blood vessel channels, and the *osteocytes,* which are deeper in the bone (Fig. 34.28). These two cell types have cytoplasmic projections that interact intimately through the presence of tight cell junctions. For labile bone calcium to reach the blood, calcium must cross the membrane barrier created by the osteoblasts and osteocytes. Movement of calcium from stable bone into the ECF also occurs but has little effect on the acute regulation of calcium concentrations. The process of remodeling bone, which occurs on a continuous basis, involves the breakdown of hydroxyapatite crystals by osteoclasts, a laying down of organic matrix by osteoblasts in the tunnels made by the osteoclasts, and finally the mineralization of the organic matrix by hydroxyapatite crystals. If an animal is subjected to prolonged changes involving calcium metabolism, the slowness of bone calcium exchange can have a significant effect on calcium metabolism. Chronic abnormalities of calcium homeostasis can cause skeletal pathology.

Parathyroid Hormone

Parathyroid Hormone (PTH) and Calcitonin Act Together to Provide Calcium Homeostasis With PTH Increasing, and Calcitonin Decreasing, Ionized Calcium. The main organ involved in the control of calcium and phosphate metabolism is the parathyroid gland (Fig. 34.29). Most domestic animals have two pairs of parathyroid glands generally located at the poles of the two lobes of the thyroid gland; the pig has only one pair of parathyroid glands, which lie anterior to the thyroid. The cranial pair of parathyroid glands in dogs and cats is at the craniolateral poles of the thyroid, and the pair of ruminants and horses is anterior to the thyroid. The caudal pair of parathyroid glands in dogs, cats,



• **Fig. 34.28** Structure of the osteon, the functional unit of bone, depicted in cross section at two magnifications. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.29** Depiction of a section of the thyroid and parathyroid glands of the rat as seen under low power of the microscope. Notice that the parathyroid gland lies near the surface of the thyroid gland and is surrounded on three sides by the thyroid follicles. (From Turner CD, Bagnara JT. *General Endocrinology.* 6th ed. Philadelphia: Saunders; 1976.)

and ruminants is located within the medial surface of the thyroid, whereas in the horse it is near the bifurcation of the carotid trunk. The parathyroid cells in the active process of hormone secretion are called *chief cells,* whereas inactive, or degenerate, cells are called *oxyphil cells.*

The synthesis of PTH is similar to the synthesis of other protein hormones; a prepro-PTH of 115 amino acids is synthesized in the rough endoplasmic reticulum and then cleaved by 25 amino acids to form pro-PTH. A 6–amino acid "pro" portion is removed by the Golgi apparatus; the resulting PTH has 84 amino acids. PTH is secreted by the process of exocytosis. PTH is rapidly metabolized by the liver and kidneys and has a relatively short half-life (5–10 minutes) in blood.

The effect of PTH is to increase calcium and decrease phosphate concentrations in ECFs. PTH has direct effects on bone and kidney metabolism of calcium and indirect effects on GI metabolism of calcium. The initial effect of PTH on bone is to promote the transfer of calcium across the osteoblast-osteocyte membrane. This level of action occurs without the movement of phosphate and therefore has no effect on phosphate concentrations in blood. PTH has additional effects on stable bone, which results in the resorption of the bone. This effect involves increased osteoclast activity and an inhibition of osteoblast activity. The effect of PTH on stable bone results in the release of both calcium and phosphate.

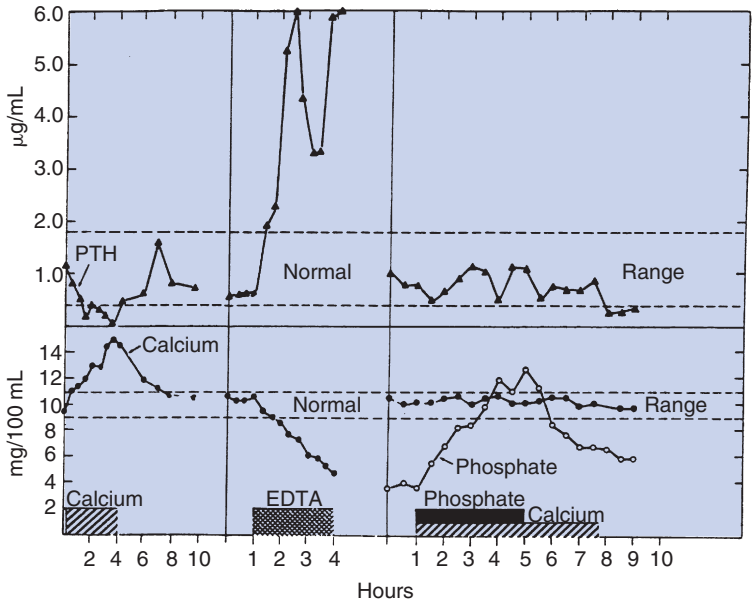
PTH acts on the distal convoluted tubules of the kidneys to increase absorption of calcium and decrease renal phosphate reabsorption through an effect on the proximal tubules. PTH is also involved in the activation of vitamin D at the kidney level. PTH mediates the absorption of calcium from the gut indirectly through its effect on vitamin D.

PTH secretion is controlled by free (ionized) calcium concentrations in blood; decreases in calcium levels stimulate PTH secretion, and increases in calcium turn off secretion (Fig. 34.30). Both actions are mediated by an effect on cAMP metabolism. Epinephrine stimulates PTH secretion through stimulation of β-adrenergic receptors. Magnesium affects PTH secretion in the same manner as calcium, but its physiological effect is much less. Sleep affects the secretion of PTH; values are highest immediately after waking.

Calcitonin

Calcitonin, a hormone produced by cells in the thyroid gland, also affects calcium metabolism. Cells of the type involved in the synthesis of calcitonin—parafollicular, or C, cells—are scattered throughout the thyroid gland and are distinctly different from the cells that synthesize thyroid hormones. During the early studies of calcitonin in animal classes such as fish, amphibians, reptiles, and birds, which have separate thyroid and ultimobranchial glands, it was found that all the calcitonin activity was in the ultimobranchial glands. Therefore the calcitonin cells represent ultimobranchial gland tissue that has been incorporated into the thyroid during embryonic development.

Calcitonin, synthesized as a preprohormone, has 32 amino acids; a ring structure at the amino terminus contains a disulfide link that bridges between amino acids 1 and 7. The processing of the molecule is interesting because calcitonin is located in the middle of procalcitonin, so an additional enzyme cleavage is required for the formation of the active molecule. The secretion of calcitonin is by exocytosis from granules.



• **Fig. 34.30** Changes in plasma immunoreactive parathyroid hormone (*PTH*) levels in response to hypercalcemia induced by calcium infusion; in response to hypocalcemia produced by ethylenediaminetetraacetic acid (*EDTA*) infusion; and in response to hyperphosphatemia with normocalcemia in a cow. (From Capen CC. The calcium regulating hormones: parathyroid hormone, calcitonin, and cholecalciferol. In: McDonald LE, ed. *Veterinary Endocrinology and Reproduction.* 4th ed. Philadelphia: Lea & Febiger; 1989.)

Calcitonin acts as a counterbalance to PTH because it causes hypocalcemia and hypophosphatemia. The effect of calcitonin on mineral metabolism is mainly on bone (Fig. 34.31). Calcitonin decreases the movement of calcium from the labile bone calcium pool (behind the osteoblast-osteocyte barrier) to the ECF and decreases bone resorption through an inhibitory effect on osteoclasts. Whereas the inhibition of bone resorption explains one aspect of the hypophosphatemic effects of calcitonin, calcitonin also increases movement of phosphate from the ECF into bone. Calcitonin decreases GI activity directly by inhibiting gastric acid secretion and indirectly by inhibiting gastrin secretion. The physiological importance of this is not known. Calcitonin also increases renal excretion of calcium and phosphate.

The control of calcitonin secretion is by calcium; increased calcium concentrations cause increased secretion of calcitonin. The physiological control of calcium metabolism by calcitonin operates in situations of hypercalcemia with increased secretion of calcitonin and concomitant inhibition of PTH secretion. During hypocalcemic conditions (see Fig. 34.31), calcitonin synthesis is inhibited, and PTH becomes responsible for reestablishing normal calcium concentrations in the ECFs. GI hormones, including gastrin, cholecystokinin, secretin, and glucagon, stimulate the secretion of calcitonin, with gastrin the most potent. These hormones limit postprandial hypercalcemia.

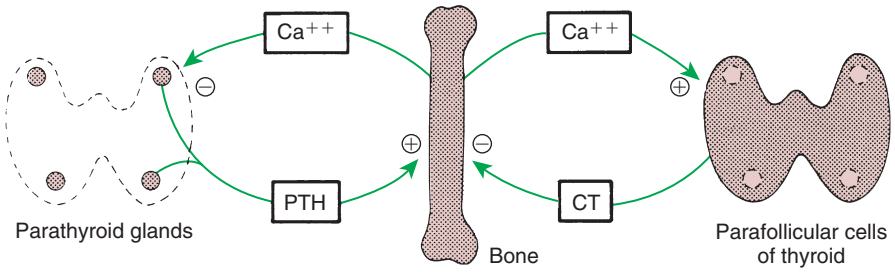
Vitamin D

Vitamin D is important for the absorption of calcium from the gut. It is a steroid-like molecule, and because it is produced in one tissue and transported by the blood to a distant site of action, it should probably be called a hormone instead of a vitamin. Epithelial cells of the skin synthesize the immediate precursor of vitamin D, 7-dehydrocholesterol, from acetate. Exposure of the skin to ultraviolet light results in cleavage of the C-9 and C-10 bonds of 7-dehydrocholesterol, which results in the formation of vitamin D (Fig. 34.32). Much vitamin D in animal species comes from the diet, and the contribution of epithelial synthesis is more variable than in humans. The vitamin D molecule, as such, is inactive and must be transformed by both the liver and the kidney before the molecule is biologically activated. The liver first hydroxylates the molecule at the C-25 position producing 25-hydroxyvitamin D, and the kidney subsequently hydroxylates the molecule at C-1 to produce the active compound, 1,25-$(OH)_2$-vitamin D (1,25-vitamin D; calcitriol).

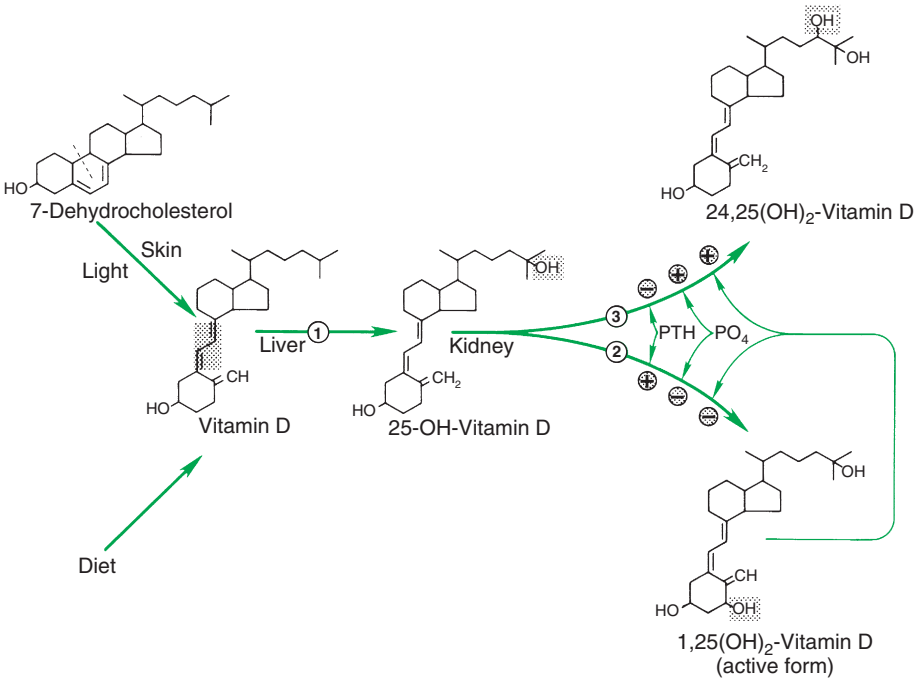
Control of the C-1 hydroxylase in the kidney by PTH is the most important control linkage for the synthesis of 1,25-vitamin D. Decreases in calcium concentrations stimulate PTH secretion, which in turn favors the synthesis of active vitamin D and increased intestinal absorption of calcium. Phosphate also regulates vitamin D metabolism. Increased serum phosphate concentrations stimulate an enzyme that promotes hydroxylation of C-24 (instead of C-1) by the kidney, which leads to the formation of 24,25-$(OH)_2$-vitamin D, an inactive molecule. The active molecule, 1,25-vitamin D, also regulates itself by decreasing C-1 hydroxylase and increasing C-24 hydroxylase activity; decreased amounts of active vitamin D is the result.

Because of its lipid nature, 1,25-vitamin D is transported by binding to proteins in the plasma. Most of vitamin D is carried in association with a specific α globulin called *transcalciferin* (vitamin D binding protein), a molecule synthesized by the liver.

The most important effects of vitamin D involve increased absorption of calcium by the GI tract. Vitamin D stimulates the synthesis of protein within the mucosal cells, which aids the rate-limiting step in calcium absorption, that is, movement of calcium into the mucosal cell (Fig. 34.33). Because the intestinal effect of vitamin D depends on the activation of protein synthesis by mucosal cells, the effect on calcium absorption usually requires several hours. Although the stimulation of protein synthesis relates mostly to active transport of calcium, vitamin D also stimulates passive transfer of calcium. Vitamin D also has effects on bone, promoting the movement of calcium ions from the labile pool into ECFs and



• **Fig. 34.31** Negative-feedback loops controlling parathyroid hormone (*PTH*) and calcitonin (*CT*) secretion. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



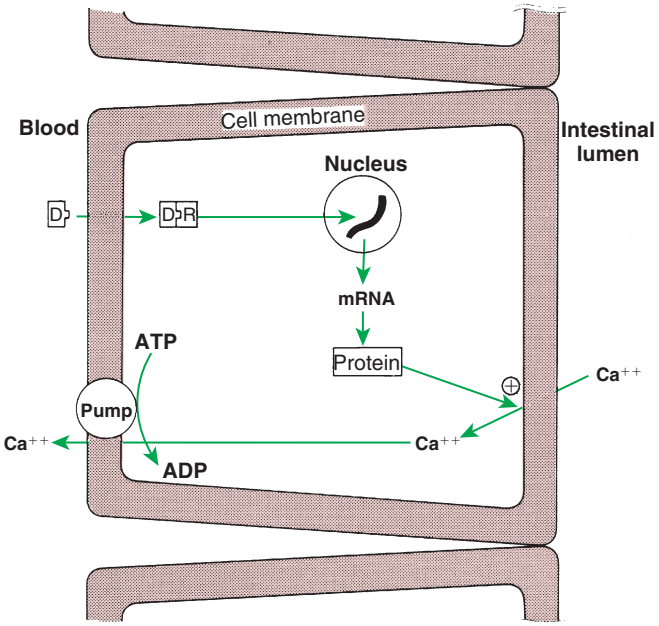
• **Fig. 34.32** Synthesis and metabolism of vitamin D. The position of hydroxylation of 25-OH-vitamin D in the kidney is controlled by parathyroid hormone (*PTH*), phosphate (PO_4), and 1,25-$(OH)_2$-vitamin D. *Shading* indicates structural change at each step; *dashed line* indicates the position of cleavage of 7-dehydrocholesterol to produce vitamin D. Enzymes: *1,* 25-hydroxylase, *2,* 1α-hydroxylase, *3,* 24-hydroxylase. *Plus signs* indicate stimulation; *minus signs* indicate inhibition. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)

the resorption of bone, as well as enhancing the effects of PTH on bone metabolism of calcium.

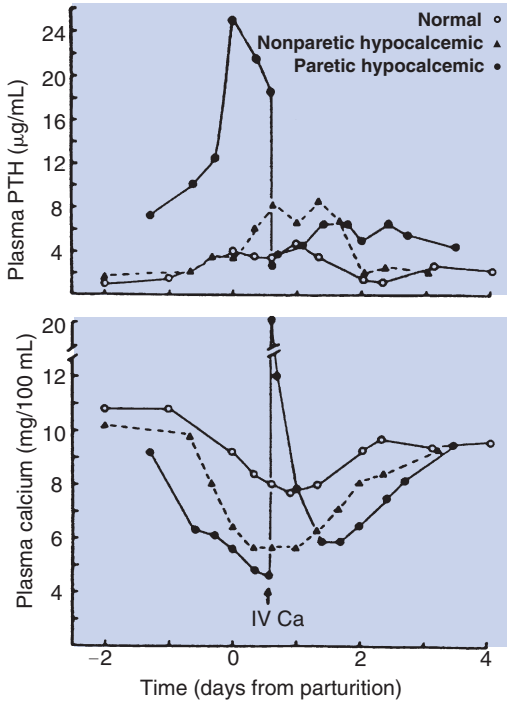
The control of 1,25-vitamin D synthesis is by PTH and phosphate. A decrease in calcium concentrations results in increased PTH secretion and increased formation of 1,25-vitamin D through enhancement of C-1 hydroxylation (Fig. 34.34). This action leads to the correction of hypocalcemia by increasing absorption of calcium by the gut. A decline in phosphate concentrations results in decreased inhibition of the C-1 hydroxylation, which indirectly results in increased 1,25-vitamin D production and increased absorption of phosphate. Some evidence suggests that hormones associated with pregnancy, such as growth hormone and prolactin, increase 1,25-vitamin D production by stimulating C-1 hydroxylation.

In the overall control of calcium metabolism, PTH is primarily responsible for the maintenance of calcium homeostasis. The primary target tissue for PTH in calcium homeostasis is the labile pool in bone; changes in renal absorption of calcium are also important. In the case of long-term calcium deficit in the diet, both PTH and 1,25-vitamin D are important for correction of the deficit. Decreased dietary calcium leads to decreased concentrations of calcium in the ECFs and the release of PTH. PTH affects resorption of calcium by the kidneys, but most importantly, for long-term correction of the problem, it causes increased 1,25-vitamin D secretion with increased absorption of dietary calcium. PTH also contributes to the overall calcium pool through its effect on stable bone, that is, the promotion of resorption.

Hypercalcemia has a variety of causes, including malignancy, hyperparathyroidism, fungal disease, osteoporosis, hypoadrenocorticism, chronic renal disease, and hypervitaminosis D. The initial signs of hypercalcemia are polydipsia and polyuria resulting from impaired response of distal renal tubules to antidiuretic hormone. Listlessness, depression, and muscle weakness result from depressed excitability of neuromuscular tissue. Mild GI signs of hypercalcemia include inappetence, vomiting, and constipation. Persistent mild elevations in serum calcium (12–14 mg/dL) can cause uroliths and



• **Fig. 34.33** Mechanism of action of 1,25-$(OH)_2$-vitamin D (*D*) to increase calcium absorption in the intestine. *Plus sign* indicates stimulation. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; *mRNA,* messenger ribonucleic acid; *R,* receptor. (From Hedge GA, Colby HD, Goodman RL. *Clinical Endocrine Physiology.* Philadelphia: Saunders; 1987.)



• **Fig. 34.34** Development of varying degrees of hypocalcemia in cows near parturition, with corresponding increases in plasma parathyroid hormone (*PTH*) levels. The cow developing severe hypocalcemia (<5 mg/100 mL) had a considerably greater increase in plasma PTH levels than the moderate rise detected in the nonparetic hypocalcemic and the normal cow. Note that PTH levels declined rapidly after treatment of the paretic cow with intravenous calcium (*IV Ca*). (From Mayer CP. The roles of parathyroid hormone and thyrocalcitonin in parturient paresis. In: Anderson JJB, ed. *Parturient Hypocalcemia.* New York: Academic Press; 1970.)

signs of urinary tract disease such as hematuria and stranguria. On the other hand, severe hypercalcemia (>14 mg/dL) can progress rapidly to acute renal failure when the calcium-phosphorus product ($Ca \times PO_4$) exceeds 60 to 80 mg/dL because of mineralization of renal tissue.

The diagnostic approach to hypercalcemia consists of ruling out the most common causes: hypercalcemia of malignancy, renal disease, and primary hyperparathyroidism. A thorough history and physical examination, including lymph node and rectal examination (for anal sac adenocarcinoma), complete blood cell count, urinalysis, serum chemistry profile, and chest and abdominal radiographs, are necessary to search for underlying neoplastic processes. Measurement of ionized calcium and PTH is useful to confirm hypercalcemia and gain insight into parathyroid function. Parathyroid hormone-related peptide (PTHrP) is elevated with hypercalcemia in some malignancies and can be measured in plasma. If lymphoma is not detected on the minimum database, a bone marrow examination and survey skeletal radiographs may be necessary. When a diagnosis of neoplasia has been excluded, the next primary differential for hypercalcemia is chronic renal failure. This is the most difficult differential to exclude because other causes of hypercalcemia may result in renal damage because of soft tissue mineralization of the kidneys. Therefore an animal with hypercalcemia and azotemia could have primary hyperparathyroidism, primary renal failure with secondary renal hyperparathyroidism, or vitamin D intoxication. Furthermore, patients with hypercalcemia secondary to renal disease may also exhibit elevations in intact PTH. Diagnosis of primary hyperparathyroidism is based on the findings of hypercalcemia (preferably ionized), hypophosphatemia (unless azotemic), high-normal to elevated serum PTH concentrations, and a mass in the cervical region. A PTH concentration above midnormal in the presence of elevated total and/or ionized calcium is considered inappropriate for the calcium level and would be considered diagnostic for primary hyperparathyroidism. The classic biochemical findings in animals with hypoparathyroidism are *hypocalcemia* (both total and ionized) and *hyperphosphatemia.* Other causes of hypocalcemia include iatrogenic (postthyroidectomy) hypoparathyroidism, chronic and acute renal failure, acute pancreatitis, hypoalbuminemia, puerperal tetany (eclampsia), ethylene glycol intoxication, intestinal malabsorption, and nutritional secondary hyperparathyroidism. Early signs of hypocalcemia are nonspecific and include anorexia, facial rubbing, nervousness, and a stiff stilted gait. Later signs progress to paresthesias, hyperventilation, and finally generalized tetany and seizures.

Primary hypoparathyroidism is diagnosed by means of an intact PTH assay. Serum or plasma PTH concentrations should be measured on a freshly drawn morning sample in a fasting animal. Handling of the sample is crucial to appropriate diagnosis because PTH may degrade if subjected to warm temperatures. *Intact PTH* refers to the entire 85–amino acid sequence of PTH; this is measured in a double-antibody "sandwich" assay in most endocrine laboratories that perform PTH measurement. For the diagnosis of primary hypoparathyroidism, the sample should be analyzed for both ionized calcium and intact PTH. Low ionized calcium and undetectable intact PTH concentrations are diagnostic for hypoparathyroidism.

Acknowledgement

The authors would like to acknowledge and thank the late Dr. George H. Stabenfeldt as a previous contributor to this chapter. Some of his original work has been incorporated into this edition.

CLINICAL CORRELATIONS

Diabetes Mellitus

History

You are presented with a 10-year-old, intact, female poodle whose owner is upset because the dog urinates in the house. In addition, the owner has noticed that the animal now drinks larger amounts of water than in the past. Although the owner indicates the dog has a good appetite, it appears to have lost weight over the past few months.

Clinical Examination

During the examination, you check the dog's breath and detect a sweet odor. Among the organ systems you check are the eyes, and you find developing cataract formation. Because you have seen this dog many times, you check its weight and find that it has lost 2 pounds since its last admittance a year ago. You are able to run a blood glucose determination in your hospital and tell the owner that the glucose concentration is 278 mg/dL.

Comment

The findings in diabetes mellitus (DM) are all attributable to inadequate availability of insulin. Glycogen synthesis decreases in tissues, whereas glycogenolysis and gluconeogenesis increase, contributing to the high concentrations of glucose found in blood. When glucose concentrations exceed the reabsorption capacity in the tubular cells of the kidney, glucose appears in the urine. The loss of glucose in urine causes an osmotic diuresis (polyuria), and the dog compensates for this by drinking additional amounts of water. The sweetness of the breath is caused by the presence of ketone bodies. These form as the result of decreased triglyceride synthesis in adipose tissue, which stimulates lipase activity and the release of free fatty acids. These fatty acids are metabolized to ketone bodies (acetoacetate, acetone, β-hydroxybutyrate) by the liver in a situation of excess fatty acids. The end result is both a ketonemia and a ketonuria. Protein metabolism shifts toward catabolism during DM with decreased protein synthesis and increased protein degradation by muscle cells. This process increases the circulating concentrations of amino acids available to the liver for gluconeogenesis. The end result is nitrogen loss and a decrease in the muscle mass of the animal. The changes noted in the lenses of the eyes represent only one of a number of changes that occur in the presence of DM as a result of glycosylation of proteins, including lens proteins and hemoglobin.

Treatment

Insulin administration is essential in the treatment of insulin-dependent (type 1) DM. During the initial stages of treatment, considerable care must be taken to ensure that the dosage is correct. The goal of treatment is to maintain glucose concentrations between a low of 80 mg/dL and a high of 200 mg/dL, with a serum fructosamine less than 400 μmol/L using one or two insulin injections every 24 hours. Too much insulin has the potential for producing a hypoglycemic coma. Two other important aspects of treatment include feeding the diabetic dog a diet high in soluble fiber in conjunction with insulin administration and adequate exercise. Finally, the owner needs to be educated and prepared for the necessity of intensive involvement in the management of the disease.

Pancreatic Tumor in a German Shepherd

History

A 10-year-old female, spayed, overweight German shepherd presents for a progressive history of collapse and weakness over the last few weeks. The owners noticed that she did not want to walk as far, that she seemed worse in her back legs, and had less energy. She also seemed slightly dull and disoriented recently, particularly after exercise. They also thought she had been eating, drinking, and urinating more frequently. They wondered if she had gained weight. Today, a few hours before presentation, she had what appeared to be one or two seizures.

Clinical Examination

On examination, she was ataxic (delayed conscious proprioception/ability to know where her legs are) more so in her hind end than her front legs. She also appeared to be weak in all four limbs, with her hind legs being more affected than her front end. The rest of the physical and neurological examination was normal.

Comment

Routine blood work demonstrated hypoglycemia. Based on the concern for pancreatic and insulin function, insulin levels were submitted. Abdominal radiographs were normal. Abdominal ultrasound revealed a mass in the pancreas. No evidence of metastatic lesions was found in the abdomen. Computed tomography (CT) could be used to further delineate the insulinoma. Insulin levels are increased.

When function is normal, glucose is metabolized to ATP within the pancreatic β cells. This results in closure of the K^+-ATP-sensitive channels. There is a decrease in K^+ efflux, which then depolarizes the β cells and opens the voltage-sensitive calcium channels. The increase in Ca^{2+} causes the exocytosis of insulin.

With abnormal function (neoplasia), the neoplastic cells secrete insulin independent of blood glucose. In some cases, localized growth hormone levels are increased in dogs with insulinoma. It has been proposed that growth hormone may increase through paracrine or autocrine mechanisms.

With hypoglycemia, there is an increase in glucagon, catecholamines, growth hormones, and glucocorticoids. Glucagon and catecholamines are most important in regulation of blood glucose.

Clinical signs are associated with hypoglycemia. Decreased glucose concentrations in the CNS result in decreased cerebral oxygenation and can cause lethargy, weakness, and seizures. Clinical signs can be episodic due to the counterregulatory mechanisms described.

Treatment

Immediate treatment includes intravenous dextrose. Long-term management includes surgical removal of the mass(es) if possible, based on size and extent of tumor(s). Some dogs will develop diabetes. Diet recommendations include small, more frequent feedings of high protein, fat, and complex carbohydrates with decreased simple sugars. Medical treatment includes streptozocin, which destroys the β cells; treatment of the hypoglycemia with prednisone, which increases gluconeogenesis, or with dexamethasone; diazoxide, which decreases insulin secretion through inhibiting closure of ATP-dependent K^+ channels of β cells; and octreotide, which inhibits insulin secretion. Median survival time for dogs with partial pancreatectomy was 12 to 14 months. Prognosis is better in dogs with stage I versus stage II or III (50% vs. 20% normothermic at 12–14 months). Stage I was defined as primary tumor with no evidence of regional lymph nodes or metastasis; stage II was primary tumor with regional lymph node metastasis; and stage III was primary tumor, no regional lymph node metastasis, but liver involvement. Most dogs are reported as stage II or III involvement.

Practice Questions

1. The other main hormone secreted by the thyroid gland, in addition to tetraiodothyronine and triiodothyronine, is:
- **a.** Calcitonin
- **b.** Insulin
- **c.** Parathyroid hormone
- **d.** Glucagon
- **e.** Somatostatin

2. The most important function of mineralocorticoids is control of:
- **a.** Carbohydrate metabolism
- **b.** Glucose metabolism
- **c.** Electrolyte metabolism
- **d.** Protein metabolism

3. The pancreas has four types of cells, each of which produces a specific hormone. For example, the α cells of the pancreas produce:
- **a.** Insulin
- **b.** Glucagon
- **c.** Somatostatin
- **d.** Pancreatic polypeptide

4. Two hormones play an important role in calcium homeostasis. The two hormones, _____ and _____, cause an increase and a decrease in calcium concentrations, respectively:
- **a.** Calcitonin; glucagon
- **b.** Somatostatin; calcitonin
- **c.** Calcitonin; parathyroid hormone
- **d.** Parathyroid hormone; calcitonin
- **e.** Parathyroid hormone; glucagon

5. The main functions of the catecholamines are to allow rapid body responses to acute stimuli, which include the mobilization of glucose. The catecholamines are secreted by the sympathetic portion of the autonomic nervous system. The main neurotransmitter of the sympathetic nervous system is _____, whereas _____ is the main hormone produced by the postganglionic cells of the adrenal medulla.
- **a.** Serotonin; epinephrine
- **b.** Epinephrine; serotonin
- **c.** Epinephrine; norepinephrine
- **d.** Norepinephrine; epinephrine
- **e.** Serotonin; melatonin

Bibliography

Court MH, Freeman LM. Identification and concentration of soy isoflavones in commercial cat foods. *Am J Vet Res*. 2002;63(2):181–185.

Doerge DR, Sheehan DM. Goitrogenic and estrogenic activity of soy isoflavones. *Environ Health Perspect*. 2002;110(suppl 3):349–353.

Dye JA, Venier M, Ward CA. Brominated-flame retardants (BFRs) in cats—possible linkage to feline hyperthyroidism (abstract). *J Vet Intern Med*. 2007;21:595.

Dye JA, Venier M, Zhu L, et al. Elevated PBDE levels in pet cats: sentinels for humans? *Environ Sci Technol*. 2007;41(18):6350–6356.

Edinboro CH, Scott-Moncrieff JC, Janovitz E, et al. Epidemiologic study of relationship between consumption of commercial canned food and risk of hyperthyroidism in cats. *J Am Vet Med Assoc*. 2004;224(6):879–886.

Eiler H. Endocrine glands. In: Reece WO, eds. *Dukes' Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Engelking LR. *Metabolic and Endocrine Physiology*. 3rd ed. Jackson, Wyo: Teton New Media; 2012.

Feldman EC, Nelson RW, Reusch CE, et al, eds. *Canine and Feline Endocrinology*. 4th ed. Philadelphia: Saunders; 2015.

Fernandez NJ, Barton J, Spotswood T. Hypoglycemia in a dog. *Can Vet J*. 2009;50(4):423–426.

Kass PH, Peterson ME, Levy J, et al. Evaluation of environmental, nutritional, and host factors for feline hyperthyroidism. *J Vet Intern Med*. 1999;13(4):323–329.

Martin KM, Rossing MA, Ryland LM. Evaluation of dietary and environmental risk factors for feline hyperthyroidism. *J Am Vet Med Assoc*. 2000;217(6):853–856.

Melmed S, Polonsky KS, Reed PR, et al. *Williams Textbook of Endocrinology*. 13th ed. Philadelphia: Saunders Elsevier; 2016.

Molina P. *Endocrine Physiology*. 4th ed. New York: McGraw-Hill Lange; 2013.

Mooney CT. Pathogenesis of feline hyperthyroidism. *J Feline Med Surg*. 2002;4(3):167–169.

Peterson M. Animal models of disease: feline hyperthyroidism, an animal model for toxic nodular goiter. *J Endocrinol*. 2014;223:97–114.

Peterson ME, Kintzer PP, Cavanaugh PG, et al. Feline hyperthyroidism: pretreatment clinical and laboratory evaluation of 131 cases. *J Am Vet Med Assoc*. 1983;183(1):103–110.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames, Iowa: Iowa State University Press; 2003.

Schoeman JP. Insulin secreting tumors. In: Ettinger SJ, Feldman EC, eds. *Textbook of Veterinary Internal Medicine*. 8th ed. St Louis: Elsevier Saunders; 2017.

Stephens MJ, O'Neill DG, Church DB, et al. Feline hyperthyroidism reported in primary-care veterinary practices in England: prevalence, associated factors and spatial distribution. *Vet Rec*. 2014;175:458.

35

Reproductive System Development and Sexual Differentiation

AUTUMN P. DAVIDSON AND GEORGE H. STABENFELDT†

KEY POINTS

Development of the Reproductive System

1. Organization of the gonads is under genetic control (genetic sexual differentiation).
2. Sexual organization of the genitalia and brain depends on the presence or absence of testosterone.

Hypothalamopituitary Control of Reproduction

1. The hypothalamus and anterior pituitary (adenohypophysis) secrete protein and peptide hormones, which control gonadal activity.
2. The adenohypophysis (*pars distalis*) produces follicle-stimulating hormone (FSH), luteinizing hormone (LH), and prolactin, all of which control reproductive processes.

Modification of Gonadotropin Release

1. The pulsatile release of gonadotropin-releasing hormone (GnRH) induces the critical pulsatile production of the gonadotropins, FSH and LH.
2. Gonadotropin release is then modulated by the process of negative feedback from estrogen and progesterone.

Ovarian Follicle Development

1. Gamete development occurs initially without gonadotropin support and subsequently with pulsatile gonadotropin secretion.
2. In the preantral follicle, gonadotropin receptors for LH develop on the theca, which results in androgen synthesis; FSH directs the granulosa to transform the androgens to estrogens.
3. Late in the ovarian follicular phase, LH receptors develop on the granulosa, which permits the preovulatory surge of LH to cause ovulation.

Development of the Reproductive System

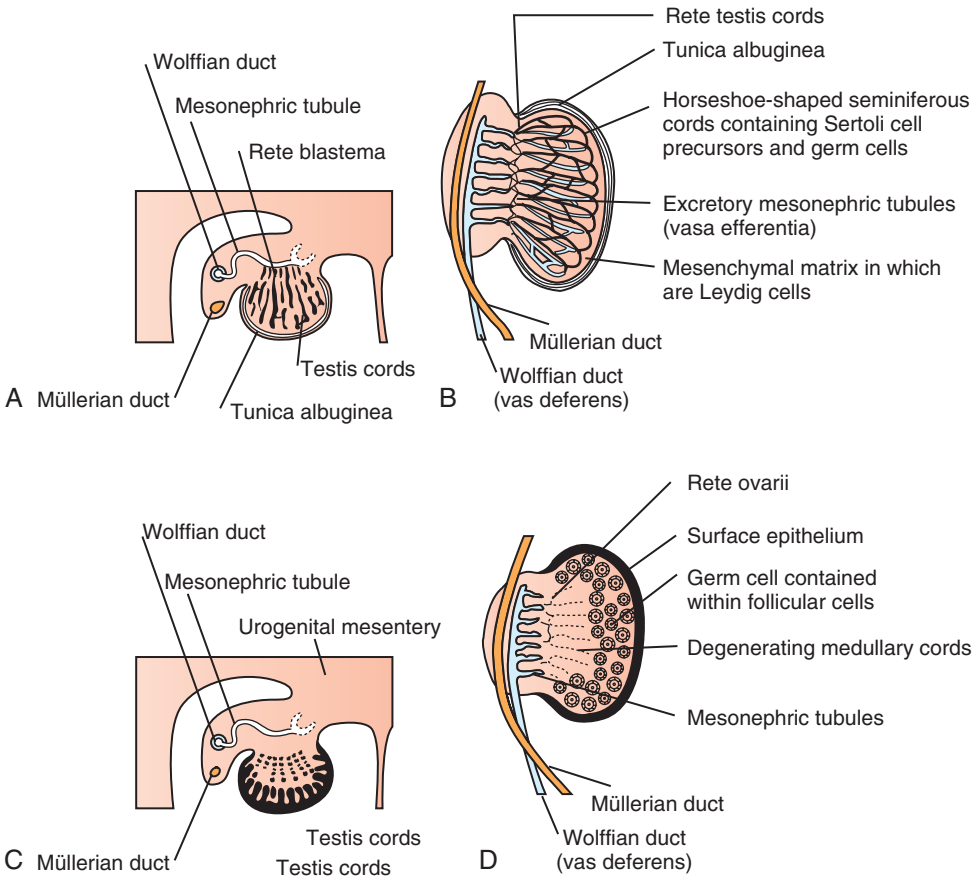
Organization of the Gonads Is Under Genetic Control (Genetic Sexual Differentiation)

The initial development of the embryonic ovary involves the migration of germ cells into the genital ridge from the yolk sac. These primordial germ cells populate sex cords that have formed in the cortical region of the embryonic gonad from the proliferation of cells from the *coelomic epithelium* (so-called germinal epithelium) of the genital ridge. The sex cords contribute cells, known initially as *follicle cells* and subsequently as *granulosa cells,* which immediately surround the oocyte. The mesenchyme of the genital ridge contributes cells that will become the theca. The entire structure is called a *follicle,* which includes oocyte, granulosa, and theca cells.

No direct connections are formed between the oocytes and the tubes destined to become the oviducts, which are derived from *müllerian ducts.* The final result is that oocytes are released through the surface of the ovary by rupture of tissue elements that surround the ovary; this process is called *ovulation.* A specialized end of the oviduct, the *fimbria,* develops to enable the oocyte to be removed efficiently from the surface of the ovary. In some animals, oocytes are funneled to the fimbria through the use of a bursa, which tends to encompass the ovary; oocytes are directed to a relatively small opening in the bursa.

The development of the embryonic testis is similar to that of the ovary; germ cells migrate into the genital ridge and populate sex cords that have formed from an invagination of the surface (coelomic) epithelium (Fig. 35.1). *Sertoli cells* (male counterparts of granulosa cells) develop from the sex cords, and *Leydig cells*

†Deceased

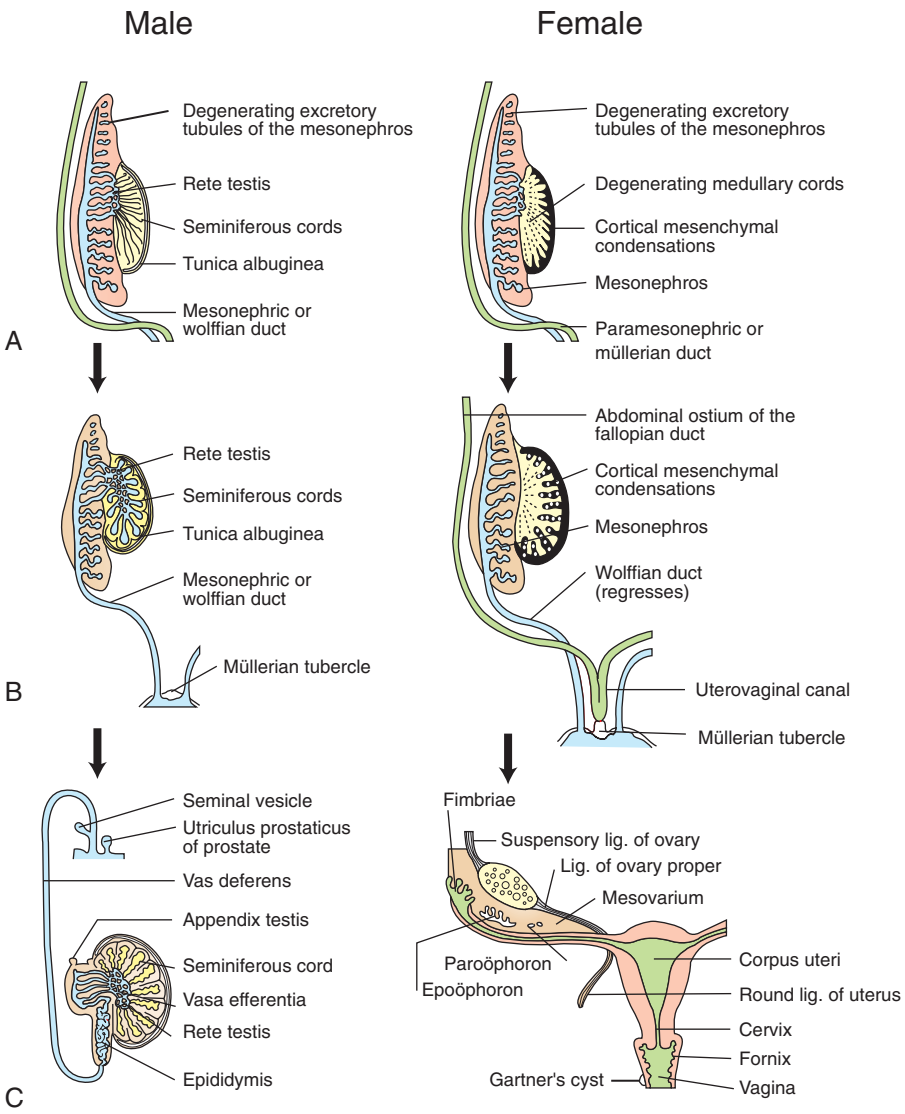


• **Fig. 35.1** Testicular development during the eighth week (A) and the sixteenth to twentieth weeks (B) of human fetal life. (A) The primitive sex cords proliferate in the medulla and establish contact with the rete testis. The tunica albuginea (fibrous connective tissue) separates the testis cords from the coelomic epithelium and eventually forms the capsule of the testis. (B) Note the horseshoe shape of the seminiferous cords and their continuity with the rete testis cords. The vasa efferentia, derived from the excretory mesonephric tubules, connect the seminiferous cords with the wolffian duct (see text). Comparable diagrams of ovarian development around the seventh week (C) and the twentieth to twenty-fourth weeks (D) of development. (C) Any primitive medullary sex cords degenerate and are replaced by the well-vascularized ovarian stroma. The cortex proliferates, and mesenchymal condensations later develop around the arriving primordial germ cells. (D) In the absence of medullary cords and a true persistent rete ovarii, no communication is established with the mesonephric tubules. Therefore, in the adult, ova are shed from the surface of the ovary and are not transported by tubules to the oviduct. (From Johnson M, Everitt B, ed. *Essential Reproduction*. 3rd ed. London: Blackwell Scientific; 1988.)

(male counterparts of thecal cells) develop from the mesenchyme of the genital ridge. One fundamental difference from ovarian development is that the invagination of the sex cords in the male continues into the medulla of the embryonic gonad, where connections are made with medullary cords from the mesonephros (primitive kidney). The duct of the mesonephros (*wolffian duct*) becomes the epididymis, vas deferens, and urethra, which has a direct connection to the seminiferous tubules. Thus male germ cells pass to the exterior of the animal through a closed tubular system.

Sexual Orientation of the Genitalia and Brain Depends on the Presence or Absence of Testosterone

The development of the genital tubular system and the external genitalia (*genital sexual differentiation*) is under the control of the developing gonad. If the individual is female—that is, the developing gonad is an ovary—the müllerian duct develops into the oviduct, uterus, cervix, and vagina, whereas the wolffian duct regresses; the absence of testosterone is important for both changes (Fig. 35.2). If the individual is male, the rete testis produces a müllerian-inhibiting factor, anti-müllerian hormone (AMH), which causes regression of the müllerian ducts. The *wolffian duct* is maintained in the male because of the influence of androgens produced by the testis. AMH is a hormone best known for its production by fetal testes in mammals and as the inhibitor of müllerian (paramesonephric) duct development in males. However, following the development of the müllerian ducts into the oviduct, uterus, and upper vagina in female mammals, the ovaries also produce AMH, which can be found in measureable amounts within the peripheral circulation, especially in adults following puberty. The ovaries appear to be the sole source of AMH in the circulation postpuberty; therefore it is a useful marker in clinically relevant situations when



• **Fig. 35.2** Differentiation of the internal genitalia in the human male and female at (A) the sixth week of gestation, (B) the fourth month of gestation, and (C) the time of descent of testis and ovary. Note that the müllerian and wolffian ducts are present in both sexes early on; the müllerian ducts eventually regress in the male and persist in the female, and the wolffian ducts regress in the female and persist in the male. The appendix testis and utriculus prostaticus in the male and epoöphoron, paroöphoron, and Gartner's cyst in the female are remnants of the degenerated müllerian and wolffian ducts, respectively. *Lig,* Ligament. (From Johnson M, Everitt B, ed. *Essential Reproduction*. 3rd ed. London: Blackwell Scientific; 1988.)

an assessment of the presence or absence of ovaries or ovarian remnants in dogs and cats is important.

The presence of an enzyme, *5α-reductase,* is important for the effect of the androgens because testosterone must be converted intracellularly into dihydrotestosterone for masculinization of the tissues to occur. The use of synthetic 5α-reductase inhibitors for the treatment of benign prostatic disease in humans is contraindicated without concurrent birth control measures, because drug levels in semen deposited repeatedly in the female can lead to disorders of sexual development in male fetuses.

Development of the external genitalia follows the development and direction of the gonads. If the individual's genotype is female, folds of tissue called *labia* form the *vulva,* and a *clitoris* develops. If the individual is male, androgens from the testis direct formation of the *penis* (male counterpart of the clitoris) and the *scrotum* (male counterpart of the labia). Again, the absence or presence of androgens is an important factor influencing the formation of external genitalia.

The final organization of the individual with regard to sex comes with sexual differentiation of the hypothalamus. Exposure of the hypothalamus to androgens at about the time of birth causes the hypothalamus to be organized as male. A paradoxical finding is that conversion (*aromatization*) of androgens to estrogens is essential for maleness, mediated by enzymes in the neural tissue. In the absence of androgens, the *hypothalamus* is organized as female.

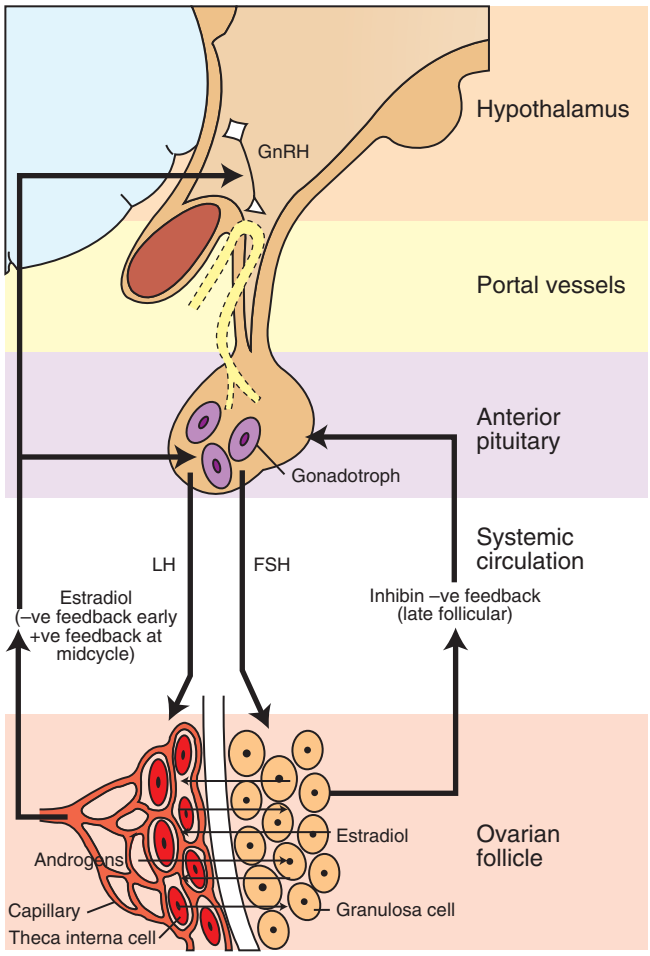
The fundamental concept of organization of the reproductive system with regard to genotype is that the female system is organized in the absence of testes. If the individual is to be male, there must be active intervention by the testes through the production of

androgens and appropriate tissue enzymes in two circumstances: (1) within the internal genitalia for conversion to more potent androgens and (2) within the hypothalamus for conversion to estrogens.

Hypothalamopituitary Control of Reproduction

The Hypothalamus and Anterior Pituitary (Adenohypophysis) Secrete Protein and Peptide Hormones, Which Control Gonadal Activity

Gonadal activity is under the control of both the *hypothalamus* and the *anterior pituitary gland* (Fig. 35.3). The hypothalamus lies near the ventral midline of the diencephalon. It is divided into halves by the third ventricle and actually forms the ventral and lateral walls of the third ventricle. The hypothalamus has clusters of neurons, collectively called *nuclei,* which secrete peptide hormones important for controlling pituitary activity. As described in more detail later, these peptides move to the pituitary either directly by passage through the axons of neurons or by a vascular portal system.



• **Fig. 35.3** Summary of hypothalamic-pituitary-ovarian interactions during the follicular phase of the cycle. *FSH,* Follicle-stimulating hormone; *GnRH,* gonadotropin-releasing hormone; *LH,* luteinizing hormone; *–ve,* negative; +ve, positive. (From Johnson M, Everitt B, ed. *Essential Reproduction*. 3rd ed. London: Blackwell Scientific; 1988.)

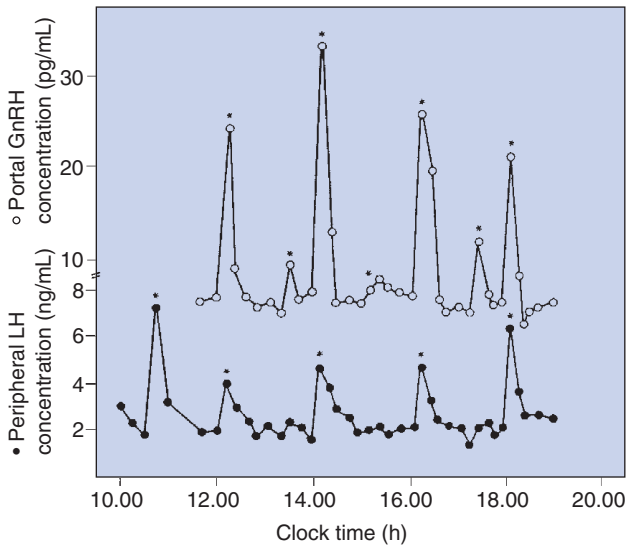
The pituitary responds to the hypothalamic peptides to produce hormones important for the control of the gonads.

The Adenohypophysis (*Pars Distalis*) Produces Follicle-Stimulating Hormone, Luteinizing Hormone, and Prolactin, All of Which Control Reproductive Processes

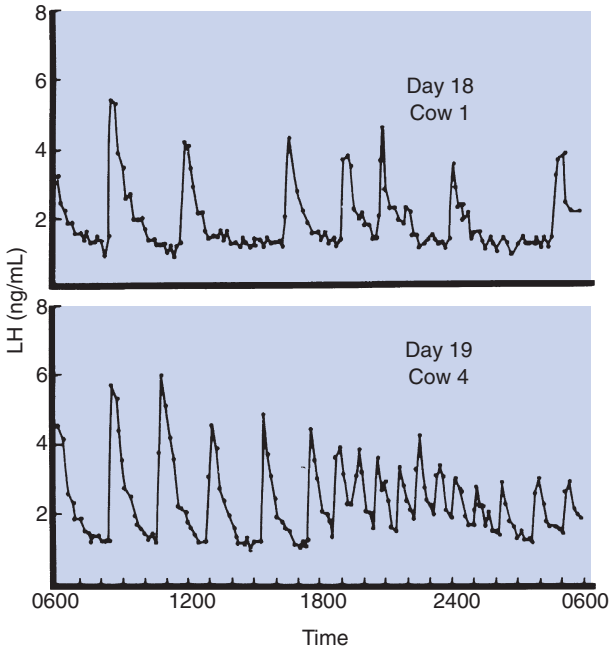
The pituitary gland is composed of three parts: an anterior lobe called the *adenohypophysis,* or *pars distalis;* an intermediate lobe called the *pars intermedia;* and a posterior lobe called the *neurohypophysis,* or *pars nervosa.* The lobes are of different embryological origins; the *pars distalis* is derived from the surface ectoderm (derived in turn from a small diverticulum off the dorsal pharynx, called *Rathke's pouch*), and the *pars intermedia* and *pars nervosa* are derived from neuroectoderm. The adenohypophysis produces protein hormones important for the control of reproduction: two *gonadotropins* (follicle-stimulating hormone [FSH] and luteinizing hormone [LH]) and a third hormone called *prolactin;* other pituitary hormones include *growth hormone* (GH), *corticotropin* (adrenocorticotropic hormone, ACTH), and *thyroid-stimulating hormone* (TSH). FSH and LH are synergistic in *folliculogenesis* and ovulation in the ovary. FSH plays a more dominant role during the growth of follicles, and LH plays a more dominant role during the final stages of follicle maturation through ovulation. The *gonadotropins,* as well as TSH, are called *glycoproteins* because their molecules contain carbohydrate moieties that contribute to their function. *Oxytocin,* which is released by the neurohypophysis, is a hormone of importance in reproduction.

Besides being an important center for the control of reproduction, the *hypothalamus* regulates appetite and temperature, and integrates the activity of the autonomic nervous system. Because of a common embryological origin, the *hypothalamus* has a direct connection to the neurohypophysis. This connection is through the neural stalk, which contains axons that originate from neuronal cell bodies located in the hypothalamus. Two sets of neurons within the hypothalamus, the *supraoptic* and *paraventricular* nuclei, are responsible for the synthesis of vasopressin and oxytocin, respectively. These small peptide hormones are coupled to larger peptide molecules, called *neurophysins,* and are transported from the site of synthesis in the hypothalamus (neuronal cell bodies) through axons to the site of storage and eventual release, the neurohypophysis.

The connection of the hypothalamus to the adenohypophysis does not involve the direct passage of axons through the *neural stalk.* A *venous portal system* connects the median eminence within the hypothalamus to the adenohypophysis. Hypothalamic substances that control the adenohypophysis are carried from the median eminence of the hypothalamus to the pituitary by a venous portal system. For example, gonadotropin-releasing hormone (GnRH), a peptide, is produced in the medial preoptic nucleus, and *dopamine,* an organic chemical of the catecholamine and phenethylamine families, is produced in the arcuate nucleus. Axons transport both substances from the hypothalamus to the *median eminence,* where they are released into the venous portal system. The synthesis of GnRH, as with oxytocin and vasopressin, involves the production of a larger precursor molecule with a C-terminal region of 56 amino acids called *GnRH-associated peptide* (GAP). Although GAP can stimulate the release of FSH and LH, GnRH is still thought to be the critical hormone for gonadotropin release. An even more important function of GAP may be its ability to inhibit prolactin secretion.



• **Fig. 35.4** Concentrations of gonadotropin-releasing hormone (*GnRH*) in portal plasma (*open circles*) and luteinizing hormone (*LH*) in jugular venous plasma (*solid circles*) of four ovariectomized ewes. *Asterisks* indicate secretory episodes (pulses) of GnRH and LH. (From Johnson M, Everitt B, ed. *Essential Reproduction*. 3rd ed. London: Blackwell Scientific; 1988.)



• **Fig. 35.5** Pattern of plasma luteinizing hormone (*LH*) concentration on day 18 or 19 of the estrous cycle in two cows. (From Rahe CH, Owens RE, Fleeger JL, et al. Pattern of plasma luteinizing hormone in the cyclic cow: dependence upon the period of the cycle. *Endocrinology*. 1980; 107(2):498–503.)

Modification of Gonadotropin Release

The Pulsatile Release of Gonadotropin-Releasing Hormone Induces the Critical Pulsatile Production of the Gonadotropins Follicle-Stimulating Hormone and Luteinizing Hormone

The main secretory pattern of gonadotropins is *pulsatile;* the pattern is driven by pulsatile secretion of GnRH from the hypothalamus (Fig. 35.4). The pulsatile release of GnRH induces the critical pulsatile production of the gonadotropins, FSH and LH.

Gonadotropin Release Is Then Modulated by the Process of Negative Feedback From Estrogen and Progesterone

Gonadotropin release is then modulated by the process of negative feedback from estrogen and progesterone. The importance of this mode of delivery is shown by the fact that if GnRH is administered in a continuous (pharmacological) manner, the system can be downregulated. Continual occupancy of GnRH receptors on gonadotrophs by GnRH interrupts the intracellular signal for the synthesis and release of gonadotropins. Successful induction of a fertile estrus in bitches can be performed by administering canine analogue GnRH; however, the dose must be diminished as ovulation approaches or downregulation will occur.

In general, *the pulse generator system* for gonadotropin secretion is increased in the *follicular phase* and decreased in the *luteal phase* of the estrous cycle (Fig. 35.5). Estrogen decreases the pulse amplitude, and progesterone decreases the pulse frequency of *gonadotropin* secretion. This means that during the follicular phase, pulse frequency increases because of the absence of progesterone, and pulse amplitude decreases because of the presence of estrogen. This combination of increased pulse frequency and decreased pulse amplitude is important for nurturing the final growth phase of the developing antral follicle.

The hypothalamus and adenohypophysis are capable of responding to a sustained increase in estrogen secretion by increased secretion of gonadotropins, a relationship termed *positive feedback.* The sudden sustained increase in estrogen levels, which occurs over one to several days during final antral follicle development, causes an increase in gonadotropin secretion by increasing the frequency of pulsatile release of GnRH and, as a result, gonadotropin secretion. In essence, the frequency of pulsatile release of gonadotropins overcomes the metabolic clearance rate. The purpose of the gonadotropin surge is to induce changes within the follicle that lead to its rupture (ovulation). The duration of the gonadotropin surge is relatively short, usually 12 to 24 hours, possibly because the main factor driving the response, estrogen, declines in concentration as the follicles respond to the preovulatory gonadotropin surge. This particular physiological mechanism for initiating the onset of ovulation is effective because the follicle is able to signal its stage of maturity to the hypothalamus and adenohypophysis by a product (estrogen) produced in increasing amounts with increasing follicle maturity.

The secretion of gonadotropins is modified by the ovarian steroid hormones *estrogen* and *progesterone.* With time, the effect of these hormones is suppressive for gonadotropin secretion. Estrogens, in particular, cause *negative-feedback inhibition* of gonadotropin secretion, which is characterized by its sensitivity (effective at low concentrations) and its rapid onset (within a few hours). The substantial increase in gonadotropin concentrations that occurs after ovariectomy is caused largely by the removal of estrogens.

Because progesterone affects gonadotropin pulse frequency, it is thought that its modulatory effect is at the level of the hypothalamus. Estrogens are thought to affect gonadotropin secretion through an effect on both the pituitary gland and the hypothalamus.

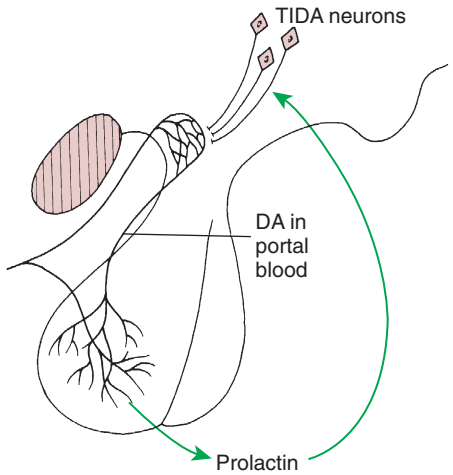
Although there are differences in the site of action among species, it appears that the hypothalamic site for negative-feedback inhibition of gonadotropins by both progesterone and estrogen is in an area immediately above the median eminence, known as the *arcuate nucleus.* The hypothalamic site for positive-feedback stimulation of gonadotropin release by estrogen is probably further anterior, that is, in the preoptic anterior hypothalamic region.

The secretion of gonadotropins can be modified by peptide and protein hormones produced by both the hypothalamus and the ovary. β-Endorphin, an opioid peptide produced from the hypothalamic precursor molecule pro-opiomelanocortin, can inhibit LH secretion when pharmacologically administered systemically. Its role in the physiological modulation of gonadotropin secretion, however, remains to be identified. Another hormone, inhibin, a protein produced by the granulosa cells of the developing follicle, also inhibits gonadotropin secretion, particularly FSH, during the final stages of follicle development (see Fig. 35.3). As described in the section on *folliculogenesis,* this depression of FSH secretion may be important to the animal for controlling the number of follicles that are brought to final maturation.

Control of gonadotropin secretion in the male is similar to that in the female; pulses of GnRH, arising in the hypothalamus, affect pulsatile secretion of the gonadotropins. This in turn causes the secretion of testosterone, also in pulsatile form, from the testes. One major difference between the sexes is that the need for positive-feedback release of gonadotropins in males does not exist; gametes are produced and released on a continuous basis within a tubular system that opens to the exterior. This negates any need for a surge release of gonadotropins as is required in the female to rupture the ovarian surface for the release of oocytes.

Prolactin is the third adenohypophysis-produced hormone that is important in the reproductive process, mainly because of its effect on the mammary gland and lactation in mammals. Although the secretion of prolactin is pulsatile, the control of secretion has more emphasis on inhibition than does stimulation of secretion. This concept is supported by the finding that prolactin secretion increases if the pituitary gland is disconnected from the hypothalamus by either cutting the pituitary stalk or transplanting the pituitary gland to another site (e.g., kidney capsule). Thus most attention has focused on factors that inhibit prolactin secretion. The catecholamine dopamine, which is produced by neurons in the ventral hypothalamus (arcuate nucleus), is a potent inhibitor of prolactin secretion (Fig. 35.6). Other factors that inhibit prolactin secretion are γ-aminobutyric acid (GABA) and GAP. Dopamine agonists, such as the ergot-type compounds *bromocriptine* and *cabergoline,* can be used to suppress prolactin secretion in cases of hyperprolactinemia. Cabergoline, a potent prolactin inhibitor, can be used to shorten interestrous intervals in female dogs and promote luteolysis in female dogs and female cats during the latter half of pregnancy (prolactin is a luteotropin). The negative-feedback control of prolactin is shown in Fig. 35.6.

One of the first known prolactin-releasing factors was *thyrotropin-releasing hormone* (TRH). The physiological relevance of TRH in prolactin secretion is still unknown despite receptors for TRH have been identified on lactotropes within the adenohypophysis. *Vasoactive intestinal peptide* (VIP), a potent stimulator of prolactin secretion, may play a physiological role in prolactin secretion through inhibition of dopamine synthesis within the hypothalamus. Estrogens can increase prolactin secretion by lactotropes by decreasing lactotrope sensitivity to dopamine and increasing the number of TRH receptors. Interestingly, female dogs undergoing ovariohysterectomy with a cesarean section usually maintain the ability to lactate effectively afterward, but lactation should always be evaluated before ovariohysterectomy or ovariectomy is performed. Removal of the ovary, the primary source of estrogen, could be detrimental if lactation is marginal. Profoundly hypothyroid dogs can have mammary development, gynecomastia, secondary to elevated TRH levels.



• **Fig. 35.6** Diagrammatic summary of the proposed negative-feedback relationship between prolactin and dopamine (*DA*). Prolactin is believed to accelerate dopamine turnover in the arcuate nucleus neurons (tuberoinfundibular dopamine [*TIDA*] neurons), and the amine is then released into the portal capillaries to gain access to the lactotropes. Hyperprolactinemia could be caused by either a failure of prolactin inhibitory factor activity at the dopamine receptor level in the anterior pituitary or a reduction of TIDA neuron activity in the hypothalamus. (From Johnson M, Everitt B, ed. *Essential Reproduction*. 3rd ed. London: Blackwell Scientific; 1988.)

Ovarian Follicle Development

Gamete Development Occurs Initially Without Gonadotropin Support and Subsequently With Pulsatile Gonadotropin Secretion

Oocyte proliferation, which occurs by *mitotic division* during fetal development, ends at about the time of birth in most mammalian species. Oocytes begin the process of reduction of chromosome numbers to the haploid state by *meiosis* shortly after birth under the influence of *meiosis-initiating factor,* thought to be produced by the *rete ovarii.* The process is soon interrupted at the *diplotene,* or *dictyate,* stage of meiosis I by the meiosis-inhibiting factor, which is probably produced by the developing follicle cells. Oocytes remain in this stage until the follicle begins its final development, an interval that can be as long as 50 years or more in humans. The *follicle,* at this point, is delineated by an outer basement membrane (*membrana propria*), which is secreted by the follicle cells.

The initial development of the follicle involves growth of the oocyte. This growth is accompanied by intense synthetic activity; a large amount of ribonucleic acid (RNA) is synthesized. At the same time, follicle cells begin to divide and form a *granulosa* several cells thick. Another boundary substance is secreted by both the oocyte and the ovarian follicles, the *zona pellucida,* that lies within the granulosa and that immediately surrounds the oocyte. Granulosa cells maintain contact with the oocyte through the zona pellucida

by means of the development of cytoplasmic processes. Interaction among granulosa cells is facilitated by the development of gap junctions. This form of communication is important because the granulosa has no blood supply; blood vessels are excluded at the level of the membrana propria. The *thecal layer* of the follicle forms around the membrana propria to complete the layers of the follicle. Follicles at this stage are called *primary,* or *preantral, follicles.*

Factors that control initial follicle growth are not known. External factors, such as gonadotropins, are not required because preantral follicles can develop in hypophysectomized animals. In species such as cattle and horses (perhaps sheep and goats as well) in which several dominant follicles develop during the estrous cycle, it is likely that a few follicles begin to develop each day. In animals in which a cohort of follicles develops synchronously (pigs, cats, dogs), there appears to be less tendency to have competing follicle growth waves during the luteal phase (pig) and a tendency to have only one cohort of follicles during the preovulatory period (cat and dog). Thus the development of a cohort of follicles may limit follicle development from the primordial state, at least during the period of active follicle development leading to ovulation. Initial follicle growth is under genetic control, and the pattern reflects the needs of the particular species.

In the Preantral Follicle, Gonadotropin Receptors for Luteinizing Hormone Develop on the Theca, Which Results in Androgen Synthesis; Follicle-Stimulating Hormone Directs the Granulosa to Transform the Androgens to Estrogens

For follicles to progress beyond the preantral stage, the granulosa and theca need to develop *receptors* for gonadotropins. FSH and LH receptors develop on the granulosa and theca, respectively. The onset of the antral follicle is marked by the appearance of fluid that begins to divide the granulosa. The follicular fluid, a secretory product of the granulosa, coalesces to form an increasingly larger fluid cavity (*antrum*) within the granulosa. In later development of the antral follicle, the oocyte remains surrounded by a layer of granulosa cells called the *cumulus oophorus,* which are attached to the wall of the follicle by a small stalk of granulosa cells.

The proximity of the granulosa and theca cells allows cooperative estrogen synthesis. The theca produces *androgens* (testosterone and androstenedione) under the influence of LH, which diffuses across the membrana propria into the granulosa, where the androgens are enzymatically transformed into *estrogen* (estradiol-17) (see Fig. 35.3). At this time of development, the granulosa is incapable of forming androgens, the precursors of estrogen biosynthesis, and the theca has limited capacity for producing estrogens. This concept of cooperative effort, called the *two-cell mechanism* for estrogen secretion, is generally accepted as being the way most follicular estrogen is produced. These estrogens have a positive-feedback effect on the granulosa; they stimulate the cells to undergo mitotic division, and thus the follicle grows in size as the granulosa proliferates in response to its own secretory product (estrogen).

One effect of estrogen is the formation of additional receptors for FSH as follicle development proceeds. In this situation the antral follicle becomes increasingly sensitive to FSH as it develops and is able to grow under a relatively steady state of FSH secretion.

Late in the Ovarian Follicular Phase, Luteinizing Hormone Receptors Develop on the Granulosa, Which Permit the Preovulatory Surge of Luteinizing Hormone to Cause Ovulation

Late in antral follicle development, FSH and estrogens initiate the formation of LH receptors on the granulosa, whereas FSH receptors begin to diminish. Increasing secretion of estrogen by the antral follicle finally results in the initiation of the *preovulatory surge of gonadotropins.* Thus in the last stages of development, the follicle falls progressively under the control of LH as it makes its last growth spurt to the point of ovulation.

CLINICAL CORRELATIONS

Androgen Insensitivity

History

You are called to examine a mare that has recently been brought to a broodmare farm after a successful racing career. It is late spring, but the mare has shown estrous behavior only on an intermittent basis.

Clinical Examination

As you approach the mare, you notice that she is large. The genital examination reveals a normal vulva, but when you introduce the speculum, it can be inserted only about 5 to 6 inches. Digital examination of the genital tract through the vulva results in a finding of complete blockage at the level of the vestibulovaginal conjunction, with no evidence of the external os of the cervix. On examination per rectum, you find the vagina, cervix, uterus, and oviducts to be absent; the gonads are symmetric in shape without the usual indentation caused by the ovulation fossa characteristic of equine ovaries.

Comment

You tell the shocked owner that you are suspicious that the animal is not really a mare but a male masquerading as a female. One of the easiest ways to confirm the diagnosis is to have a testosterone analysis done on plasma. If the gonads are testes, they still retain the ability to secrete significant amounts of testosterone, even though they are retained (cryptorchid, in a sense) within the abdominal cavity. You could also have a chromosomal analysis to verify that the animal has an XY sex chromosome complement. In this case, it is likely that the testes were able to secrete the müllerian-inhibiting factor, which resulted in regression of the tubular system of the genital tract that forms the female system (oviducts, uterus, cervix, vagina). But why, asks the owner, did the external genitalia not turn out to be male? There is evidence, in cases such as this, that the tissues of the external genitalia lacked critical receptors for androgens; thus the external genitalia were female in type. The rule of sexual development is that the female state develops in the absence of testicular input, the latter including müllerian-inhibiting factor and testosterone. In this case, the lack of sexual differentiation also appeared to involve the hypothalamus, because the "mare" did not exhibit male behavior despite relatively high testosterone concentrations.

Treatment

There is obviously no treatment for this syndrome. It would be unethical to take her back to the track and race her again as a female when the owner knows that "she" is a male. The horse could be used as a performance horse (i.e., hunter/jumper, dressage, eventing) or for pleasure riding.

CLINICAL CORRELATIONS—cont'd

Ovarian Remnant Syndrome

History

Your client acquired a dog from an animal rescue association 2 years ago, who was reportedly ovariohysterectomized (spayed) prior to placement. However, the dog is attractive to intact male dogs at the park and periodically has a prominent vulva.

Clinical Examination

The dog is normal on physical examination, with no enlargement of the external genitalia at the time. A vaginal cytology is taken and shows no estrogen influence on the vaginal epithelium currently. A urinalysis is performed and is normal. You perform an anti-müllerian hormone (AMH) test, which comes back positive, indicating ovarian tissue is present.

Comment

One important consideration is whether the dog could be exposed to estrogen or progesterone exogenously, most commonly from transdermal human hormone replacement therapeutic creams. This would most likely cause continuous signs of inappropriate hormone presence, rather than periodic signs.

Treatment

An abdominal ultrasound is performed, and an oval structure is noted in the region of the right kidney that appears to be an ovary. An exploratory surgery is performed, and the structure is removed and confirmed with biopsy to be an ovarian remnant. The opposite ovarian site is also resected and is free of ovarian tissue. The periodic signs are no longer occurring.

Practice Questions

1. Which of the following statements is *true*?
- **a.** Müllerian ducts develop in the female because of the presence of estrogen.
- **b.** Müllerian ducts develop in the female because of a müllerian-stimulating factor.
- **c.** Wolffian ducts develop in the male because of a wolffian-stimulating factor.
- **d.** Wolffian ducts develop in the male because of the presence of androgen.

2. The most potent factor involved in the organization of the internal and external parts of the genital tract is:
- **a.** Müllerian-inhibiting factor
- **b.** Müllerian-stimulating factor
- **c.** Estrogen
- **d.** Androgen

3. Which of the following groups of hormones is transported to the anterior pituitary by the hypothalamohypophyseal portal system?
- **a.** Oxytocin, GnRH, and dopamine
- **b.** GnRH, dopamine, and vasopressin
- **c.** Dopamine, vasopressin, and oxytocin
- **d.** Dopamine and GnRH

4. Which of the following groups of hormones controls the synthesis and release of hypophyseal hormones involved in reproductive processes?
- **a.** Oxytocin, GnRH, VIP, and dopamine
- **b.** GnRH, dopamine, VIP, and vasopressin
- **c.** Dopamine, vasopressin, VIP, and oxytocin
- **d.** GAP, dopamine, VIP, and GnRH
- **e.** GAP, GnRH, VIP, and oxytocin

5. Which of the following factors is responsible for causing oocytes to remain in a diplotene or dictyate state?
- **a.** Müllerian-inhibiting factor
- **b.** Müllerian-stimulating factor
- **c.** Meiosis-inhibiting factor
- **d.** Meiosis-stimulating factor
- **e.** Wolffian-inhibiting factor
- **f.** Wolffian-stimulating factor

Bibliography

Hafez ESE, Hafez B, eds. *Reproduction in Farm Animals*. 7th ed. Baltimore: Lippincott Williams & Wilkins; 2000.

Johnson MH. *Essential Reproduction*. 7th ed. London: Wiley-Blackwell; 2013.

Lyle SK. Disorders of sexual development in the dog and cat. *Theriogenology*. 2007;68(3):338–343.

Plant TM, Zeleznik AJ, eds. *Knobil and Neill's Physiology of Reproduction*. 4th ed. vols 1 and 2. London: Academic Press; 2015.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames: Iowa State University Press; 2003.

Place NJ, Hansen BS, Cheraskin JL, et al. Measurement of serum anti-Müllerian hormone concentration in female dogs and cats before and after ovariohysterectomy. *J Vet Diagn Invest*. 2011;23(3):524–527.

Romagnoli S, Schlafer DH. Disorders of sexual differentiation in puppies and kittens: a diagnostic and clinical approach. *Vet Clin North Am Small Anim Pract*. 2006;36(3):573–606.

36

The Ovarian Cycle

AUTUMN P. DAVIDSON AND GEORGE H. STABENFELDT†

KEY POINTS

Ovulation

1. Ovulatory follicles are selected at the onset of luteolysis (in large domestic animals).
2. Ovulation is caused by an estrogen-induced preovulatory surge of gonadotropins.

Corpus Luteum

1. The corpus luteum secretes progesterone, which is essential for pregnancy.
2. Luteinizing hormone is important for the maintenance of the corpus luteum.
3. Regression of the corpus luteum in nonpregnant large domestic animals is controlled by uterine secretion of prostaglandin $F_{2\alpha}$.
4. Changes in luteal life span in large domestic animals occur because of changes in prostaglandin $F_{2\alpha}$ synthesis by the uterus.

Ovarian Cycles

1. In spontaneously ovulating animals, ovarian cycles have two phases: follicular and luteal; animals that require copulation for ovulation can have only a follicular phase.
2. The luteal phase is modified by copulation in some species.

Ovulation

Ovulatory Follicles Are Selected at the Onset of Luteolysis (in Large Domestic Animals)

Until the advent of ultrasonography, it was difficult to identify growth patterns of follicles in domestic animals, especially those of follicles that develop during the luteal phase of the cycle. The concept that follicles do develop during the luteal phase was emphasized by the earlier work of Rajakowski, who described the midcycle follicle in the cow. With ultrasonography, it has been possible to define follicular growth and regression during the luteal phase of the cycle in the cow and mare. In cattle, the predominant pattern is for several dominant (large) antral follicles to develop sequentially during the cycle (Fig. 36.1). The follicular cycles are distinct to the extent that follicle regression usually begins (as indicated by follicle size) before the onset of the growth of the next follicle. The first dominant follicle regresses about midluteal phase, with a second dominant follicle beginning growth immediately. Whether the second dominant follicle is the ovulatory follicle, or whether a third develops, depends on the stage of the follicle at the time of regression of the *corpus luteum* (CL). If the second dominant follicle has begun to regress at the time of CL regression, a third follicle develops. Thus the selected ovulatory follicle is, by chance, the dominant follicle still in a developmental stage at the time that regression of the CL is initiated. The duration required for the development of the antral follicle to the point of ovulation has been estimated by various techniques to be about 10 days in domestic animals, perhaps slightly longer in some primates.

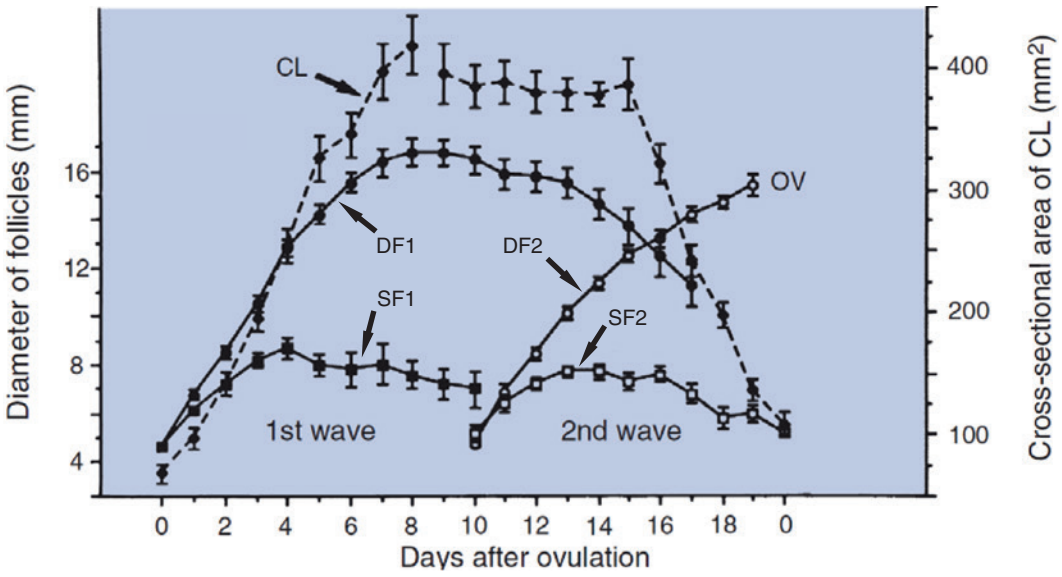
From ultrasonographical and endocrinological studies, two different phases in final antral follicle development apparently occur in large domestic animals: a relatively slow phase that lasts 4 to 5 days, followed by a second phase of accelerated growth, again lasting 4 to 5 days, that terminates in ovulation (Fig. 36.2). Because the final growth phase of follicle development can be initiated during the luteal phase, the initiation of this phase can occur under the influence of a relatively slow pulse rate of gonadotropin release that occurs during the luteal phase. The rapidly growing follicle requires exposure to a faster gonadotropin pulse rate by the third, or fourth, day in order for the follicle(s) to complete the normal growth pattern through ovulation. This situation usually occurs in conjunction with the onset of CL regression, which passively allows an increase in pulsatile rate of gonadotropin secretion (see Fig. 35.4).

One of the ways the *dominant follicle* maintains its status is to produce substances that inhibit the development of other antral follicles. One of the substances is *inhibin,* a peptide hormone produced by the granulosa, which inhibits the secretion of follicle-stimulating hormone (FSH). The dominant follicle is able to compensate for the lower FSH concentrations and continue to grow because of the numbers of FSH receptors it has compared with competitor follicles. Follicle development is dynamic once the rapid growth phase is achieved; the follicle(s) must be acted on through proper gonadotropin stimulation within a few days, or the result is death of the follicle. If the rapidly growing antral follicle is not exposed to the proper gonadotropin environment, *atresia* (regression) of the follicle begins almost immediately. Follicles that regress are invaded by inflammatory cells, and the area previously occupied by the antral follicle is eventually filled by connective tissue; that is, the follicle is replaced by an ovarian scar.

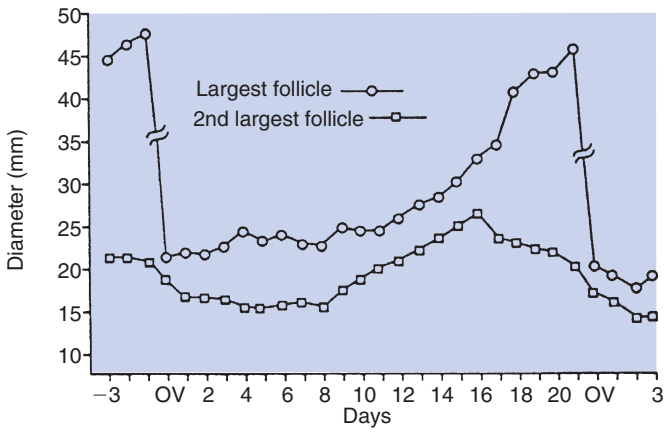
Ovulation Is Caused by an Estrogen-Induced Preovulatory Surge of Gonadotropins

The preovulatory surge of luteinizing hormone (LH), which begins about 24 hours before ovulation in most domestic species, including

†Deceased



• **Fig. 36.1** Mean (± standard error of the mean) profiles of diameters of dominant follicles and the largest subordinate follicle, and the cross-sectional luteinized area of the corpus luteum (*CL*), for interovulatory intervals with two follicular waves in 18 cattle. Regression (significant decrease in cross-sectional area, $P < 0.05$) of the CL began between days 15 and 16 for these two-wave intervals. Growth of subordinate follicles is suppressed by the respective dominant follicles. See text for further explanation. *DF1*, Dominant follicle of first wave; *DF2*, dominant follicle of second wave; *SF1*, subordinate follicle of first wave; *SF2*, subordinate follicle of second wave, *OV,* ovulation. (From Ginther OJ, Knopf L, Kastelic JP. Temporal associations among ovarian events in cattle during oestrous cycles with two and three follicular waves. *J Reprod Fertil*. 1989;87(1):223–230.)



• **Fig. 36.2** Development of the dominant and second-largest follicle during the estrous cycle of the mare. Note the divergence in diameter between the largest and second-largest follicles 1 day after ovulation. (From Pierson RA, Ginther OJ. Follicular population dynamics during the estrous cycle of the mare. *Anim Reprod Sci*. 1987;14:219.)

the cow, dog, goat, pig, and sheep, initiates the critical changes in the follicle that affect its endocrine organ status and result in release of the oocyte (Fig. 36.3). Two important tissues, the oocyte and the granulosa, have been kept under control by the production of inhibitory substances that are probably of granulosa origin. An *oocyte-inhibiting factor* prevents the oocyte from resuming meiosis, and a *luteinizing-inhibiting factor* prevents the granulosa from prematurely being changed into luteal tissue. The impact of the LH surge blocks the production of both these factors. In most animals the resumption of meiosis results in the first division of meiosis (meiosis I), or formation of the *first polar body,* which is complete before ovulation. In animals with the potential for reasonably long reproductive longevity (e.g., cattle), the initiation of the meiotic process could have begun as many as 10 years or more before its completion.

The effect of the LH surge on the granulosa is to allow initiation of the process of *luteinization,* which transforms the cells from estrogen to progesterone secretion. This process is underway before ovulation occurs. With the advent of the LH surge, estrogen secretion declines concomitantly with the onset of progesterone secretion.

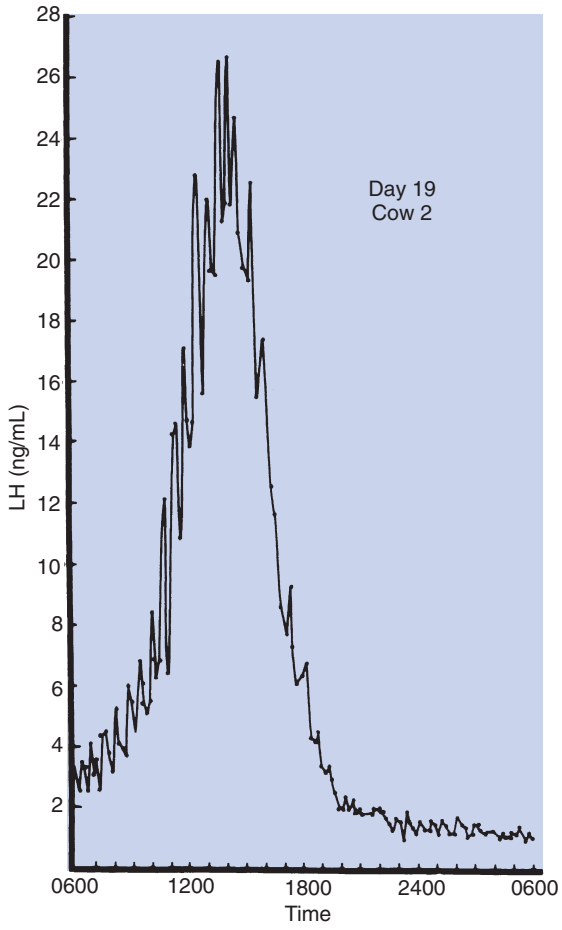
Another function of the preovulatory surge release of LH is to cause the granulosa to produce substances, such as relaxin and *prostaglandin* $F_{2\alpha}$ ($PGF_{2\alpha}$), that affect the continuity of the connective tissue of the thecal layers of the follicle. These and other unknown substances disrupt the theca through the development of vesicles (within fibrocytes) that contain hydrolytic enzymes capable of breaking down the collagen matrix of connective tissue. The rupture of the follicle is caused by the disintegration of the connective tissue.

In summary, estrogen is used by the follicle(s) (1) to stimulate the growth and development of the granulosa and (2) to signal the hypothalamus and anterior pituitary as to the readiness of the follicle(s) for ovulation.

Corpus Luteum

The Corpus Luteum Secretes Progesterone, Which Is Essential for Pregnancy

The main function of the CL is the secretion of progesterone, which prepares the uterus for the initiation and maintenance of pregnancy. The CL forms from the wall of the follicle, which is



• **Fig. 36.3** Preovulatory surge of luteinizing hormone (*LH*) on day 19 of the estrous cycle in a cow. (From Rahe CH, Owens RE, Fleeger JL, et al. Pattern of plasma luteinizing hormone in the cyclic cow: dependence upon the period of the cycle, *Endocrinology*. 1980;107(2):498–503.)

collapsed and folded after ovulation. With rupture of the follicle, there is a breakdown of the tissues that surround the granulosa, particularly the membrana propria, and hemorrhage into the cavity can occur from vessels in the theca. The folds of tissue that protrude inward into the cavity contain granulosa and theca cells and, importantly, the vascular system that will support cell growth and differentiation. Although the granulosa cell is the dominant cell of the CL, theca cells also contribute significantly to the composition of the structure. The process that granulosa cells undergo during the change from estrogen to progesterone secretion, *luteinization,* begins with the onset of the preovulatory LH surge and accelerates with ovulation.

In most domestic species, significant production of progesterone by the CL begins within 24 hours of ovulation. In some species, including the dog and primates, small amounts of progesterone are produced during the preovulatory LH surge; in the dog, this is important for the expression of sexual receptivity, which occurs as estrogen levels decline while progesterone levels increase.

Luteinizing Hormone Is Important for the Maintenance of the Corpus Luteum

For most domestic animals, LH is the important *luteotropin,* with the CL maintained in either nonpregnant or pregnant animals by a relatively slow pulsatile pattern of LH release (one pulse every 2 to 3 hours). In rodents, prolactin is the important luteotropin; daily biphasic release of prolactin is initiated by copulation, which is essential for the maintenance of the CL. Of the domestic species, prolactin has been implicated as a luteotropin in sheep and dogs.

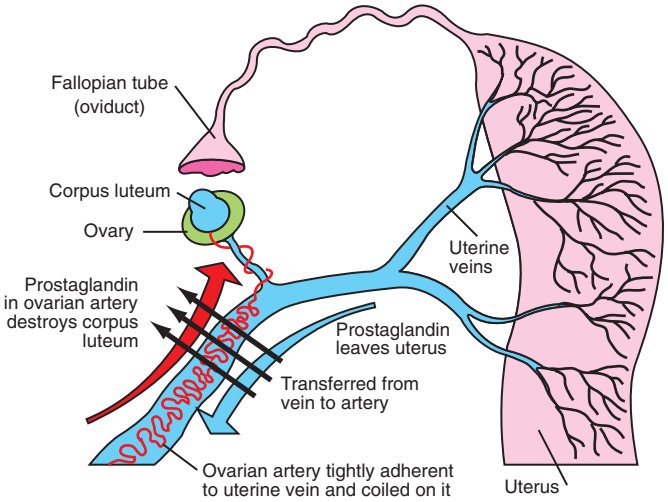
Normal folliculogenesis, a prerequisite for ovulation, sets the stage for the subsequent development of the postovulatory CL. However, more clinical attention is paid to factors controlling the regression of the CL than to luteotropic factors, especially when artificial estrus synchronization among animals is desired.

Regression of the Corpus Luteum in Nonpregnant Large Domestic Animals Is Controlled by Uterine Secretion of Prostaglandin $F_{2\alpha}$

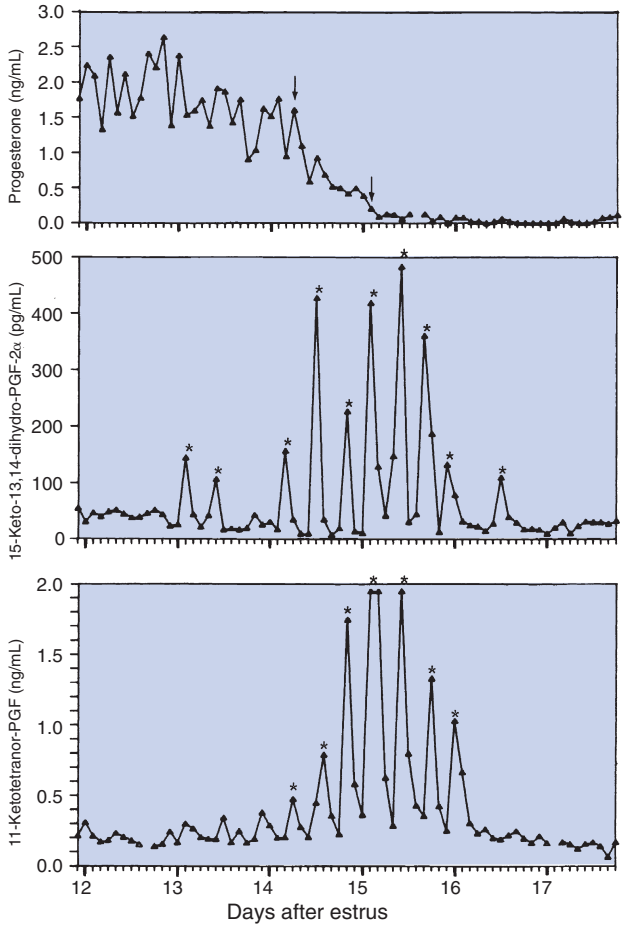
Regression of the CL is important in large domestic nonpregnant animals so that animals reenter a potentially fertile state as soon as possible. The CL life span after ovulation must be of sufficient duration to allow a newly developing conceptus (the fertilized ovum and its associated membranes) to synthesize and release factors that allow the CL to be maintained, but it must be relatively short so that a nonpregnant animal can return to a potentially fertile state. In large domestic animals, the duration of the luteal phase is about 14 days in the absence of pregnancy. This allows large domestic animals to recycle at relatively frequent intervals, approximately every 3 weeks.

Leo Loeb first showed (in 1923) the importance of the uterus for the regression of the CL through hysterectomy studies that extended the luteal phase in guinea pigs. He concluded that the uterus must produce a substance that terminated luteal activity. This information lay dormant for many years, until hysterectomy studies in cattle, pigs, and sheep in the 1950s produced similar results, that is, a prolongation of the luteal phase of the estrous cycle. Through these studies, the concept developed that the uterus is responsible for control of the duration of the life span of the CL, at least in large domestic species (and guinea pigs).

It is now accepted that $PGF_{2\alpha}$, a 20-carbon unsaturated fatty acid, is the uterine substance that causes regression of the CL in large domestic animals, including cattle, goats, horses, pigs, and sheep; $PGF_{2\alpha}$ has no known natural role in CL regression in cats and dogs or in primates. Prostaglandin ($PGF_{2\alpha}$ and PGE) therapy has been used clinically to cause luteolysis in the bitch and queen, for the treatment of pyometra (a uterine infection associated with abnormal endometrial growth and bacterial invasion), or to induce abortion. In large domestic species, regression of the CL is initiated by uterine synthesis and release of $PGF_{2\alpha}$ (likely of endometrial origin) at about 14 days postovulation. The mode of transfer of $PGF_{2\alpha}$ from the uterus to the ovary is thought to occur either by *local countercurrent transfer* or *general systemic transfer.* Countercurrent transfer involves the movement of molecules across the blood vascular system from higher concentrations in the venous effluent (utero-ovarian vein) to an area of lower concentration (ovarian artery) (Fig. 36.4). Systemic transfer involves passage of the molecules through the general circulatory system. In some species (cow and ewe), $PGF_{2\alpha}$ synthesis from a uterine horn only influences the life span of the CL in the ipsilateral ovary. In other species (sow and perhaps mare), $PGF_{2\alpha}$ synthesis from one horn is sufficient to cause regression of CL in both ovaries. This effect likely occurs because of greater production of $PGF_{2\alpha}$ by uterine tissue, as well as a difference in the rate of metabolism of $PGF_{2\alpha}$. $PGF_{2\alpha}$ is rapidly



• **Fig. 36.4** Postulated route by which prostaglandin secreted by the progesterone-primed uterus is able to enter the ovarian artery and destroy the corpus luteum in sheep. (From Baird DT. The ovary. In Austin CR, Short RV, ed. *Reproduction in Mammals,* vol 3. Cambridge, UK: Cambridge University Press; 1986.)



• **Fig. 36.5** Concentrations of progesterone, 15-keto-13,14-dihydro-$PGF_{2\alpha}$, and 11-ketotetranor-PGF metabolites in a nonpregnant ewe. Values identified as significant pulses of either $PGF_{2\alpha}$ metabolite are indicated by asterisks. The times of initiation and completion of functional luteolysis are indicated by arrows. *PGF,* Prostaglandin F. (From Zarco L, Stabenfeldt GH, Basu S, et al. Modification of prostaglandin F-2 alpha synthesis and release in the ewe during the initial establishment of pregnancy. *J Reprod Fertil*. 1988;83(2):527–536.)

metabolized systemically, with more than 90% changed by one passage through the lungs. Thus the system involving the use of $PGF_{2\alpha}$ as the luteolytic agent in large domestic species requires that $PGF_{2\alpha}$ be conserved through a special transfer system or that it be produced in relatively large amounts.

The pattern of synthesis and release of $PGF_{2\alpha}$ is essential to its luteolytic effect. For example, $PGF_{2\alpha}$ synthesis and release must be pulsatile, with pulses occurring at about 6-hour intervals, for *luteolysis* to be affected (Fig. 36.5). The concept has developed that a minimum of four to five pulses within 24 hours is required to cause complete *luteolysis.* If pulse intervals increase significantly before complete luteolysis (e.g., to 12 hours), the CL can recover and continue to function, even if at a lower level of steroid synthetic activity. The uterus must be exposed to estrogen and progesterone to synthesize and release $PGF_{2\alpha}$. Although the initiation of $PGF_{2\alpha}$ synthesis that leads to *luteolysis* is not completely understood, one possible explanation is that estrogen (from an antral follicle) causes the initial synthesis and release of $PGF_{2\alpha}$. In sheep, it is thought that interplay occurs between the uterus and ovary after the initial $PGF_{2\alpha}$ pulse. $PGF_{2\alpha}$ affects the CL to cause both a reduction in progesterone production and the release of *luteal oxytocin.* Oxytocin then interacts with receptors within the uterus to initiate another round of $PGF_{2\alpha}$ synthesis. $PGF_{2\alpha}$ synthesis ceases 6 to 12 hours after progesterone concentrations have become basal, that is, with the completion of *luteolysis.* A system for early recycling is not present in nonpregnant dogs and cats as far as regression of CL; the luteal phase is about 70 and 35 to 40 days, respectively. Bitches experiencing infertility as a result of frequent estrous cycles may have either a pathologically shortened diestrus or anestrus.

Changes in Luteal Life Span in Large Domestic Animals Occur Because of Changes in Prostaglandin $F_{2\alpha}$ Synthesis by the Uterus

Significant changes in the length of the life span of the CL in nonpregnant large domestic species occur only because of changes within the uterus. As discussed in Chapter 38, the presence of an embryo results in the blockage of $PGF_{2\alpha}$ synthesis and a continuance in luteal activity. Prolonged luteal phases also typically occur in mares in the absence of uterine infection. This deficit in mares appears to be a genetic propensity toward inadequate synthesis and release of $PGF_{2\alpha}$. The absence of a uterine horn can also result in a lengthened luteal phase in animals in which the ipsilateral horn controls the CL (local control). In this situation (e.g., in the cow), if ovulation occurs in the ovary ipsilateral to the missing horn, the luteal phase is prolonged because of the need for the ipsilateral uterine horn to control the life span of the CL.

In nonpregnant large domestic animals, inflammatory responses of the endometrium caused by bacterial contamination can result in significant synthesis and release of $PGF_{2\alpha}$, leading to premature luteolysis and a shortening of the estrous cycle. It should be emphasized that luteal activity is almost always normal in the absence of uterine abnormality in large domestic species. Thus short estrous cycles in large domestic animals are pathognomonic for uterine infection.

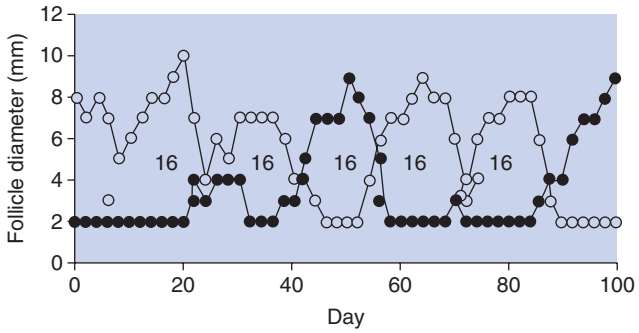
Ovarian Cycles

In Spontaneously Ovulating Animals, Ovarian Cycles Have Two Phases: Follicular and Luteal; Animals That Require Copulation for Ovulation Can Have Only a Follicular Phase

An *ovarian cycle* in a nonpregnant animal is defined as the interval between successive ovulations. The cycle is composed of two phases, an initial *follicular phase* and a subsequent *luteal phase,* with ovulation separating the phases. In most domestic animals and primates, the ovulatory process is governed by internal mechanisms; estrogen from the antral follicle initiates the ovulatory release of gonadotropins. These animals are called *spontaneous ovulators.*

Fundamental differences exist among animals regarding the relationship of the *follicular* and *luteal phases* of the cycle. In higher primates, there is complete separation of follicular and luteal phases, with no significant follicle growth occurring until luteolysis is complete. In large domestic animals, significant follicle growth does occur during the luteal phase of the cycle. For example, in the cow a large antral follicle is present at the onset of *luteolysis,* and in the mare, follicle growth can even result in ovulation of follicles during the luteal phase (about 5% of cycles). Thus in large domestic animals, much of the follicle growth is telescoped into the luteal phase. This situation results in shorter cycles in large domestic animals versus primates (17 to 21 days vs. 28 days); the interval of *luteolysis* to ovulation is shorter in large domestic animals (5 to 10 days) than in primates (12 to 13 days). The period of antral follicle growth leading to ovulation is not appreciably different, however, with the final progression of antral follicle growth requiring about 10 days in large domestic animals and about 12 to 13 days in primates.

Animals that require copulation for ovulation are known as *induced ovulators.* They include cats, rabbits, ferrets, mink, camels, llamas, and alpacas. Copulation replaces estrogen as the stimulus that induces the ovulatory release of gonadotropins. However, these animals require exposure to elevated estrogen concentrations before they can respond to copulation by the release of gonadotropins.



• **Fig. 36.6** Ovarian follicular activity over a period of 100 days in a llama, indicating follicle growth alternating between the left (*open circles*) and the right (*solid circles*) ovaries. (From Vaughan JL, Macmillan KL, D'Occhio MJ. Ovarian follicular wave characteristics in alpacas. *Anim Reprod Sci*. 2004;80:353–361.)

Induced ovulators have follicle growth patterns (in the absence of coitus) in which cohorts of follicles develop, are maintained in a mature state for a few days, and then regress. Follicle growth patterns can be distinctly separated, as in the cat, in which follicles develop and regress over 6 to 7 days, with a minimum of 8 to 9 days between follicle growth waves. Follicle waves can also have some overlap, as in llamas and alpacas (Fig. 36.6), or can closely overlap, as in the rabbit.

The Luteal Phase Is Modified by Copulation in Some Species

In rodent species, the luteal phase of the ovarian cycle is extended by copulation. The life span of CL is only 1 to 2 days in the absence of copulation. Copulation initiates the release of prolactin, which results in prolongation of luteal activity for up to 10 or 11 days in the absence of pregnancy. This phenomenon is often called *pseudopregnancy.* In the canine, spontaneous regression of the CL marking the end of diestrus occurs in association with increased levels of prolactin, causing clinical pseudopregnancy. Nonpregnant bitches can nest, lactate, and nurture objects during this time. The queen can exhibit pseudopregnancy if copulation occurred with an infertile tom.

CLINICAL CORRELATIONS

Pseudopregnancy in the Bitch Presented for Infertility

History

A 5-year-old Irish Setter is presented for failure to conceive. She has heat cycles every 8 months, which appear normal in character and length to her owner. She has been bred three times with no apparent pregnancy resulting. The bitch is sent to the owner of the stud dog early in proestrus and has been bred by artificial insemination, performed by the stud owner. Her owner complains that she always has very misleading false pregnancies about 6 weeks after the breedings, with lactation and reclusive behavior. She was last in heat 3 months previously.

Clinical Examination

Physical examination, complete blood count, chemistries, and a urinalysis are all normal. Ultrasound of the abdomen found normal-appearing ovaries and uterus. Serum progesterone was 0.2 ng/mL, indicating anestrus.

Comment

The most common causes of failure to conceive in the bitch are problems with husbandry (timing and method of breeding) and problems with the stud dog (subfertile or infertile). Sending a bitch to the stud dog too early can interfere with ovulation due to stress, and breeding based on the stud dog owner's opinion of her fertile period can be problematic. The occurrence of pseudopregnancy actually indicates this bitch has a normal ovarian cycle, with ovulation followed by progesterone elevation for ~45 days and eventual elevation of prolactin (at ~6 weeks).

Treatment

Ovulation timing and natural breedings with a proven, fertile male based on the results of timing are advised.

Persistent Luteal Phase in the Mare

History

You have been called to examine a mare that foaled this spring but was not bred at the foal "heat" because of a retained placenta. It has been 40 days since the foal "heat," and the owner wants to know why the mare has not returned to estrus.

Clinical Examination

The main clinical findings are a cervix found (through speculum examination) to be relatively small, tightly constricted, and (through palpation per rectum)

CLINICAL CORRELATIONS—cont'd

to have considerable tone. Rectal palpation also reveals a uterus that has considerable tone. The ovaries are normal in size; in fact, one ovary has a 35-mm follicle. This prompts you to ask the owner whether the mare has been vigorously teased by a stallion for the detection of estrus. The owner brings the teasing stallion to the mare to demonstrate the farm's teasing technique, and as predicted, the mare vigorously rejects the stallion.

Comment

A history of a mare that has been previously in estrus and has not returned to estrus within 30 days usually indicates the presence of a persistent corpus luteum (CL). The CL persists because of inadequate $PGF_{2\alpha}$ synthesis and release, which normally occurs approximately 14 days after ovulation and causes regression of the CL in the absence of pregnancy. The incidence of the syndrome may be as high as 15% to 20%. The CL can remain active for as long as 3 months before the mare is able to synthesize and release $PGF_{2\alpha}$ in amounts sufficient to cause regression of the CL. It is difficult to palpate a persistent CL per rectum because it tends to shrink into the interior of the ovary. The structure may be visualized by ultrasonography, but this is not always possible. The appearance of the cervix and the tone of the cervix and uterus suggest that the genital tubular system is under the influence of progesterone; these findings, together with the history, support a tentative diagnosis. A tentative diagnosis can also be made if the mare returns to estrus within a few days after the administration of $PGF_{2\alpha}$. A definitive diagnosis can be made by progesterone analysis of blood; values are often 1 to 2 ng/mL in this syndrome, versus 3 ng/mL or more in mares with normal–estrous-cycle CLs. Additional supportive diagnostics would be to repeat the palpation and ultrasound examination in several days. If the mare maintains her uterine tone, does not have edema, and maintains cervical tone, these findings would also support a persistent CL.

The clinical finding that can be confusing in this syndrome is the presence of a large follicle in the absence of estrus. Ovarian follicles develop in this syndrome, and sometimes ovulation even occurs. However, mares do not show sexual receptivity in the presence of large follicles if luteal-phase concentrations of progesterone are present. Additionally, they do not develop marked uterine edema or cervical relaxation if progesterone is still present. One possibility that should be considered in a differential diagnosis is that ovarian activity has stopped (i.e., the mare has become anestrous). Although this does not occur often in foaling mares, mares that foal early can be adversely affected by the relatively short photoperiod that is present. In this case, the clinical signs do not support the diagnosis of anestrus.

Treatment

The administration of $PGF_{2\alpha}$ (or one of its analogues) usually initiates regression of the persistent CL and results in the appearance of estrus within 5 to 7 days. The early return to estrus is based on the fact that ovarian follicles tend to develop on a continuous basis throughout the persistent luteal-phase syndrome. Regression of the CL allows the current dominant follicle to continue to develop and produce estrogen, which brings the mare into estrus. One caveat: If a large follicle (e.g., 40 to 45 mm) is present at treatment, the follicle may ovulate before the mare manifests estrus, and the treatment will be judged as failing. In this case, the animal needs to be monitored daily; if ovulation occurs within a few days of treatment, the animal may need to be inseminated artificially if breed rules allow.

Practice Questions

1. The main hormone secreted by the dominant follicle that allows the follicle to maintain its dominant state is:
- **a.** Estrogen
- **b.** Inhibin
- **c.** Oocyte-inhibiting factor
- **d.** Progesterone

2. The factor that is most important in deciding whether a luteal-phase dominant follicle will go on to ovulation is:
- **a.** Inadequate pituitary stimulation
- **b.** Regression of the CL
- **c.** Atresia of the follicle

3. The initiation of the preovulatory LH surge that leads to ovulation in spontaneous ovulators results from:
- **a.** Estrogen
- **b.** Inhibin
- **c.** Progesterone
- **d.** FSH
- **e.** Prolactin

4. The substance responsible for the regression of the CL in large domestic animals is:
- **a.** Estrogen
- **b.** Inhibin
- **c.** Oxytocin
- **d.** Prolactin
- **e.** $PGF_{2\alpha}$

5. Ovarian follicle patterns in animals that are induced ovulators—that is, those that require copulation for the induction of ovulation—are as follows:
- **a.** Ovarian follicle waves greatly overlap.
- **b.** Ovarian follicle waves slightly overlap.
- **c.** Ovarian follicle waves are distinctly separated.
- **d.** All the above.

Bibliography

Adams GP, Ratto MH, Huanca W, Singh J. Ovulation-inducing factor in the seminal plasma of alpacas and llamas. *Biol Reprod*. 2005;73(3):452–457.

Bocci F, Di Salvo P, Zelli R, et al. *Ovarian ultrasonography and progesterone concentration during the periovulatory period in bitches*. Presented at 5th Biannual Congress, European Veterinary Society for Small Animal Reproduction (EVSSAR), Budapest, Hungary, 2005.

Concannon PW. Endocrinologic control of normal canine ovarian function. *Reprod Domest Anim*. 2009;44(Suppl 2):3–15.

Evans AC. Characteristics of ovarian follicle development in domestic animals. *Reprod Domest Anim*. 2003;38(4):240–246.

Feldman EC, Nelson RW, eds. *Canine and Feline Endocrinology and Reproduction*. 4th ed. Philadelphia: Saunders; 2009.

Hafez ESE, Hafez B, eds. *Reproduction in Farm Animals*. 7th ed. Baltimore: Lippincott Williams & Wilkins; 2000.

Johnson MH. *Essential Reproduction*. 7th ed. London: Wiley-Blackwell; 2013.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames: Iowa State University Press; 2003.

Plant TM, Zeleznik AJ, eds. *Knobil and Neill's Physiology of Reproduction*. Vol 1 and 2. 4th ed. London: Academic Press; 2015.

37

The Estrous/Menstrual Cycle

AUTUMN P. DAVIDSON AND GEORGE H. STABENFELDT†

KEY POINTS

Reproductive Cycles

1. The two types of reproductive cycles are estrual and menstrual.

Puberty and Reproductive Senescence

1. Puberty is the time when animals initially release mature germ cells.
2. Reproductive senescence in primates occurs because of ovarian inadequacy, not inadequacy of gonadotropin secretion.

Sexual Behavior

1. Sexual receptivity is keyed by the interaction of the hormones estrogen and progesterone via gonadotropin-releasing hormone in the female and testosterone in the male.

External Factors Controlling Reproductive Cycles

1. Photoperiod, lactation, nutrition, and animal interaction are important factors that affect reproduction.
2. Inadequate nutrition results in ovarian inactivity, especially in cattle.

Reproductive Cycles

The Two Types of Reproductive Cycles Are Estrual and Menstrual

Two types of reproductive cycles are recognized, *estrual* and *menstrual,* with the term *ovarian cycle* representing the interval between two successive ovulations. These terminologies have developed to use certain external characteristics for accurately identifying a particular stage of the reproductive cycle and, most importantly, relating it to the time of ovulation.

Domestic animals have limited periods of *estrus* (the time of *sexual receptivity*); the term *estrous cycle* is used, and the onset of *proestrus* defines the start of the cycle (Fig. 37.1). Primates are sexually receptive during most of the reproductive cycle; the term *menstrual cycle* is used with the onset of *menstruation* (vaginal discharge of blood-tinged fluids and tissues) designated as the start of the cycle (Fig. 37.2). The first day of the cycle for both estrual and menstrual cycles in many species begins shortly after the end of the *luteal phase.* In the dog, the anestrous period separates *diestrus* and *proestrus* (the stages of the cycle are described later).

In domestic animals, proestrus usually begins within 48 hours after the end of the luteal phase; the dog and pig are exceptions with proestrus in the dog delayed by the anestrus phase (2–3 months) and proestrus in the pig not occurring for 5 to 6 days. In primates, menstruation usually begins within 24 hours of the end of the luteal phase. Even though both cycles begin at the same time in relation to the luteal phase (shortly after), the time of ovulation differs. This is because, as previously discussed, luteal and follicular phases are separated in primates with ovulation occurring at a minimum of 12 to 13 days after the onset of menses. In most domestic animals, the *follicular phase* overlaps the luteal phase, therefore ovulation occurs relatively earlier in the estrous cycle. Ovulation is easier to predict in domestic animals (versus primates) because estrus is usually tightly coupled to the preovulatory release of *gonadotropins* and ovulation. The onset of follicular development in primates can be delayed for a variety of reasons (e.g., stress), making the time of ovulation less predictable for primates than for domestic animals.

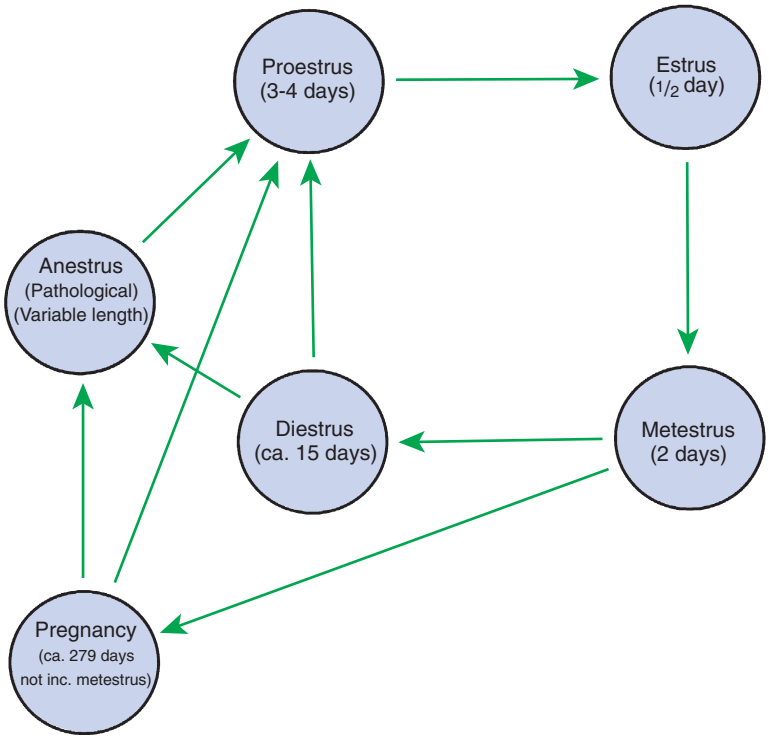
The estrous cycle has been classically divided into stages that represent either behavioral or gonadal events (see Fig. 37.1). The terminology, originally developed for the guinea pig, rat, and mouse, is as follows:

- *Proestrus.* Period of follicle development occurring subsequent to luteal regression and ending at estrus.
- *Estrus.* Period of sexual receptivity.
- *Metestrus.* Period of initial development of the corpus luteum (CL).
- *Diestrus.* Period of mature phase of the CL.

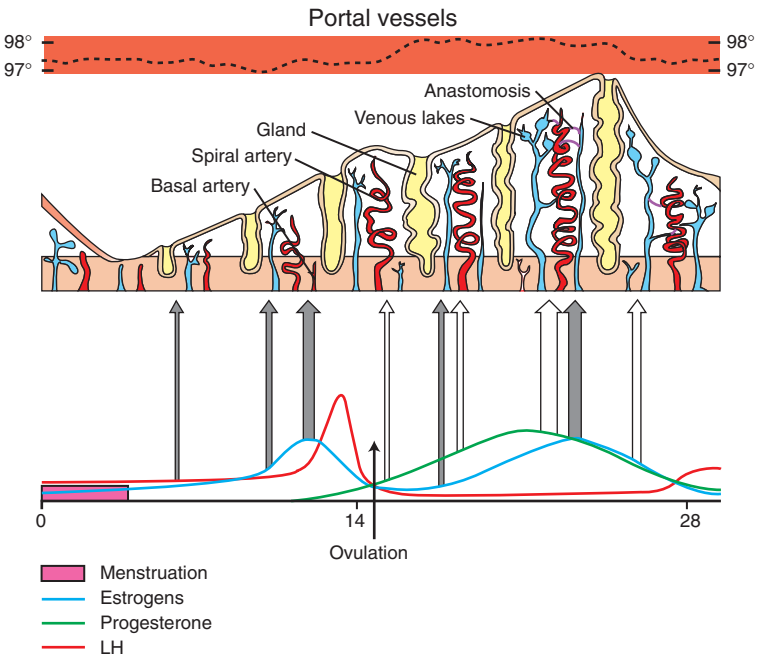
The common terminologies used for domestic animals involve either *behavioral* or *gonadal* activity. The cycle can be described in a behavioral manner by indicating whether animals are in *estrus* (sexually receptive) or not (proestrus, metestrus, and diestrus). The cycle can also be described with reference to the activity of the gonads if differentiation of follicles and the CL is possible. Animals can be in the *follicular phase* (proestrus and estrus) or the *luteal phase* (metestrus and diestrus).

Because the equine CL is relatively difficult to identify by palpation per rectum, horses are usually classified by sexual behavior: estrus or nonestrus. The behavioral classification is also used in other domestic species, including the goat, pig, and sheep, because of the difficulty of determining their ovarian status. The ovarian status of cattle can be determined accurately by palpation per rectum, and cows are usually classified by ovarian status: follicular or luteal. The ovarian status of the dog and cat can be determined by performing vaginal cytology (estrogen effect) and measuring serum progesterone levels. If a CL can be identified, the judgment

†Deceased



• **Fig. 37.1** Various stages of the ovarian cycle of the cow. (From McDonald LE. *Veterinary Endocrinology and Reproduction.* 5th ed. Philadelphia, 2003: John Wiley & Sons.)



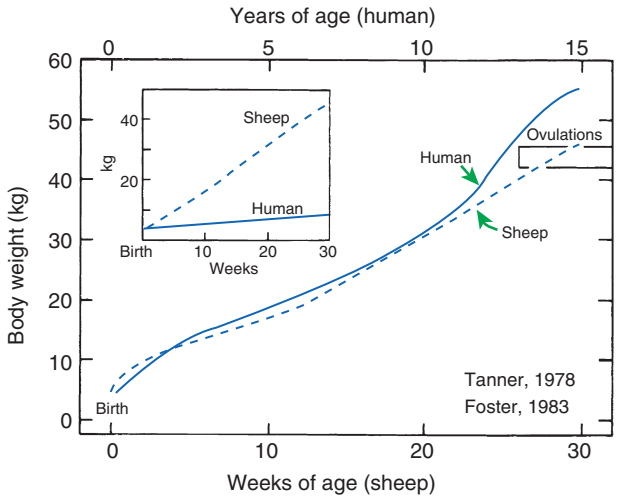
• **Fig. 37.2** Changes in human endometrium during the menstrual cycle. Underlying steroid changes are indicated below, and basal body temperature is indicated above. Thickness of *arrows* (estrogens, *shaded;* progestogens, *white*) indicates strength of action. *LH,* Luteinizing hormone. (From Johnson M, Everitt B, eds. *Essential Reproduction.* 3rd ed. London: Blackwell Scientific; 1988.)

can be made that ovarian activity is normal in the particular animal, because the CL represents the culmination of follicle growth and ovulation.

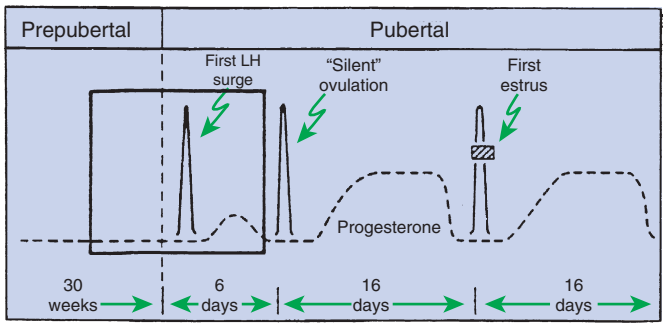
Puberty and Reproductive Senescence

Puberty Is the Time When Animals Initially Release Mature Germ Cells

For females to begin reproductive cycles, they must undergo a process called *puberty.* The term *puberty* is used to define the onset of reproductive life. For the female, although the onset of sexual activity (in domestic animals) or first menstrual bleeding (in primates) is often used as the onset of puberty, the most precise definition is the time of first ovulation. For all species, there is a critical requirement for the attainment of a certain size for puberty to be initiated, in cattle about 275 kg, for example, and in sheep about 40 kg (Fig. 37.3). If this critical requirement is not met because of inadequate nutrition, puberty is delayed. The typical age at puberty for domestic animals is as follows: cats, 6 to 12 months; cows, 8 to 12 months; dogs, 6 to 12 months; goats, 7 to 8 months; horses, 12 to 18 months; and sheep, 7 to 8 months.



• **Fig. 37.3** Body weight from birth through the initiation of ovulation for sheep (mean) and humans (50th percentile). *Inset* shows absolute growth during the first 30 weeks. (From Foster DL, Karsch FJ, Olster DH, et al. Determinants of puberty in a seasonal breeder, *Recent Prog Horm Res.* 1986;42:331–384.)



• **Fig. 37.4** Schematic overview of major events during the transition into adulthood in the female sheep. *LH,* Luteinizing hormone. (From Foster DL, Ryan KD. Mechanisms governing onset of ovarian cyclicity at puberty in the lamb, *Ann Biol Anim Biochem Biophys.* 1979;19:1369.)

Classically, bitches have attained 75% of their adult size before puberty occurs.

The physiological mechanisms involving control of puberty in domestic animals are best known in sheep. One of the fundamental concepts of the onset of puberty involves an increase in the synthesis and release of *gonadotropin-releasing hormone* (GnRH) from the hypothalamus, which drives *gonadotropin* secretion (in pulsatile form) and follicle growth. Before puberty, GnRH and gonadotropin secretion are kept in check because the hypothalamus is highly sensitive to *negative-feedback inhibition by estrogens.* One of the keys to puberty in lambs is a maturation of the hypothalamus, which results in reduced sensitivity to negative feedback by estrogen. Puberty onset is not held back because of lack of responsiveness of the prepubertal gonads, because ovarian follicle development can be elicited by gonadotropin administration.

Changes in *photoperiod* are important for allowing lambs to enter puberty. It has been shown that lambs must have some exposure to a long photoperiod during their prepubertal development; the period can be as short as 1 to 2 weeks (under experimental conditions). Termination of the long photoperiod, which occurs with the summer solstice, allows the sensitivity of the hypothalamus to decrease in response to negative estrogen feedback. The minimal interval from the end of the long photoperiod exposure to the onset of puberty is 10 weeks under experimental conditions. This aspect agrees well with the timing of spontaneous puberty, in which the first ovulation often occurs in the latter part of September (in the Northern hemisphere) or about 13 weeks from the occurrence of the summer solstice. Note that this concept of the initiation of puberty does not involve a decreasing photoperiod; the emphasis is on a turning point that involves the cessation of exposure to a long photoperiod.

With appropriate growth and photoperiod exposure, the secretion of gonadotropins in lambs causes significant follicle growth. This growth is maintained because of decreased sensitivity of the hypothalamus to estrogens produced by growing follicles. The first endocrine event of puberty in the ewe lamb is the appearance of a preovulatory-type surge of gonadotropins, presumably induced by estrogens produced by developing follicles (Fig. 37.4). The gonadotropin surge results in the production of a luteal structure, through luteinization of a follicle(s), which has a short life span of 3 to 4 days. After the demise of the initial luteal structure, another gonadotropin surge occurs, leading to ovulation and the formation of a CL, usually of a normal life span. At this time, cyclical ovarian activity is finally initiated in the ewe lamb.

Photoperiod can have a suppressive effect on the timing of puberty in animals whose ovarian cyclicity is controlled by light. Kittens born in the spring may be large enough to enter puberty by late autumn, but puberty could be delayed a few months if the kittens are under the natural photoperiod with shortening light exposure in fall.

Photoperiod influences the timing of puberty onset in macaque monkeys, depending on the physiological maturity of the individual. The first ovulation, or puberty onset, can occur during the late autumn or early winter at about 30 months of age (20% of animals) or 12 months later at about 42 months of age (80% of animals). The animals undergoing puberty at about 30 months of age have an earlier maturation of the neuroendocrine system, in which significant gonadotropin secretion begins during the previous spring. Thus there is a window of opportunity for the onset of puberty in macaques that must be entered within the favorable photoperiod of decreasing light if puberty is to occur at an earlier time; nutrition and growth are likely determinants of the earlier time for onset of puberty.

The onset of puberty usually results in the establishment of cyclical ovarian activity within a relatively short period (i.e., within a few weeks to a month in lambs). Ewe lambs can initiate normal ovarian activity at the onset of puberty, which can lead to pregnancy (if mated) at the first estrus or they can have false starts with the establishment of limited luteal phases and cessation of ovarian activity for several weeks to a month before they resume ovarian activity. In general, the onset of ovarian cyclicity starts later and ends earlier for ewe lambs compared with adults of the same breed. The earlier cessation of ovarian activity results from an earlier response to negative estrogen feedback.

The initiation of cyclical ovarian activity in pubertal primates takes longer; the first significant follicle growth usually ends in ovulatory failure. In macaque monkeys, 3 to 6 months is usually required after the onset of menarche, or first vaginal bleeding, before the occurrence of the first ovulation of puberty. In humans, follicle growth without ovulation can occur for up to a year before the establishment of normal ovarian cyclicity, including ovulation and CL formation.

For male lambs, the onset of puberty is first keyed when lambs begin to lose their sensitivity to estrogen feedback inhibition, usually by about 15 weeks of age. For many males, this occurs during the

period of increasing, or long, photoperiod, which is in contrast to the ewe lamb. Spermatogenesis (process of sperm production resulting in the presence of mature sperm) usually begins at this time, but because of the length of the process, lambs are usually not capable of successful breeding until about 30 weeks of age or more, or in concert with the onset of puberty in ewe lambs. Thus puberty is a relatively gradual phenomenon in male sheep compared with the abrupt process in females.

Because adult ewes experience the same double gonadotropin surge at the onset of the breeding season, it has been suggested that adult animals recapitulate puberty each year as they enter the breeding season. Recent studies in adult ewes, however, indicate that *refractoriness* to the long photoperiod experienced by animals during the spring and summer is the most critical aspect for the establishment of ovarian activity. Thus the concept that the renewal of ovarian activity in sheep recapitulates puberty appears not to be accurate, at least in some aspects.

Reproductive Senescence in Primates Occurs Because of Ovarian Inadequacy, Not Inadequacy of Gonadotropin Secretion

The end to ovarian activity that occurs in primates is called *menopause.* In humans, for example, it usually occurs between 45 and 50 years of age. Menopause results from the depletion of oocytes, which has occurred throughout the reproductive life of the individual; in essence, it represents ovarian failure. It is not clear whether follicles fail to develop from their primordial state because of an absolute, or relative, reduction in follicle numbers or whether the absence of gonadotropin receptors prohibits follicles from entering the gonadotropin-dependent stage of growth. The initiation of menopause often involves cyclical irregularity caused by failure of follicle development and ovulation. Gonadotropin secretion can be increased, or can remain normal, because of the lack of estrogen and therefore lack of negative feedback on gonadotropin secretion. In the end, ovarian follicle activity ceases, estrogen concentrations decline, and in the absence of negative-feedback inhibition, gonadotropin concentrations increase dramatically.

Reproductive senescence is not recognized in domestic animals. This is partly because some domestic species have lives that are shortened for economic or humane reasons. Nevertheless, a phenomenon such as menopause clearly does not occur in domestic animals. One effect of age can be noted in the dog: estrous-cycle intervals gradually increase from the norm of 7.5 to 12 to 15 months toward the end of the life span. Also, litter size diminishes, and increased neonatal mortality, probably associated with dystocia (difficult birth), occurs with increasing age of the dam. Reproductive senescence in the cheetah has been reported to be a consequence of uterine rather than ovarian changes.

Sexual Behavior

Sexual Receptivity Is Keyed by the Interaction of the Hormones Estrogen and Progesterone via Gonadotropin-Releasing Hormone in the Female and Testosterone in the Male

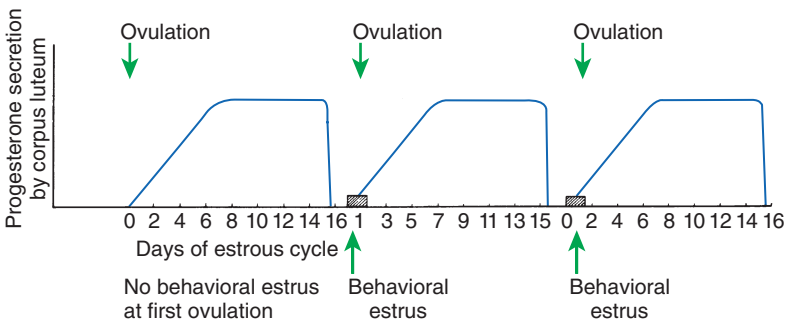
As indicated previously, the establishment of sexual behavior depends on exposure, or lack of exposure, of the hypothalamus to testosterone during the early neonatal period. In essence, testosterone (aromatized to estrogen) causes masculinization of the sexual centers in the hypothalamus; in the absence of testosterone, the hypothalamus becomes feminized. An area within the hypothalamus, the *medial preoptic area,* has been identified in the rat as an area modified structurally by exposure to testosterone.

Several principles exist regarding the effects of hormones on the sexual behavior of domestic animals. First, the magnitude of change in hormone concentration that affects sexual behavior is small; in the cat, for example, an increase in estradiol-17 concentration from 10 to 20 pg/mL of plasma results in signs of proestrus. Second, synergism between hormones is often important for sexual receptiveness; in the dog, for example, estrogen priming and decline followed by progesterone elevation is important. Third, the sequence of exposure to hormones can be important; in the ewe, for example, progesterone priming is required before estrogen exposure for manifestation of estrus.

Estrogen, from the developing antral follicle, is the one hormone required for sexual receptivity in all domestic animals. *Progesterone,* derived from either the granulosa of the preovulatory follicle or the CL, is also important for estrus in some animals (dog).

In sheep and goats, estrus occurs in response to estrogen only if the animal has been exposed previously to progesterone (through the presence of a previous CL). Estrus usually begins within a short period after the end of the luteal phase (i.e., 24–36 hours) because of the presence of large antral follicles at luteolysis; thus the period from last exposure to progesterone and the onset of estrus is short (Fig. 37.5). The requirement of progesterone for sexual receptivity means that the first follicular phase of the breeding season, which leads to ovulation in the ewe, is not accompanied by estrus. Most adult ewes show estrus after the first luteal phase. Ewe lambs often require the exposure of two or more luteal phases before they express estrus.

Of the domestic species, dogs are unusual in that sexual receptivity is keyed by progesterone, produced initially by the granulosa during the preovulatory luteinizing hormone (LH) surge and subsequently by the developing CL. Prior exposure to estrogen makes the female attractive to males but does not produce sexual receptivity; estrus requires the additional exposure to progesterone.



• **Fig. 37.5** Estrous cycle of the ewe, showing how the first ovulation of the season is unaccompanied by estrus. Note the short interval between regression of the corpus luteum and the next ovulation. (From Austin CR, Short RV, eds. *Reproduction in Mammals, Volume 4: Reproductive Fitness,* 2nd ed. © Cambridge University Press 1972, 1985.)

Estrus is often maintained for 7 to 9 days in the presence of a developing luteal phase. In other domestic species, progesterone is inhibitory for estrual activity.

The importance of prior progesterone priming for estrus manifestation has been suggested for dairy cattle by the finding of a reduced incidence of estrus at the first postpartum ovulation (days 15–20). Complete progesterone withdrawal occurs in the cow immediately before delivery, and animals would not have been exposed to progesterone for 2 to 3 weeks in this situation. Sows also have a reduced incidence of estrus at the first ovulation, which usually does not occur until after weaning, usually not until at least 45 days after parturition. Other domestic species (i.e., cats, goats, horses) all show estrus with the first ovulation of the season with no apparent requirement for progesterone priming.

Testosterone is important for *libido* in female primates. The theca layer from degenerating follicles forms an active interstitium that secretes the androgens androstenedione and testosterone. Androgens are also essential for the maintenance of libido in males. Occasionally, castrated males, particularly horses, are able to maintain libido despite the lower concentrations of androgens (of adrenal origin) that are present after castration. These animals can sometimes be differentiated from those with retained (undescended) testicles (*cryptorchid* animals) by testosterone analysis of plasma; however, serum testosterone levels in intact males vary by the minute. A GnRH stimulation test more accurately identifies remaining testicular tissues (2.2 μg/kg intravenously, sampling before and 1–3 hours later). When commercially available, serum LH levels are better for differentiating bilaterally cryptorchid individuals (LH <1 ng/mL) from castrated animals (LH >1 ng/mL).

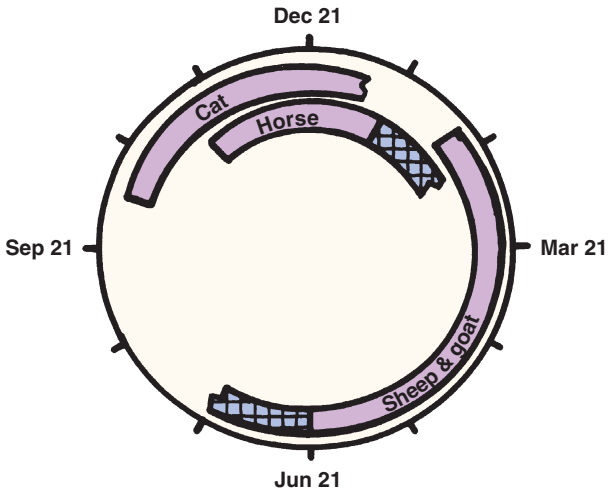
Both experimental and circumstantial evidence indicates that GnRH plays a role in sexual receptivity. The administration of GnRH to ovariectomized rats produced sexual (lordosis) responses, and in prepubertal gilts, GnRH administration resulted in the occurrence of estrus within 24 hours. The circumstantial evidence is that the onset of sexual receptivity in animals is tightly coupled to the onset of the preovulatory gonadotropin surge. Because the preovulatory gonadotropin surge is the result of an increased rate of pulsatile release of gonadotropins driven by increased GnRH synthesis and release, it is likely that this increased GnRH secretory activity affects sexual centers within the hypothalamus for the promotion of sexual receptivity. This allows the onset of the ovulatory process, triggered by the gonadotropin surge, to be tightly coupled with sexual receptivity.

External Factors Controlling Reproductive Cycles

Photoperiod, Lactation, Nutrition, and Animal Interaction Are Important Factors That Affect Reproduction

Photoperiod

Photoperiod controls the occurrence of reproductive cycles in a number of domestic species, including cats, goats, horses, and sheep. The result is that these animals have an annual period in which they have continuous (cyclical) ovarian activity, as well as another period of no ovarian activity, termed *anestrus.* The response to photoperiod is different among these species; cats and horses are positively affected by increasing light, and goats and sheep are positively affected by a decreasing photoperiod (Fig. 37.6).



• **Fig. 37.6** Diagrammatic representation of the effect of photoperiod on ovarian activity in the typical cat, horse, sheep, and goat. The *bars* represent periods of ovarian inactivity (anestrum). The transitional periods for the horse, sheep, and goat are shown by the *hatched portions* of the bars. (Reprinted from *Dukes' Physiology of Domestic Animals,* 10th ed, edited by Melvin J. Swenson. Copyright © 1984 Cornell University Press. Used by permission of the publisher, Cornell University Press.)

A positive response to a change in the photoperiod usually occurs relatively soon after the occurrence of the summer or winter solstice (i.e., within 1–2 months). A negative response to a change in the photoperiod usually requires a longer duration for an effect (i.e., 2–4 months to suppress ovarian activity after the occurrence of the particular solstice). The net result is that, in the absence of pregnancy, cyclical ovarian activity usually occupies more than half the year for these four seasonally breeding species.

In cats, cyclical ovarian activity can range from late January through October (in the Northern hemisphere). In horses, the usual range of ovarian activity is from March through October. Conversely, sheep and goats have ovarian activity from late July through February or March (depending on the breed). As indicated previously, progesterone priming immediately before follicle development is required for sexual receptivity in sheep. The full length of the reproductive season of sheep is not manifested externally because (1) the first ovulation is not preceded by the presence of a CL, and (2) the last follicle phase may be delayed because of a negative photoperiod, with the priming effects of progesterone lost before follicle growth.

The main translator of photoperiod is the *pineal gland,* which produces *melatonin* in response to darkness. The central nervous system (CNS) pathway involved with the translation of light includes the retina, suprachiasmatic nucleus, superior cervical ganglion, and pineal gland. Whereas melatonin has been previously described as *antigonadal,* this is obviously not true, because both short and long phases of darkness, with resultant short and long duration of melatonin secretion, can have a positive effect on reproductive cycles. In sheep, however, exposure to increasing darkness may be important only for maintaining ovarian activity. The onset of ovarian activity is thought to occur in response to the development of refractoriness to the long photoperiod. The development of *photorefractoriness* to a long photoperiod as a requisite to ovarian cyclicity is consonant with the fact that sheep can begin cyclical ovarian activity even before the onset of the summer solstice.

Of the seasonal breeders, the cat is the most sensitive to photoperiod change; estrus, in conjunction with the presence of mature antral follicles, can occur as early as January 15 (in the Northern hemisphere). Initial follicle activity likely begins at least 10 days before the first expression of estrus, or 15 days after the winter solstice. Thus a total photoperiod change as brief as 15 minutes can be perceived and translated by the cat into ovarian activity.

The suppressive effects of photoperiod can be overcome by exposure to artificial lighting regimens. This is relatively easy in the case of cats and horses, in which environments with photoperiods are compatible with ovarian activity (i.e., at least 12 hours of light per day). If the photoperiod is established before the end of ovarian activity in the autumn, cyclical ovarian activity continues through the time associated with anestrus. If mares are allowed to become anestrous in the autumn, it can take a minimum of 2 months of exposure of mares to increased light to reestablish ovarian activity. The usual time for placing mares under lights is December 1 (in the Northern hemisphere), with cyclical ovarian activity expected by early February.

It is usually not possible to place sheep and goats in light-tight barns to increase their exposure to dark to overcome the suppressive effects of increasing light. One recent development in this regard has been the oral or systemic (implant) administration of melatonin to sheep during the spring. This exposure to melatonin has resulted in an early onset of ovarian activity and increased the number of multiple ovulations above that normally observed at the beginning of the breeding season.

Lactation

Lactation can have suppressive effects on ovarian activity. In pigs, suppression of ovarian activity is complete; sows do not come into estrus until after piglets are weaned. Cats can have ovarian activity suppressed throughout lactation, although they occasionally come into estrus during the latter part of lactation. Ovarian activity tends to be suppressed in lactating beef cows, with the first estrus and ovulation not occurring before day 45 postpartum. The suckling process appears important to ovarian suppression; dairy cows are not suppressed by lactation unless it involves a large nutritional deficit.

Goats and sheep usually begin lactation during a photoperiod that is increasingly suppressive for ovarian activity, and therefore the reestablishment of ovarian activity in these species is confounded by the photoperiod. Ewes delivering in the autumn ovulated as early as postpartum day 12 (average, postpartum day 23), indicating that lactation has little suppressive effect on ovarian activity in sheep. Mares usually ovulate by postpartum days 10 to 15, with lactation having no suppressive effect on ovarian activity regarding this ovulatory interval.

One of the concepts of *lactational suppression of ovarian activity* involves the importance of suckling with its related stimulation of prolactin synthesis. Inhibiting factors for prolactin synthesis, including dopamine and the GnRH-associated peptide, need to be suppressed for prolactin synthesis to proceed. The sensory input from suckling suppresses the production of these prolactin-inhibiting factors. Because both dopamine and GnRH-associated peptide are essential links in the synthesis of gonadotropins, their reduced output results in reduced ovarian activity through decreased gonadotropin synthesis and release.

Pheromones

Pheromones are chemical compounds that allow communication among animals through the olfactory system. When sexual behavior is affected, the compounds are called *sex pheromones.* Pheromones arise from several tissue sources; the most prominent sources for animals are sebaceous glands, the reproductive tract, and the urinary tract.

Some of the first experiments that demonstrated the potency of male odors to influence reproductive behavior were done in mice. One syndrome, called the *Whitten effect,* involves the synchronization of estrus in female mice through the sudden introduction of a male (or male odor through bedding), with a large number of animals cycling within 3 days of introduction of the male. The effect of the pheromones in this case is to stimulate the synthesis and release of gonadotropins. Another syndrome, called the *Bruce effect,* involves the blockage of pregnancy development by the introduction of a different (strange) male in proximity to a recently bred female. The effect of the odor of the strange male is to block the release of prolactin, the hormone responsible for the maintenance of the CL in association with pregnancy in rodents. Regression of the CL in this case produces fetal loss. Thus pheromones can strongly affect reproductive cycles.

Pheromones are important for the attraction of the male to the female at the time of sexual receptivity. Sexual attractiveness of the female evolves from the pheromones that she elicits on a limited, cyclical basis in association with estrus. For example, methyl-p-hydroxybenzoate, isolated from the vaginal secretions of dogs in proestrus and estrus, has produced intense anogenital interest by males when applied to anestrous females. Females are also influenced by male odors; sows in estrus assume a breeding (*rigidity*) stance when exposed to the urine of males. Androgens can serve as pheromones, or they can influence the production of substances within the kidney that influence female sexual behavior. The attractiveness of the female to the male involves a change in perception of the male by the female resulting from a changing physiological state within the female, not because of changes that are occurring in the male.

The classic way for males to delineate their territory has been for them to mark the area with urine. In general, pheromones that affect sexual behavior tend to have a musk type of odor. The classic pheromone used by humans is perfume, which is derived from civetone, a cyclical 17-carbon compound obtained from the civet cat.

The Whitten effect has been used to manipulate the estrous cycles of animals. In sheep, males are introduced into a flock of ewes before the breeding season to advance, or ensure, ovarian cyclicity at the beginning of the breeding season. Whereas it was previously thought the effect of the introduction of a male was short-lived (i.e., a gonadotropin response could only be obtained within the first few days from ewes that had antral follicles), it is now clear that the interaction of rams with ewes over extended periods of the anestrus results in earlier ovarian activity.

As discussed, pheromones can account for some of the effect of the male. More recent studies, however, have shown that *sight* of the male by the female and *physical contact* are important factors that influence gonadotropin secretion and thus ovarian activity. This occurs in llamas; the sight of a male llama causes females to enter estrus immediately.

The Whitten effect has also been used to influence the onset of puberty in pigs. The introduction of males into groups of gilts beginning several weeks before the expected time of puberty (180–200 days) has been used to ensure, or advance, the onset of puberty. The *dormitory effect,* the well-recognized synchronization of menstrual cycles in roommate women, occurs in bitches kenneled together as well.

Inadequate Nutrition Results in Ovarian Inactivity, Especially in Cattle

In dairy cattle, genetically selected for high productivity, the ability to produce up to 100 pounds of milk per day is a remarkable achievement. It is almost impossible for dairy cows to consume enough feed during the first part of the lactation cycle to maintain their body weight, and they are often in a negative nutritional balance for up to 100 days postpartum. Animals must have an adequate level of nutrition to initiate ovarian activity, so ovarian activity is suppressed until a positive energy balance is established. If an owner wants a dairy cow to produce large quantities of milk, the owner must be willing to wait for nutrition to "catch up" with milk production.

Inadequate nutrition can affect ovarian activity in the postpartum period. A management practice sometimes used to enhance production efficiency is to maintain beef cows on a marginal plane of nutrition during the winter. This approach forces animals to use fat that has been developed and stored during the grazing season. If pregnant beef cows are not returned to a positive nutritional balance by the last month of gestation, the reestablishment of ovarian cyclicity, which usually occurs between days 45 and 60 postpartum, will be delayed. Another situation that can affect ovarian activity involves pregnant beef heifers. These animals often need extra nutrition in the postpartum period to reestablish ovarian activity because they have requirements for growth as well as for lactation.

CLINICAL CORRELATIONS

Sexual Attractiveness in the Spayed Bitch

History

You are called by a veterinary colleague who has seen a bitch owned by one of her important clients. The client is upset because the dog is attracting males despite recently having undergone an ovariohysterectomy. You inquire whether the dog allows intromission by males. Although the answer is "no," the owner is sure that a portion of an ovary was left *in situ*. Your colleague is sure she removed the ovaries during the surgical procedure. You are asked to examine the dog as a favor to your colleague.

Clinical Examination

The dog has a vulva that is slightly swollen with a small amount of creamy discharge present. An examination of a vaginal mucosal smear reveals noncornified epithelial cells and an increased number of neutrophils. You indicate to the owner that you believe the (nondiscriminatory) male dogs are being attracted by the presence of an infection in the urogenital tract; the owner needs more convincing. You decide to obtain a reproductive endocrine panel (progesterone level and an LH test) and a urinalysis from the dog. The values for progesterone are low (<1.0 ng/mL) and the LH is positive (indicating lack of feedback) and thus do not support the active presence of ovarian tissue. The presence of white blood cells and bacteria in the urine suggests a urinary tract infection. A culture and sensitivity can be submitted to determine the specific bacteria and its antibiotic sensitivity.

Comment

It is common for bitches with genitourinary infections to attract male dogs, presumably because of the odors generated by the infection. One of the most important points of differentiation as to cause (i.e., bladder infection versus presence of ovarian remnant) is to know about the sexual behavior of the animal. The bitch normally allows intromission by a male only if she has been exposed to progesterone after priming with estrogen. This situation occurs only if an ovarian follicle is present and has begun to luteinize after the preovulatory LH surge. If the animal in question had allowed intromission, the presence of an ovarian remnant would be more likely. Because of the failure of the bitch to allow intromission, you can conclude that the animal almost certainly lacks ovarian tissue. With regard to the endocrine analysis, if the animal was completely ovariectomized, the LH level will be high, indicating a lack of negative feedback from either estrogen (not seen in the vaginal cytology) or progesterone (< 1.0 ng/mL). More recently, anti-müllerian hormone (AMH) assays have become available and are a reliable indicator of the presence of ovarian tissue. Nonestrogenized vaginal cytology and a low progesterone level do not rule out the presence of ovarian tissue, but if the sample is obtained when the animal is showing the sexual attractiveness, it can be stated with assurance that the behavior is not caused by hormones or, by extension, activity of ovarian tissue (some other source of blood in the vulva such as a urinary tract infection could be the cause). Had vaginal cytology indicated estrogen exposure, the clinician must differentiate whether the exposure is exogenous (from contact with an estrogen containing transdermal medication commonly used by menopausal women) or endogenous (ovarian remnant); the AMH test will differentiate this.

Treatment

The bladder infection is treated, and the owner is instructed to keep the female away from males until the infection has cleared.

Trying to Get a Mare Pregnant

History

You work in Minnesota. Your clients have a 4-year-old thoroughbred mare that they would like to breed early in the season, so that they will have as old of foal as possible, when it is old enough to race.

Clinical Examination

The mare has never been bred before. The breeding soundness examination (which assesses conformation, ultrasonography of her reproductive tract, cytology, and culture of her uterus) is normal.

Comment

If the mare can be bred as early in the year as possible, the owner will hopefully have a foal early in January (in the Northern hemisphere). This will hopefully result in the foal being at its maximal development to race. The photoperiod is regulated by the pineal gland, which produces melatonin. Processing of light signals through the CNS is regulated by the retina, suprachiasmatic nucleus, superior cervical ganglion, and the pineal gland. Ovarian activity begins in response to refractoriness of the photoperiod. There are two major methods for stimulating ovulatory activity in mares. One is to increase the total light per day. The other involves pulsatile amounts of light. Alternatively, drugs can be used to alter the cycle.

Treatment

The typical method to stimulate ovulation involves mimicking the photoperiod for 60 days before breeding. For this mare, 16 hours of day light should be started in November. In other cases, short periods of "flashes" of light (i.e., 1 hour of light 9.5 hours after onset of darkness) during the photosensitive period (10 hours after onset of darkness) can be used to stimulate ovulatory activity. An alternative option is to stimulate the mare with dopamine antagonists (i.e., domperidone) while increasing the photoperiod (i.e., increase photoperiod for 2 weeks, and then add dopamine antagonist until mare starts to cycle). Using one of these methods, the mare should begin ovulatory activity early in January and hopefully will be bred and become pregnant for an early January foal.

Practice Questions

1. The first estrous cycle of the cow subsequent to parturition follows which sequence?
- **a.** Anestrus, diestrus, estrus, metestrus, proestrus
- **b.** Anestrus, estrus, diestrus, metestrus, proestrus
- **c.** Anestrus, metestrus, diestrus, estrus, proestrus
- **d.** Anestrus, proestrus, estrus, diestrus, proestrus
- **e.** Anestrus, proestrus, estrus, metestrus, diestrus

2. The usual situation in large domestic animals is for a dominant follicle or dominant follicles to be present at the time of luteal regression, with sexual receptivity manifested within 1 to 2 days after luteal regression; the one large animal species that is the exception to this generalization is the:
- **a.** Cow
- **b.** Doe
- **c.** Ewe
- **d.** Mare
- **e.** Sow

3. The hormones that form the foundation for sexual receptivity are:
- **a.** Estrogen and $PGF_{2\alpha}$
- **b.** Progesterone and estrogen
- **c.** Estrogen and GnRH
- **d.** Progesterone and $PGF_{2\alpha}$
- **e.** $PGF_{2\alpha}$ and GnRH

4. Decreasing light turns off cyclical ovarian activity after a number of months, whereas increasing light reverses the process after a number of months, including the development of a transitional period. This description fits which domestic species?
- **a.** Cat
- **b.** Cow
- **c.** Dog
- **d.** Goat
- **e.** Horse
- **f.** Pig
- **g.** Sheep

5. What response results from the Whitten effect, in which the introduction of a male into a group of noncyclical animals results in the reestablishment of ovarian activity?
- **a.** Increased estrogen secretion
- **b.** Increased progesterone secretion
- **c.** Increased prolactin secretion
- **d.** Increased follicle-stimulating hormone secretion
- **e.** Increased luteinizing hormone secretion
- **f.** Increased follicle-stimulating hormone and luteinizing hormone secretion

6. Which one of the following domestic species requires progesterone priming, in addition to estrogen, to manifest estrus (therefore not manifesting estrus with the first ovarian cycle in the postpartum period)?
- **a.** Cat
- **b.** Dog
- **c.** Goat
- **d.** Horse
- **e.** Pig
- **f.** Sheep

Bibliography

Breen KM, Billings HJ, Wagenmaker ER, et al. Endocrine basis for disruptive effects of cortisol on preovulatory events. *Endocrinology*. 2005;146(4):2107–2115.

Card C. Hormone therapy in the mare. In: Samper JC, eds. *Equine Breeding Management and Artificial Insemination*. 2nd ed. St Louis: Saunders; 2009.

Davidson AP, Stabenfeldt GH. Reproductive cycles. In: Klein BG, eds. *Cunningham's Textbook of Veterinary Physiology*. 5th ed. St Louis: Saunders; 2013.

Feldman EC, Nelson RW, eds. *Canine and Feline Endocrinology and Reproduction*. 4th ed. Philadelphia: Saunders; 2009.

Gorman SP, Levy JK, Hampton AL, et al. Evaluation of a porcine zona pellucida vaccine for the immunocontraception of domestic kittens (Felis catus). *Theriogenology*. 2002;58(1):135–149.

Hafez ESE, Hafez B, eds. *Reproduction in Farm Animals*. 7th ed. Baltimore: Lippincott Williams & Wilkins; 2000.

Johnson MH. *Essential Reproduction*. 7th ed. London: Wiley-Blackwell; 2013.

Kutzler M, Wood A. Non-surgical methods of contraception and sterilization. *Theriogenology*. 2006;66(3):514–525.

National Research Council. *Nutrient Requirements of Dogs and Cats*. Washington, DC: National Academies Press; 2005.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames: Iowa State University Press; 2003.

Plant TM, Zeleznik AJ, eds. *Knobil and Neill's Physiology of Reproduction*. Vol. 1 and 2. 4th ed. London: Academic Press; 2015.

Simpson GM, England GCW, Harvey MJ, eds. *BSAVA Manual of Canine and Feline Reproduction and Neonatology*. 2nd ed. Gloucester, UK: BSAVA; 2011.

38

Pregnancy and Parturition

AUTUMN P. DAVIDSON AND GEORGE H. STABENFELDT†

KEY POINTS

Pregnancy

1. The development of an embryo involves fusion of an oocyte and spermatozoon within the oviduct.
2. Extension of the life span of the corpus luteum in large domestic species and augmentation with placental progesterone in cats is essential for pregnancy maintenance.
3. The placenta acts as an endocrine organ.

Parturition

1. Fetal cortisol initiates delivery through increased secretion of estrogen and thus prostaglandin $F_{2\alpha}$.

Pregnancy

The Development of an Embryo Involves Fusion of an Oocyte and Spermatozoon Within the Oviduct

The development of a new individual requires the transfer of male gametes to the female genital tract for fertilization of the female gamete(s). *Spermatozoa,* which have been concentrated and stored in the epididymis, gradually change from *oxidative* (aerobic) to *glycolytic* (anaerobic) *metabolism* as they progress through the epididymis. In this situation, spermatozoa are in a state of reduced metabolism. Mature sperm are only able to metabolize a certain sugar, *fructose,* within the reproductive tract. Lactose, glucose, dextrose, and fructose have all been used in commercially available semen extenders.

Sperm are ejaculated usually into the vagina, although some domestic species (horse and pig) ejaculate directly into the cervix and uterus. The volume of prostatic fluid (third fraction) after the sperm-rich second fraction of the canine ejaculate forces sperm through the cervix. The movement of sperm through the cervix is aided by estrogen-induced changes in cervical mucus, which result in the formation of channels that facilitate movement of sperm. This has been particularly emphasized in primates, in which the thinning of mucus occurs just before ovulation, a factor that can be used to predict the time of ovulation.

The environment of the female genital system is generally inhospitable to the survival of sperm; for example, white blood cells are quickly attracted to the uterine lumen because sperm cells are foreign to the female genital tract. Special reservoirs have evolved in the female tract to aid in the survival of sperm during transport; these include the cervix and oviduct, the latter involving areas at the uterotubule junction and within the ampulla. The reservoirs are progressively filled (from caudal to cranial in the tract), requiring hours before the *oviductal reservoirs* are full. Finally, the reservoir within the ampulla is able to release a few sperm on a continuous basis, so that *fertilization* can occur shortly after the arrival of mature oocytes within the oviduct.

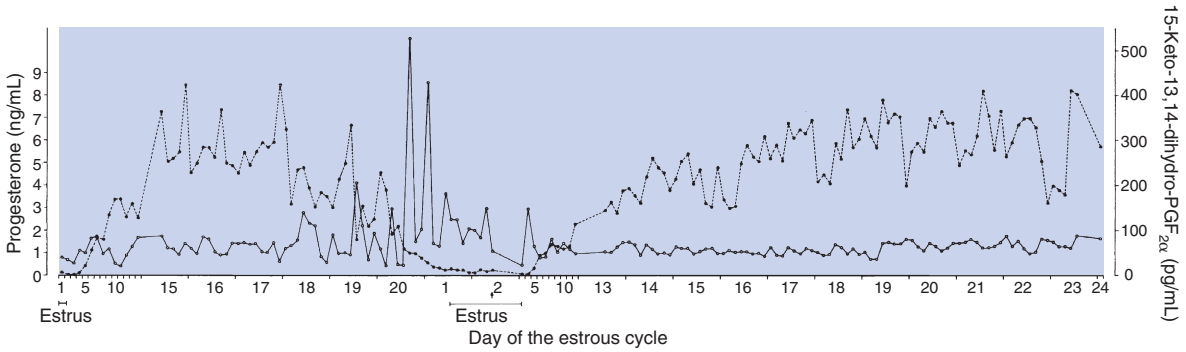
The first studies in *sperm transport* emphasized the rapidity of the process, with sperm reported passing from the vagina to the fimbriated end of the oviduct within minutes. It is now known that sperm undergoing so-called fast transport are not involved in *fertilization;* in fact, they are damaged by the rapid transport.

Sperm need to undergo changes within the female genital tract that are a prerequisite for fertilization; the process is called *capacitation.* One of the effects of *capacitation* is the removal of glycoproteins from the sperm cell surface.

The glycoproteins, perhaps added for protective purposes, interfere with *fertilization.* This change allows sperm to undergo the *acrosome reaction* when they come in contact with oocytes. The *acrosome reaction* involves the release of hydrolytic enzymes from the acrosomal cap; this may be important for penetration of the sperm through the granulosa and zona pellucida to the oocyte plasma membrane. *Hyaluronidase* causes breakdown of hyaluronic acid, an important component of the intercellular matrix of granulosa cells that surround the oocyte. *Acrosin,* a proteolytic enzyme, digests the acellular coating around the oocyte. Both enzymatic events allow the sperm to penetrate to the oocyte. The *acrosome reaction* also changes the surface of the sperm, which allows it to fuse with the oocyte. The *acrosome reaction* results in tail movements that feature a flagellar beat that tends to drive sperm in a forward direction.

Because of the changes that spermatozoa must undergo within the female reproductive tract before *fertilization,* the deposition of sperm before ovulation is the preferred timing for producing maximal fertility. An exception to this takes place when sperm with reduced longevity are used, such as the case with chilled-extended or frozen-thawed semen. In these cases, deposition of semen into the female reproductive tract should occur close to the time of ova maturation associated with fertilization. Females are usually sexually receptive for at least 24 hours before ovulation and, in the natural setting (free interaction between genders), insemination usually occurs a number of hours before the occurrence of ovulation. Even with induced ovulators, such as cats, the interval from copulation to ovulation is usually 24 hours or more. In essence, the system has evolved to have ready-to-fertilize sperm at the fertilization site when fertilizable oocytes arrive. This concurs with the finding that the life span of male gametes tends to be twice that of female gametes.

†Deceased



• **Fig. 38.1** Relationship between prostaglandin release, as indicated by the measurement of 15-keto-13,14-dihydroprostaglandin $F_{2\alpha}$, and progesterone production by the corpus luteum during a nonfertile cycle and after a conception in the same cow. (From Kindahl H, Edquist LE, Bane A. Blood levels of progesterone and 15-keto-13,14-dihydro-prostaglandin $F_{2\alpha}$ during the normal oestrous cycle and early pregnancy of heifers. *Acta Endocrinol [Copenh].* 1976;82(1):134–49.)

The presentation of male gametes before female gametes in the oviduct implies that oocytes are ready for fertilization on arrival in the ampulla; this is likely true for a majority of animals. A prerequisite for fertilization of the oocyte is that it must undergo the first meiotic division before fertilization. Although this occurs in a number of species before ovulation, in the horse and dog the first meiotic division does not occur until after ovulation (in the dog, not for at least 48 hours). In this situation, spermatozoa often wait for oocytes to mature in the oviduct before *fertilization* can occur. One means of adaptation to delayed completion of meiosis is that spermatozoa have a longer life span in the dog (6–11 days) and horse compared with other domestic species.

After *fertilization* has occurred, the *embryo* usually develops to the *morula,* or early *blastocyst* stage, within the oviduct before moving into the uterus. This period, lasting usually 4 to 5 days, affords the uterus time to finish its inflammatory response involving the removal of spermatozoa. This period also allows the endometrial glands time to secrete nutrients under the influence of progesterone from the developing corpus luteum (CL); the nutrients are essential for the development of embryos during their preimplantation stage.

An interesting finding in the mare is her ability to distinguish fertilized from unfertilized oocytes; unfertilized oocytes from previous cycles are retained within the oviduct, whereas recently fertilized oocytes (embryos) move through the oviduct to the uterus. It is likely that all animals recognize pregnancy by the presence of an embryo(s) at the early oviductal stage. However, this recognition does not necessarily result in prolongation of the CL and the continued production of progesterone, which is essential for the maintenance of pregnancy. In the bitch, despite ovulation and ova maturation spanning several hours, embryonic stages are synchronized by some mechanism inherent to the bitch's reproductive tract.

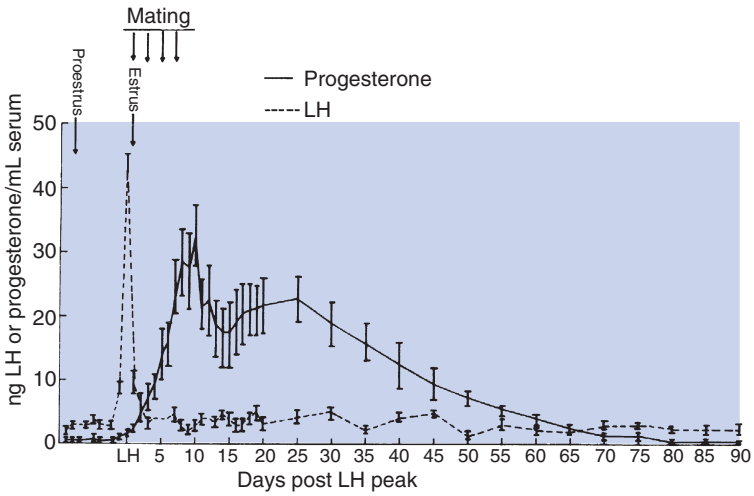
Extension of the Life Span of the Corpus Luteum in Large Domestic Species and Augmentation With Placental Progesterone in Cats Is Essential for Pregnancy Maintenance

For those domestic animals (cattle, goats, horses, pigs, sheep) whose luteal activity is controlled by the uterus, modification of *uterine prostaglandin* $F_{2\alpha}$ ($PGF_{2\alpha}$) synthesis and release is critical for the establishment of pregnancy. The embryo apparently produces substances that modify uterine production of $PGF_{2\alpha}$. *Estrogen synthesis* by the embryo is one way the endometrium may be informed regarding the presence of an embryo. A specific protein of embryonic origin called *trophoblastin,* produced before day 14 of pregnancy (or postovulation) in both sheep and cattle, is of immunological interest for the establishment of pregnancy; it has a close structural relationship to the molecule *interferon.* Movement of the embryo(s) in the tract is also important for pregnancy recognition. In the mare, the embryo moves throughout both horns before being fixed at day 16. In pigs, a minimal number of embryos need to be present (about four), presumably to occupy a sufficiently large area of the endometrium. Litter-bearing animals also use transuterine migration to maximize the opportunity for fetal development, a procedure that aids the recognition of pregnancy process. The end result is either suppression of $PGF_{2\alpha}$ synthesis, as seen in the cow (Fig. 38.1), or modification of the secretion mode (*continuous* instead of *pulsatile*), as seen in sheep. The absence of *pulsatile secretion* of $PGF_{2\alpha}$ seems to be critical for the extension of the life span of the CL and the establishment of pregnancy in large domestic species.

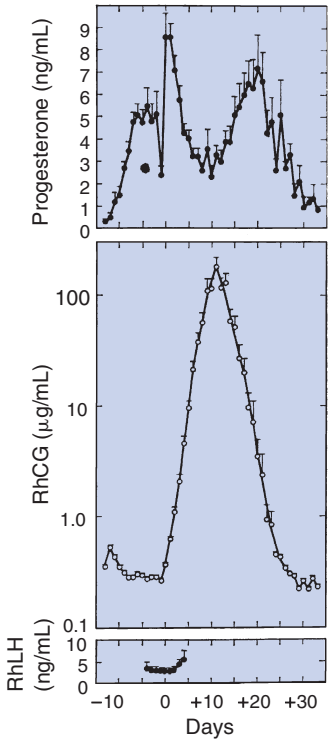
In the cat, the CL lasts for 35 to 40 days after ovulation regardless of the presence of pregnancy, and thus early modification of luteal activity is not essential for the establishment of pregnancy. Implantation occurs at about day 13, which allows the fetoplacental unit to influence and extend luteal activity that is compatible with pregnancy maintenance. The luteotropic hormone responsible for luteal maintenance in the cat is not known. One hormone that likely synergizes with progesterone for the support of pregnancy is relaxin, a placental hormone produced in the cat beginning at about day 20 of gestation (see later discussion). Placental production of progesterone occurs during the last third of pregnancy. In one study, 50% of queens ovariectomized after day 40 of gestation could maintain pregnancy, likely as a result of local placental progesterone production.

The dog does not extend its luteal phase during pregnancy; the luteal phase in the nonpregnant animal is only slightly longer (70 days) than in pregnant animals. Nevertheless, enhancement of luteal activity occurs through an unidentified placental *luteotropin,* possibly *relaxin,* with progesterone secretion enhanced beginning at about day 20 of gestation or a few days after implantation. Early in the *luteal phase,* luteal function in the bitch is likely autonomous. During the second half of the *luteal phase,* luteinizing hormone (LH) and prolactin are likely *luteotrophs* (Fig. 38.2).

The rescue of the CL at the onset of pregnancy in primates involves the production of a *luteotropin* called *chorionic gonadotropin*



• **Fig. 38.2** Luteinizing hormone (*LH*) and progesterone concentrations during pregnancy in nine dogs. *Vertical bars* represent the standard error of the mean. (From Smith MS, McDonald LE. Serum levels of luteinizing hormone and progesterone during the estrous cycle, pseudopregnancy and pregnancy in the dog. *Endocrinology.* 1974;94(2):404–12.)



• **Fig. 38.3** Summarization of 15 early pregnancies in normal rhesus monkeys normalized to the day of corpus luteum rescue (day 0). Points are means plus or minus standard error. Note the temporal relationship between luteal progesterone production (before day + 10) and chorionic gonadotropin output. *RhCG,* Rhesus chorionic gonadotropin; *RhLH,* rhesus luteinizing hormone. (From Knobil E. On the regulation of the primate corpus luteum. *Biol Reprod.* 1973;8:246.)

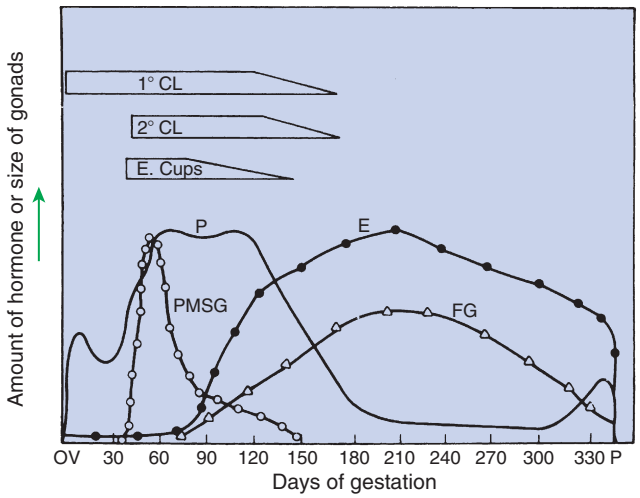
(CG; for humans, hCG), which is produced by trophoblastic cells (*syncytiotrophoblasts*) of the embryo (Fig. 38.3). For trophoblast tissue to produce CG, it must have intimate contact with the interstitium of the endometrium. This contact occurs by a type of implantation called *interstitial,* in which the embryo penetrates the endometrium at about 8 to 9 days after fertilization in humans and nonhuman primates. Secretion of CG begins 24 to 48 hours after implantation, with immediate enhancement of luteal progesterone production. Rescue of the CL in human pregnancy occurs as late as 4 to 5 days before the end of the luteal phase. hCG is thought to be responsible for nausea in early pregnancy in humans.

As indicated, interstitial implantation is essential to the development of pregnancy in primates. *Implantation* is less invasive in the dog and cat, with the type termed *eccentric.* In the large domestic species, "invasion" of the endometrium is minimal; implantation occurs within special endometrial protrusions called *caruncles* in ruminants and by relatively minor *villus* invasion of the endometrium in horses and pigs. Domestic animals depend more on uterine secretions for the support of pregnancy than do primates. For cattle and horses, the first indications of implantation begin about days 25 to 30, and another 7 to 10 days likely passes before a significant amount of embryonic nutrition is obtained through the implantation site. Subclinical uterine infections, or an inadequate number of endometrial glands, can interfere with the establishment of pregnancy in the species in which a long interval exists from fertilization to implantation. The cervix forms an important barrier to contamination of the uterine lumen in both the nonpregnant and the pregnant animal; in the latter, the cervix becomes effectively sealed.

The Placenta Acts as an Endocrine Organ

Besides the essential role of providing nutrients and oxygen for embryonic metabolism, the *placenta* functions as an endocrine organ. One of the most important functions of the placenta is the *production of progesterone.* In primates, this function is established early in gestation, and the placenta likely can maintain pregnancy within 2 to 3 weeks after implantation in primates. Placental production of sufficient progesterone to maintain pregnancy occurs later in domestic animals (sheep, day 50 of 150-day gestation; horse, day 70 of 340-day gestation; cat, day 40 of 65-day gestation); in some species, the placenta never produces enough progesterone to support pregnancy (cattle, goats, pigs, dogs).

The *production of estrogen,* in contrast to that of progesterone, requires interaction between the fetus and the placenta. This interaction has been best described in primates, in particular by the experiments of the Hungarian immigrant to Sweden Egon Dicfaluszy. He and his co-workers found that the primate placenta is unable to produce estrogen from progesterone even though the steroids are only separated by androgens in the steroid biochemical synthetic pathway. The placenta simply does not possess the enzymes



• **Fig. 38.4** Summary of the temporal relationships among changes in hormonal concentrations and morphological changes throughout the gestational period of the mare. *1° CL,* Primary corpus luteum; *2° CL,* secondary corpora lutea; *E,* estrogens; *E. Cups,* endometrial cups; *FG,* fetal gonads; *P,* progesterone; *PMSG,* pregnant mares' serum gonadotropin (equine chorionic gonadotropin). (From Daels PF, Hughes JP, Stabenfeldt GH. Reproduction in horses. In: Cupps PT, ed. *Reproduction in Domestic Animals.* 4th ed. New York, NY: Academic Press; 1991.)

necessary for the conversion of progesterone to androgens. Therefore a system has evolved in which the placenta supplies *pregnenolone,* the immediate precursor of progesterone, to the fetus, and the *fetal zone of the adrenal cortex* transforms pregnenolone to a C-19 androgen, *dehydroepiandrosterone.* This is returned to the placenta, which is able to convert dehydroepiandrosterone to an estrogen. In humans, the primary estrogen of pregnancy is *estriol.* Because the fetus is involved in the production of estriol, the well-being of the fetus can be judged by determining estriol concentrations in the plasma of the mother.

The production of estrogen in the mare also involves an interaction between the placenta and fetus (Fig. 38.4). From the work of Pashen and Allen, we know that the fetal gonads replace the fetal adrenals in primates as the key fetal endocrine organ involved in the cooperative synthesis of estrogen. The interstitial cells of the gonads appear to be the interactive cells, with fetal gonads enlarging to a size greater than the maternal gonads during the latter part of gestation. The production of estrogens during pregnancy in other domestic species, occurring relatively late in gestation, may involve the development of placental enzymes that allow progesterone to be metabolized to estrogens without the direct intervention of a fetal endocrine organ. (Fetal cortisol, however, is important for the induction of these placental enzymes, particularly in sheep; see next section.)

The protein hormones produced during pregnancy tend to be of placental origin. For example, *relaxin* is a hormone produced by the placenta in the cat, dog, and horse beginning at about days 20, 20, and 70, respectively. Besides its importance for preparing the soft tissues of the pelvic canal for passage of the fetus at birth (see the discussion on parturition), *relaxin* may be important for the support of pregnancy through a synergistic action with progesterone. In exception to the general rule of protein hormone production by the placenta, *relaxin* is produced by the CL in the pig, cow, and primates during pregnancy, with prepartum release occurring in conjunction with *luteolysis.*

The only CG identified in domestic animals to date is *equine CG* (eCG, formerly called *pregnant mares' serum gonadotropin* by its discoverer, Harold Cole) (see Fig. 38.4). The eCG is produced by trophoblast cells that initially form as a band on the chorion (*chorionic girdle*), detach themselves around day 35 of pregnancy, penetrate the endometrium, and form associations of cells called *endometrial cups.* The eCG enhances progesterone production by the primary CL of pregnancy and aids in the formation of additional (secondary) CL through the *luteinization,* or ovulation, of preformed follicles. The essentiality of eCG for pregnancy maintenance is not known, because the primary CL is adequate for maintaining pregnancy.

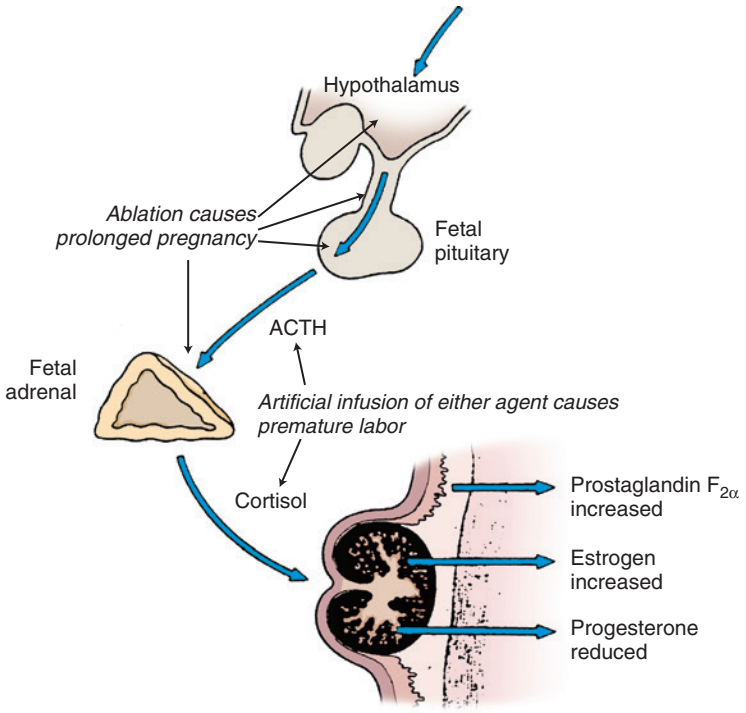
Placental lactogen is another placental protein hormone. Its production increases in primates as CG secretion wanes during pregnancy. Placental lactogen has been reported in goats and sheep, with secretion increasing during the latter part of gestation. The hormone appears to have both *somatotropic* and *lactogenic effects* on the basis of growth hormone–like and prolactin-like properties. In dairy cattle, for example, placental lactogen may be important for mammary gland alveolar development, setting the stage for the next lactation. Another hormone whose production is increased during pregnancy, *prolactin,* also is important for alveolar development during the prepartum period. Prolactin is not a hormone of placental origin; prolactin increases during the latter part of gestation because of the effect of estrogen on its release from the adenohypophysis. Prolactin is luteotrophic in the dog.

Parturition

Fetal Cortisol Initiates Delivery through Increased Secretion of Estrogen and Thus Prostaglandin $F_{2\alpha}$

During pregnancy, the uterus progressively enlarges and stretches because of the growing fetus(es). Progesterone plays an important role in maintaining the quiescence of the myometrium and promoting a tightly closed cervix. During the latter part of gestation, estrogen begins to influence uterine muscle by stimulating the production of *contractile protein* and the formation of *gap junctions;* the former increases the contractile potential of the uterus, and the latter facilitates the contractile process through increased communication among smooth muscle cells. Thus important changes that set the stage for parturition begin weeks before the actual process begins. In the end, the uterus is converted from a quiescent to a contractile organ, and importantly, the cervix relaxes and opens to allow the fetus to be delivered.

An important question about parturition concerns what initiates the process. In domestic animals, maturation of the fetus eventually brings about changes that initiate the delivery process. The key organ of the fetus responsible for initiating the process is the *fetal adrenal cortex,* with the hypothalamus and adenohypophysis playing important supporting roles. This concept came from the work at the University of California, Davis by Liggins and Kennedy, who showed that destruction of the anterior pituitary of the sheep fetus resulted in prolongation of gestation; Drost subsequently found the same results after fetal adrenalectomy. Critical changes in *cortisol secretion* by the fetus eventually result in the synthesis and release of $PGF_{2\alpha}$ from the uterus, which produces myometrial contraction and relaxation of the cervix. The following details of the initiation of parturition emphasize ruminants. It is postulated that elevated fetal cortisol levels also contribute to the initiation of parturition in the dog and cat.



• **Fig. 38.5** Diagrammatic summary explaining how the fetal lamb controls the onset of labor. Experimental procedures that lengthen or shorten pregnancy are shown. *ACTH,* Adrenocorticotropic hormone (corticotropin). (Redrawn from Liggins CG. The fetal role in the initiation of parturition in the ewe. In: Wolstenholme GEW, O'Connor M, eds. *Fetal Autonomy.* London: J & A Churchill; 1969.)

The maturation of the fetal adrenal cortex is of critical importance in the initiation of parturition. The adrenal cortex likely becomes progressively sensitive to fetal adrenocorticotropic hormone (ACTH, corticotropin) (Fig. 38.5). The time of adrenal maturation is under fetal genetic control, as shown by studies conducted on fetal lambs of different breeds in the same uterus (produced by embryo transfer) in which the prepartum initiation of cortisol production occurred at times that were characteristic (and different) for the breed. Fetal cortisol induces placental enzymes (*17-hydroxylase* and *C17–20 lyase*) that direct steroid synthesis away from progesterone to estrogen. This process occurs at different times prepartum in domestic species, beginning at prepartum days 25 to 30 in cattle, 7 to 10 in pigs, and 2 to 3 in sheep. The end result of increased estrogen secretion is the secretion of *prostaglandins,* particularly $PGF_{2\alpha}$. $PGF_{2\alpha}$ is the pivotal hormone for the initiation of parturition; once its secretion begins, and the acute phase of delivery is activated. The role of oxytocin in the initiation of delivery is not certain; it likely complements $PGF_{2\alpha}$ when the delivery process has started.

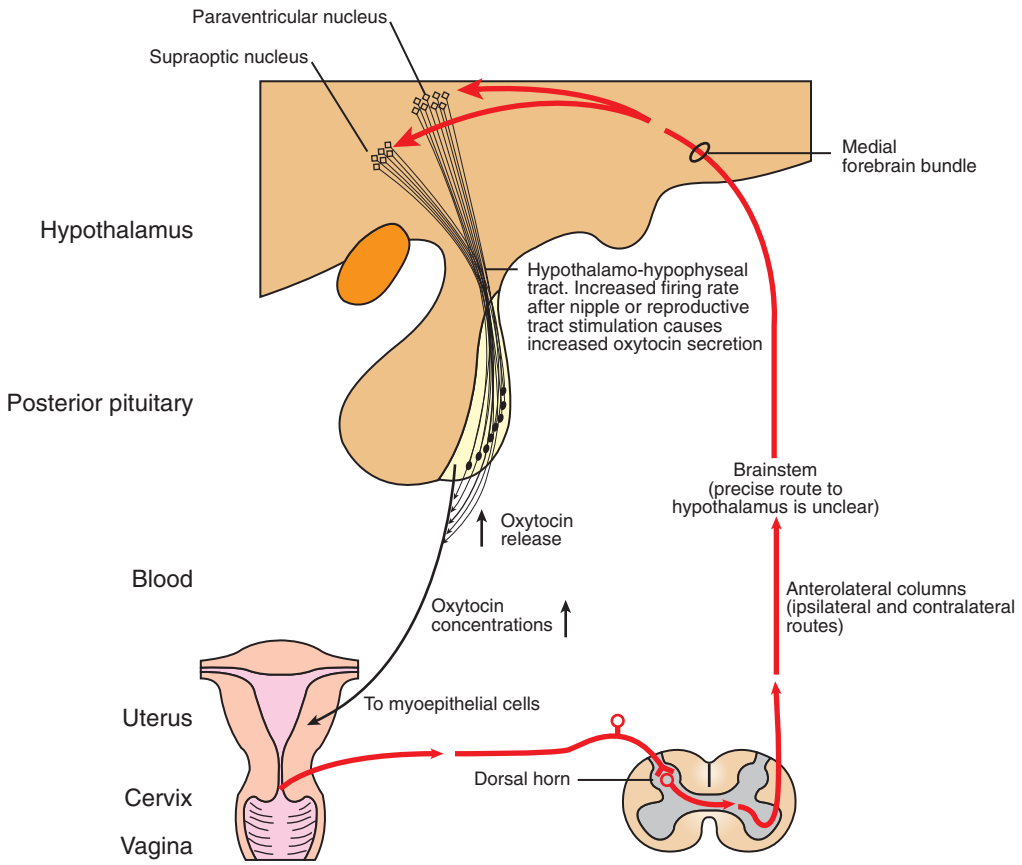
The synthesis of $PGF_{2\alpha}$ is thought to come about through increased availability of the substrate *arachidonic acid,* which is the main rate-limiting step in the synthesis of $PGF_{2\alpha}$. Estrogens are proposed to influence the system by making available the enzyme *phospholipase A,* a membrane-bound lysosomal enzyme that initiates the subsequent hydrolysis of phospholipids and release of arachidonic acid. This likely results from an increasing estrogen/progesterone ratio, with progesterone initially stabilizing, then estrogens destabilizing, lysosomal membranes. The end result is increased availability of arachidonic acid for the synthesis of $PGF_{2\alpha}$. The onset of $PGF_{2\alpha}$ synthesis results in the immediate release of the hormone because $PGF_{2\alpha}$ is not synthesized and stored. The critical effect of $PGF_{2\alpha}$ on the myometrium is to release intracellular calcium ion, which binds to *actin* and *myosin* to initiate the contractile process. Prostaglandins, both PGE and $PGF_{2\alpha}$, also have important effects on the *cervix,* which allow it to relax and dilate, permitting passage of the fetus. The end result is a direct effect of $PGF_{2\alpha}$ on the intracellular matrix of the cervix in which there is a loss of collagen with a concomitant increase in *glycosaminoglycans,* the latter affecting the aggregation of collagen fibers.

In some animals, such as the cow, goat, dog, and cat, $PGF_{2\alpha}$ synthesis and release initiate regression of the CL beginning 24 to 36 hours before delivery, with complete withdrawal of progesterone occurring 12 to 24 hours before labor. Progesterone withdrawal does not initiate labor; it is the release of $PGF_{2\alpha}$ that both causes *luteolysis* and drives myometrial contractions.

In the mare, as in primates, delivery occurs even though progesterone concentrations remain elevated during the process. In this situation, $PGF_{2\alpha}$ is able to overcome the suppressive effects of progesterone on myometrial activity. For animals dependent on placental production of progesterone for pregnancy maintenance, it is not possible to turn off one function (i.e., steroid synthesis) and still continue with other functions necessary for the support of the fetus through the time of delivery.

Oxytocin is also important to the delivery process (Fig. 38.6). Estrogen induces *oxytocin receptor* formation in the myometrium. Recent information indicates that significant amounts of oxytocin are released only with the entry of the fetus into the birth canal. Oxytocin release occurs through the *Ferguson reflex.* The afferent arm of the reflex involves passage of impulses through sensory nerves in the spinal cord to the appropriate nuclei in the hypothalamus; the efferent arm involves transport of oxytocin from the neurohypophysis by the vascular system. Oxytocin is synergistic with $PGF_{2\alpha}$ in promoting contraction of the uterus.

As noted earlier, a hormone important for the preparation of parturition is *relaxin.* This hormone was first identified as being responsible for the separation of the pubic symphysis through relaxation of the interpubic ligament. Relaxin causes the ligaments and associated muscles surrounding the pelvic canal to relax, which allows the fetus to expand the pelvic canal to its fullest potential. In the mare, a well-defined area of muscle softening can be discerned on the midline from the top of the croup through the ventral commissure of the vulva. In the cow, muscles posterior to the hip



• **Fig. 38.6** The neuroendocrine reflex (Ferguson reflex) underlying oxytocin synthesis and secretion. (From Johnson M, Everitt B, eds. *Essential Reproduction.* 3rd ed. London: Blackwell Scientific; 1988.)

become relaxed to the point that they undulate as the animal walks in the final 24 hours before parturition. In the cow and pig, the CL is the source of relaxin. In both these species, the prepartum release of $PGF_{2\alpha}$ causes luteolysis, with a concomitant decline in progesterone production and the release of preformed relaxin. In other domestic species, such as cats, dogs, and horses, the source of relaxin is the placenta. In these species, significant relaxin production begins during the first part of gestation, with values sustained through parturition. A relaxin assay has been developed for the diagnosis of pregnancy in the dog, with good accuracy after 25 days of gestation. Relaxin may be important in these species for the maintenance of pregnancy in synergism with progesterone (see Fig. 38.2).

The *first stage of parturition* involves presentation of the fetus at the internal os of the cervix. This likely results from increased myometrial activity caused by $PGF_{2\alpha}$ release. When the cervix opens and the fetus passes into the pelvic canal, myometrial contractions become less important for delivery of the fetus; *abdominal press,* accomplished by closure of the epiglottis and contraction of maternal abdominal muscles, becomes the main force involved in the delivery process. This explains why breathing exercises in women prevent bearing down (abdominal press) too early in the delivery process. The actual delivery process is called the *second stage of parturition.*

The *third stage of parturition* involves the delivery of the fetal membranes. In litter-bearing animals, such as the cat, dog, and pig, the placental membranes are delivered often with, or immediately after, the appearance of each fetus. In single-bearing species, the placenta may be delivered immediately or within a few hours. From studies done on the mare at the University of California, Davis, we know that major, sustained surges of $PGF_{2\alpha}$ occur in the immediate postpartum period and are important for expulsion of placental membranes and reduction of uterine size through myometrial contraction. $PGF_{2\alpha}$ is likely the most important component of uterine size reduction in the immediate postpartum period for all domestic species. This can be inferred from the episodes of discomfort that parturient animals undergo during the hours immediately after delivery.

The neonate must make a major physiological adjustment to life on the outside. The major change involves the vascular system, in particular the respiratory system. During fetal life, blood bypasses the lungs (except for the perfusion of lung tissue in support of development) by two routes: through the ventricles by way of the *foramen ovale* and from the pulmonary artery to the aorta by the *ductus arteriosus.* The foramen ovale is closed functionally at birth by a flap of tissue in the left ventricle through the development of higher pressures within the left versus the right ventricle. Although the ductus arteriosus immediately constricts at birth, it requires months before it is completely closed. This course of closure is also true for the *ductus venosus,* which serves as a hepatic shunt during fetal life. The rapid conversion from a fluid to a gaseous environment, as occurs at birth, is a truly remarkable adaptation.

CLINICAL CORRELATIONS

Prolonged Gestation in the Cow

History

You are called to examine a purebred Holstein cow that is 12 days overdue compared with the herd gestation average of 280 days. She was artificially inseminated, was diagnosed pregnant 35 days later, and has not been observed in estrus since insemination. You inquire about the presence of bulls on the dairy farm, but there are none.

Clinical Examination

The cow has a greatly enlarged abdomen. On palpation of the uterus per rectum, you find the presence of a large calf. The cow certainly appears to be term as far as the size of the calf. You are puzzled, however, by the lack of colostrum in the udder.

Comment

The history and physical examination findings are compatible with an animal that has a pregnancy that is defective in terms of the initiation of parturition. A normal fetal hypothalamic-pituitary-adrenocortical system is essential for the production of cortisol, which initiates the delivery process. In the cow, this can begin 3 to 4 weeks prepartum, with fetal cortisol directing the increased production of estrogen; this in turn eventually initiates $PGF_{2\alpha}$ synthesis and release. The deficit could be caused by a malformed adrenal gland, pituitary, or hypothalamus. In one syndrome described for Holsteins, the critical defect was a lack of corticotropin-producing cells in the pituitary, which led to inadequate stimulation of the adrenal cortex and inadequate fetal cortisol production. The lack of lactogenesis reflects that the endocrine changes beginning 3 to 4 weeks prepartum as a prelude to delivery are also important for lactogenesis, and in their absence, colostral formation is delayed.

Treatment

The animal can respond to glucocorticoids, with delivery usually occurring 2 to 3 days later. The placenta is normal in this situation, and the systemic administration of glucocorticoids substitutes for fetal cortisol in initiating the endocrine events that lead to parturition. Lactogenesis is usually initiated by glucocorticoid treatment, although the process is usually less advanced than that expected at normal delivery. Because the calf continues to grow in utero in this syndrome, it is often too large to be delivered *per vaginam,* and a cesarean section may have to be performed 2 to 3 days after treatment in concert with dilation of the cervix.

You need to tell the owner that the calf will likely not survive because of inadequate adrenal secretion. If the calf were an extremely valuable bull prospect, one could administer both glucocorticoids and mineralocorticoids for a number of months with the hope the animal would eventually be able to take over its own adrenal support (this actually occurred in one case at the University of California, Davis). It would be questionable, however, to initiate treatment of the calf on the basis that the disease is an autosomal recessive inherited condition.

Prolonged Gestation in the Bitch

History

A 4-year-old multiparous bitch is presented for pregnancy diagnosis. Previously, she has had a litter of six and a litter of eight without incident. She was bred by vaginal artificial insemination based on vaginal cytology serial progesterone tests 42 and 44 days previously, but her owner doesn't think she looks pregnant.

Clinical Examination

Her physical examination is normal, and uteromegaly evident on abdominal palpation. Abdominal ultrasound reveals a single viable fetus.

Comment

Evaluation of her previous timing tests determines that her true gestational length is 48 to 50 days, as she was bred on days 3 and 5 after the initial rise in progesterone suggesting the LH surge. Given her historical normal litter size, the fertility of the stud dog or the handling of the artificial insemination are questionable. Additionally, the presence of a single fetus in a breed with average litter size of eight raises concerns for prolonged gestation as a consequence of a lack of fetal crowding and stress, which normally precipitates fetal cortisol production. Fetal overgrowth occurs as a consequence, with resultant dystocia concerns.

Treatment

Reevaluation of the fetal size radiographically is advised at 62 days gestation based on ovulation timing. The fetus appears too large for the pelvic canal, and an elective cesarean section is performed successfully on day 64.

Practice Questions

1. Active rescue of luteal activity through suppression of pulsatile prostaglandin synthesis and release by the production of embryonic signals must occur in which of the following species for a developing pregnancy to have the early progestational support essential for pregnancy maintenance? (Select all that apply.)
- **a.** Cattle
- **b.** Dog
- **c.** Goat
- **d.** Horse
- **e.** Pig
- **f.** Sheep

2. In primates, it has been established that estrogen production during much of pregnancy is a cooperative venture between fetal adrenals and the placenta. The domestic species most extensively studied in this regard is the horse. In this species the main two interactive organs involved in the synthesis of estrogen during pregnancy are the placenta and the:
- **a.** Fetal adrenals
- **b.** Fetal gonads
- **c.** Fetal liver
- **d.** Fetal hypothalamus
- **e.** Fetal pituitary

3. Which of the following hormones initiates the final process that eventually leads to parturition?
- **a.** Maternal estrogen
- **b.** Maternal progesterone
- **c.** Fetal cortisol
- **d.** Maternal relaxin
- **e.** Maternal prostaglandin
- **f.** Maternal oxytocin

4. The hormone that initiates the myometrial contractile process that acutely initiates parturition is:
- **a.** Maternal estrogen
- **b.** Maternal progesterone
- **c.** Fetal cortisol
- **d.** Maternal relaxin
- **e.** Maternal prostaglandin
- **f.** Maternal oxytocin

5. The hormone released by the passage of the fetus into the pelvic canal through the cervix is:
- **a.** Maternal estrogen
- **b.** Maternal progesterone
- **c.** Fetal cortisol
- **d.** Maternal relaxin
- **e.** Maternal prostaglandin
- **f.** Maternal oxytocin

Bibliography

Austin CR. Sperm maturation in the male and female genital tracts. *Biology of Fertilization*. 2012;2:121–155.

Feldman EC, Nelson RW, eds. *Canine and Feline Endocrinology and Reproduction*. 4th ed. Philadelphia: Saunders; 2009.

Hafez ESE, Hafez B, eds. *Reproduction in Farm Animals*. 7th ed. Baltimore: Lippincott Williams & Wilkins; 2000.

Jackson PGG. *Handbook of Veterinary Obstetrics*. 2nd ed. Philadelphia: Saunders; 2004.

Kowalewski MP, Beceriklisoy HB, Aslan S, Agaoglu AR, Hoffmann B. Time related changes in luteal prostaglandin synthesis and steroidogenic capacity during pregnancy, normal and antiprogestin induced luteolysis in the bitch. *Anim Reprod Sci*. 2009;116(1):129–138.

LeBlanc MM, Lopate C, Knottenbelt DC, Pascoe RR. The mare. In: Knottenbelt DC, Pascoe RR, Lopate C, LeBlanc MM, eds. *Equine Stud Farm Medicine and Surgery*. 2nd ed. Edinburgh: Saunders; 2003.

Lennoz-Roland M. *Practical uses of aglepristone: review of a recent expert meeting. Presented at 5th Biannual Congress,* European Veterinary Society for Small Animal Reproduction (EVSSAR), *Budapest, Hungary*, 2006.

Nelson RW, Couto CG. *Small Animal Internal Medicine-E-Book*. Elsevier Health Sciences; 2014.

Pineda MH, Dooley MP, eds. *McDonald's Veterinary Endocrinology and Reproduction*. 5th ed. Ames: Iowa State University Press; 2003.

Plant TM, Zeleznik AJ, eds. *Knobil and Neill's Physiology of Reproduction*. Vol. 1 and 2. 4th ed. London: Academic Press; 2015.

Simpson GM, England GC, Harvey MJ, eds. *BSAVA Manual of Small Animal Reproduction and Neonatology*. Gloucester, UK: BSAVA; 2010.

Stout TA. The early pregnancy. In: Samper JC, eds. *Equine Breeding Management and Artificial Insemination*. 2nd ed. St Louis: Saunders; 2009.

Tsutsui T, Suzuki Y, Toyonaga M, et al. The role of the ovary for the maintenance of pregnancy in cats. *Reprod Domest Anim*. 2009;44(s2):120–124.

39

The Mammary Gland and Lactation

AUTUMN P. DAVIDSON AND GEORGE H. STABENFELDT†

KEY POINTS

Anatomical Aspects of the Mammary Gland

1. The milk-secreting cells of the mammary gland develop through the proliferation of epithelium into hollow structures called alveoli.
2. Most of the milk that accumulates before suckling or milking is stored in the alveoli, even though animals have enlarged milk-storage areas called cisterns.
3. A suspensory system involving the udder of the cow allows the animal to carry a large amount of milk.

Control of Mammogenesis

1. Initial development of the mammary gland is programmed by embryonic mesenchyme.
2. Proliferation of the mammary duct system begins at puberty, with ducts under the control of estrogen, growth hormone, and adrenal steroids, and the alveoli under the control of progesterone and prolactin.

Colostrum

1. Prepartum milk production (without removal) results in the formation of colostrum.
2. The ingestion of colostrum is important because of the passive immunity it confers through the presence of high concentrations of immunoglobulins.
3. The time immunoglobulins can be absorbed through the neonatal gut is limited to the first 24 to 36 hours of life.
4. Lipids (particularly vitamin A) and proteins (caseins and albumins) are high in concentration in colostrum; carbohydrates (lactose) are low.

Lactogenesis

1. Prolactin, inhibited by dopamine and stimulated by vasoactive intestinal peptide, is the most important hormone involved in the process of milk synthesis, or lactogenesis; growth hormone is also important for lactogenesis.
2. The release of fat into milk from the alveolar cell involves constriction of the plasma membrane around the fat droplet; fats are dispersed in milk in droplet form.
3. Milk proteins and lactose are released from alveolar cells by the process of exocytosis.

Milk Removal

1. Efficient milk removal requires the release of oxytocin, which causes contraction of muscle cells that surround the alveoli (myoepithelial cells) and movement of milk into the ducts and cisterns.

First Nursing

1. Carbohydrate stores are good in neonates born as singles or twins, whereas carbohydrate stores are low in neonates born in litters; consequently the former can stand a longer interval to first suckling than can the latter.

Composition of Milk

1. Fats are the most important energy source in milk.
2. Lactose, composed of glucose and galactose, is the main carbohydrate of mammalian milk.
3. The main proteins in milk are called *caseins* and are found in curd.

The Lactation Cycle

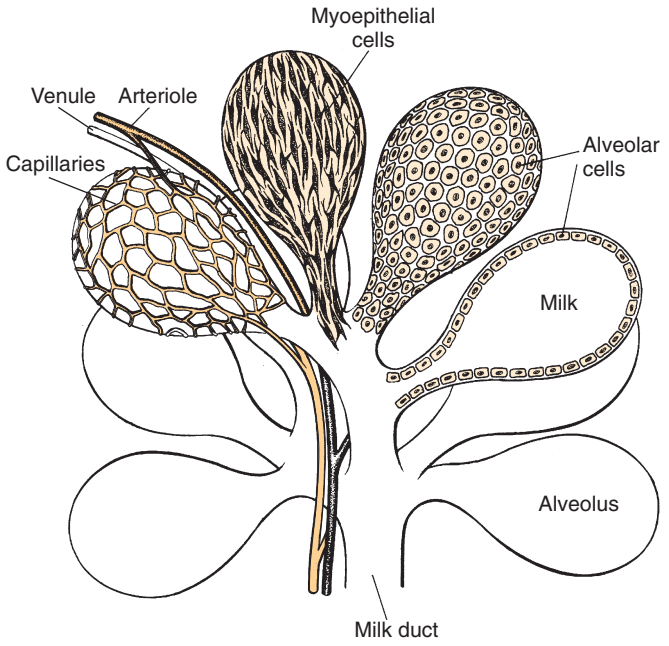
1. Milk production peaks at 1 month postpartum in dairy cattle, followed by a slow decline in production; milking usually stops at 305 days of lactation so that the animal can prepare the mammary gland for the next lactation.
2. Lactation can be induced by hormone administration (estrogen and progesterone) and enhanced by growth hormone and increased photoperiod exposure.

Diseases Associated With the Mammary Gland

1. The main diseases that affect the mammary gland directly are mastitis (prevalent in dairy cattle and dogs) and neoplasia (prevalent in intact dogs and cats).
2. The main conditions that involve the mammary gland indirectly are passive transfer of red blood cell agglutinating antibodies by the ingestion of colostrum (mare, queen) and hypocalcemia caused by the transient drain of calcium that occurs with initiation of lactation (dairy cattle) or during the perinatal period (dog).

†Deceased

Animals that belong to the class *Mammalia* are characterized as having bodies basically covered with hair, delivering live young instead of eggs (the monotremes are an exception), and, pertinent to this chapter, nurturing their young through the use of structures called *mammary glands.* The ability of mammals to nurture their young through milk secretion by mammary glands during the early part of postfetal life has given these animals survival advantages. Because the reproductive strategy of mammals involves the production of far fewer young, compared with reptiles, amphibians, and birds, mammary glands have allowed mammals to be much more efficient in the nurture of their young. Egg-laying classes of animals, such as fish, reptiles, and amphibians, depend on favorable environmental factors for the nurture of their young; the offspring are often vulnerable to the vagaries of nature. Mammalian young do not require teeth for the suckling process and thus can be delivered with immature maxillae and mandibles, which facilitates the delivery of the head. The development of teeth coincides with the need to consume food other than milk.



• **Fig. 39.1** Diagram of a cluster of alveoli in the mammary gland of a goat. (From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)

Anatomical Aspects of the Mammary Gland

The Milk-Secreting Cells of the Mammary Gland Develop Through the Proliferation of Epithelium Into Hollow Structures Called Alveoli

Embryonic ectoderm is the source of the mammary glands. The mammary ectoderm is first represented by parallel linear thickenings on the ventral belly wall. The continuity of the ridge that is formed is broken into the appropriate number of *mammary buds,* from which the functional part of the mammary gland will be derived.

The *parenchyma,* or milk-secreting cells, of the mammary gland develops through the proliferation of epithelial cells that arise from the primary mammary cord. The epithelial cells eventually form hollow, circular structures called *alveoli,* which are the fundamental milk-secreting units of the mammary gland (Fig. 39.1). In concert with this development, an enlarged area of epithelium, the *nipple,* which is the external connection to the internal milk-secreting system, develops on the surface. In males, although nipples can develop, the underlying primary mammary cord does not normally develop into substantial glandular tissue.

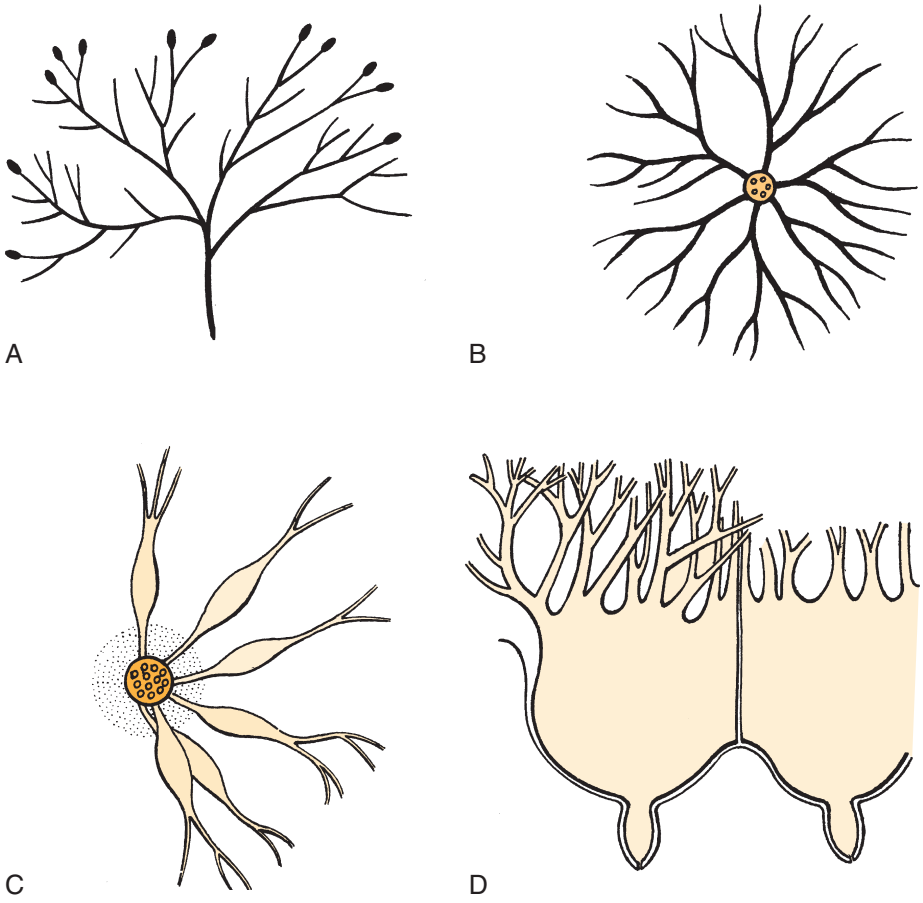
Most of the Milk That Accumulates Before Suckling or Milking Is Stored in the Alveoli, Even Though Animals Have Enlarged Milk-Storage Areas Called Cisterns

Duct systems connect alveoli with the nipple, or teat, enabling milk to pass from the area of formation to the area of delivery (nipple). The ducts may come together so that there is only one final duct per gland, which has one opening through the nipple, or teat, such as occurs in cattle, goats, and sheep. Two main ducts and associated openings occur in the mare and sow, whereas the cat and dog can have 10 or more openings in the nipple, with each opening representing separate glands (Fig. 39.2). Both the cow and the doe (goat) have specialized areas for holding milk, called *cisterns,* which are located in the ventral part of the gland and into which all main ducts empty (Fig. 39.3). This has enabled the cow, for example, to synthesize and store larger amounts of milk than would otherwise be possible. Despite this adaptation, it is important to realize that a majority of the milk present at the time of milking is stored in the duct system of the mammary glands.

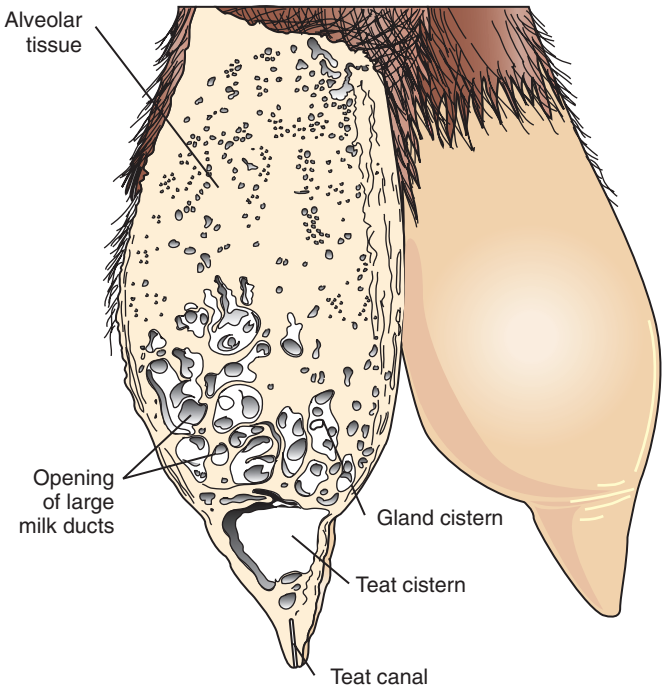
Mammary glands develop typically as paired structures. The number of pairs in domestic animals varies from one in goats, horses, and sheep; two in cattle; to seven to nine in the sow and seven to ten in the bitch and queen. Pairing (number per side and positioning) is not always matched in bitches. The position of mammary glands varies in animals, being thoracic in primates; extending the length of the thorax and abdomen in cats, dogs, and pigs; and being inguinal in cattle, goats, and horses. In domestic species, such as cattle, goats, horses, and sheep, pairs of mammary glands are closely apposed to each other; the resulting structure is called an *udder.* In the cow, for example, two pairs of glands (four quarters) compose the udder.

A Suspensory System Involving the Udder of the Cow Allows the Animal to Carry a Large Amount of Milk

One of the important anatomical adaptations of the udder that allows dairy cows to carry large amounts of milk is the development of a suspension system for the udder. This system is formed by the median suspensory ligament (formed between pairs of mammary glands) composed of elastic connective tissue that originates from the abdominal tunic. The lateral (nonelastic) suspensory ligament, which originates from prepubic and subpubic ligaments, enters the glands laterally at various levels to become part of the interstitial connective tissue framework of the udder. It is not unusual for heavy-producing dairy cows to have 25 kg (55 lb) of milk in their udder immediately before milking. If the suspensory support system were not in place, the mammary gland system would soon break down from the weight of the milk.



• **Fig. 39.2** Diagram showing different arrangements of the mammary duct system. (A) Cow, goat, sheep. (B) Horse, pig. (C) Cat, dog. (D) Cow, goat cistern. (From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)



• **Fig. 39.3** Depiction of the udder of a goat in which a section of the left mammary gland shows the dense alveolar tissues, the gland cistern with the large ducts opening into it, the teat cistern, and the teat canal. (From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)

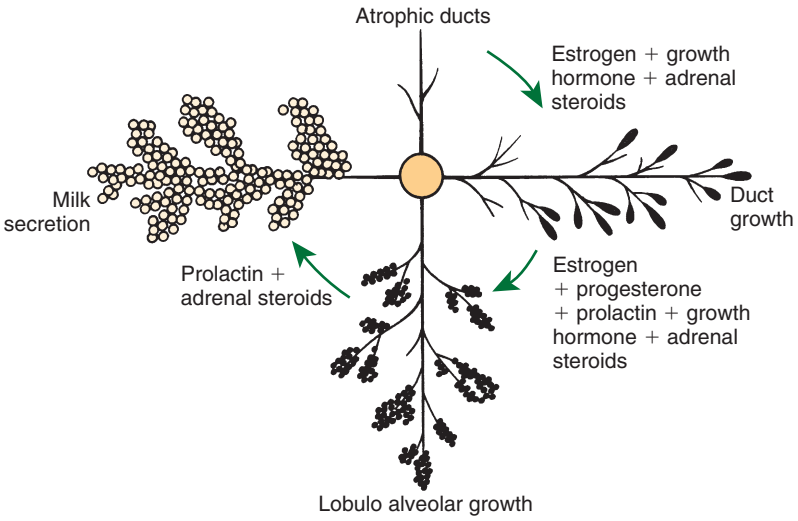
Control of Mammogenesis

Initial Development of the Mammary Gland Is Programmed by Embryonic Mesenchyme

The fetal development of the mammary gland is under both genetic and endocrine control. The initial development of the mammary bud is under control of embryonic *mesenchyme* (connective tissue). If mammary mesenchyme is transplanted to another area, mammary bud formation will occur at the site of transplantation. Although little is known about fetal mammary development, it is not thought to be driven by hormones. However, actively secreting mammary glands may be present at birth as a result of exogenous administration of certain hormones to the mother.

Proliferation of the Mammary Duct System Begins at Puberty, With Ducts Under the Control of Estrogens, Growth Hormone, and Adrenal Steroids, and Alveoli Under the Control of Progesterone and Prolactin

Development of the mammary gland in postfetal life usually starts in concert with puberty. Cyclical ovarian activity results in the production of estrogen and progesterone. Estrogen, with growth hormone (GH) and adrenal steroids, is responsible for proliferation of the duct system. The development of alveoli from the terminal ends of the ducts requires the addition of progesterone and *prolactin* (Fig. 39.4).



• **Fig. 39.4** Hormones involved in the growth of the mammary gland and in the initiation of milk secretion in the hypophysectomized-ovariectomized-adrenalectomized rat. (From Lyons WR. *Proc R Soc B.* 1958;149:303. From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)

Although the development of the mammary gland begins with the onset of puberty, the gland remains relatively undeveloped until the occurrence of pregnancy. In most domestic animals, udder development usually becomes evident by the middle of gestation; the secretion of milk often begins during the latter part of gestation (mainly from increasing *prolactin* secretion) and results in the formation of colostrum, as discussed later. By the end of pregnancy, the mammary gland has been transformed from a structure involving mostly stromal (connective tissue) elements to a structure filled with alveolar cells that are actively synthesizing and secreting milk. Groups of adjacent alveoli form *lobules* that further combine into larger structures called *lobes.* Connective tissue bands delineate the lobules and the lobes (Fig. 39.5).

Colostrum

Prepartum Milk Production (Without Removal) Results in the Formation of Colostrum

The milk formed before parturition is called *colostrum.* Its formation represents a secretory process in which lactogenesis occurs in the absence of milk removal. *Lactation* cannot fully blossom until pregnancy is terminated, however, because of the inhibitory effects of progesterone and estrogen on milk secretion, inhibitory factors that are removed at or just before delivery. Prolactin, a luteotroph, increases as progesterone levels decline late in gestation, promoting lactation.

The Ingestion of Colostrum Is Important Because of the Passive Immunity It Confers Through the Presence of High Concentrations of Immunoglobulins

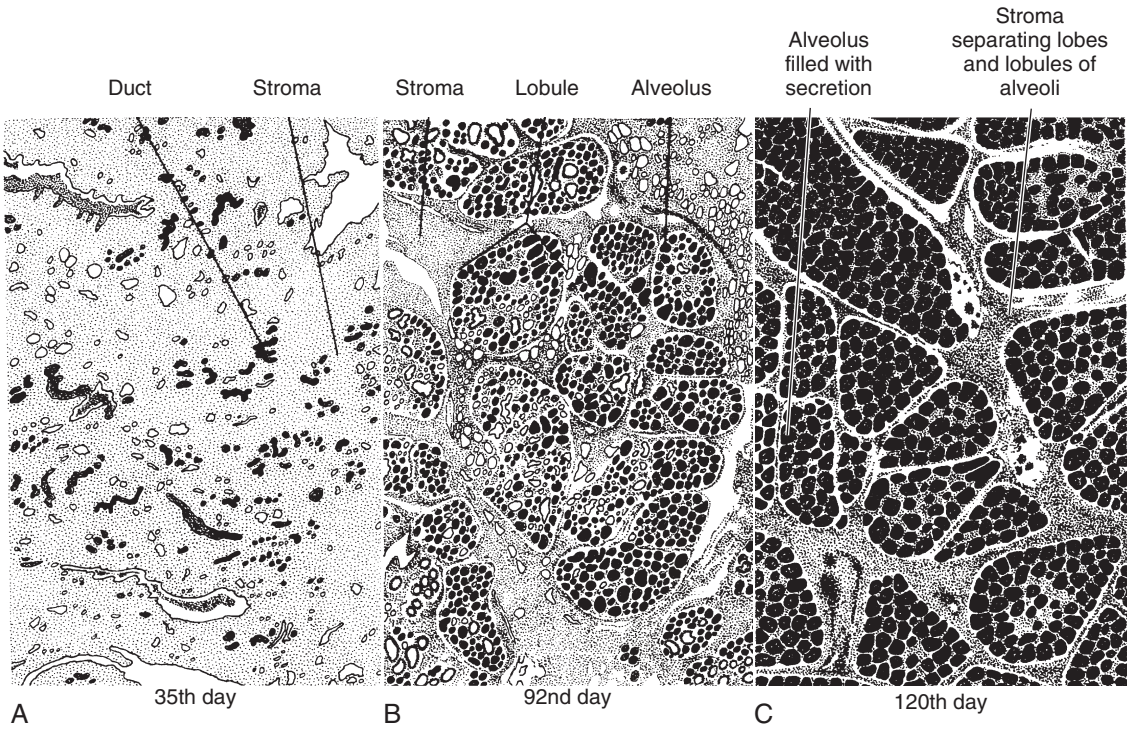
When colostrum is formed before parturition, certain substances are concentrated in the process. Ingestion of colostrum is important for the well-being of the neonate. In addition to nutrition, colostrum has an important function in temporary, or passive, protection against infectious agents. *Immunoglobulins* (e.g., immunoglobulin A, or IgA) are produced in the mammary gland by plasma cells (derived from B lymphocytes originating in the gut) as a result of exposure of the mother to certain microorganisms. The immunoglobulins gain access to the milk system through the migration of the plasma cells from adjacent tissue sites. The immunoglobulins are highly concentrated in colostrum, and through the consumption of colostrum, the neonate can receive passive immunity against pathogens experienced by the mother. This allows the young to receive immediate protection from environmental organisms. The neonates of all domestic animals acquire antibodies through the ingestion of colostrum. The absorption of antibodies through milk in domestic animals contrasts with the situation in other species, including humans, rabbits, and guinea pigs, in which a more substantial amount of antibody is passed to the fetus through the placenta.

The Time Immunoglobulins Can Be Absorbed Through the Neonatal Gut Is Limited to the First 24 to 36 Hours of Life

Neonates usually have a limited time (24–36 hours) in which immunoglobulins (proteins) can be absorbed through the gut. Thus the feeding of colostrum within this period is important to ensure the presence of immunoglobulins in the newborn. Other antimicrobial factors found in milk that are important for protection against the development of pathogenical enteric bacterial flora include lysozymes, lactoferrin, and the lactoperoxidase system.

Lipids (Particularly Vitamin A) and Proteins (Caseins and Albumins) Are High in Concentration in Colostrum; Carbohydrates (Lactose) Are Low

Colostrum is a rich source of nutrients, especially vitamin A, in addition to immunoglobulins. Placental transfer of vitamin A is limited in domestic animals, with calves and piglets being particularly low in vitamin A at birth. This deficiency is corrected by the ingestion of colostrum. Lipids and proteins, including caseins and albumins, are also present in relatively high concentration in colostrum. One exception is lactose; its synthesis is significantly inhibited by progesterone until about the time of delivery. Nevertheless, at the moment of delivery, the newborn's milk supply is nutritive (high protein, fat, and vitamin A content) and protective (immunoglobulins) (Table 39.1).



• **Fig. 39.5** Drawings of sections of the mammary gland of the goat at three different times during pregnancy (which lasts approximately 150 days). (A) Note the small collections of ducts scattered throughout the stroma on the 35th day. (B) On the 92nd day, the lobules of alveoli are forming in groups known as lobes; secretion is present in some of the alveolar lumina, and there is still considerable stromal tissue. (C) On the 120th day, the lobules of alveoli are almost fully developed; the alveoli are full of secretion, and the stromal tissue is reduced to thin bands separating lobules and thicker strands between lobes. (From Falconer IR, ed. *Lactation,* London, UK: Butterworths; 1970.)

Lactogenesis

Prolactin, Inhibited by Dopamine and Stimulated by Vasoactive Intestinal Peptide, Is the Most Important Hormone Involved in the Process of Milk Synthesis, or Lactogenesis; Growth Hormone Is Also Important for Lactogenesis

Prolactin plays an important role in the secretion of milk, or *lactogenesis.* Prolactin is released in conjunction with manipulation of the teat through either the suckling or the milking process. Sensory stimuli are carried into the hypothalamus, and the synthesis and release of *dopamine,* a major inhibitor of prolactin secretion, is blocked while neurons in the *paraventricular nucleus* are stimulated to produce and release *vasoactive intestinal peptide,* a stimulator of prolactin release (Fig. 39.6). A short-lived surge of prolactin secretion occurs immediately after the onset of milk removal; peak values are usually reached within 30 minutes after the initial stimulus. Major surges of prolactin apparently do not need to be elicited on an hourly basis to maintain lactation because 12-hour release intervals, as occur in association with the milking of dairy cows, are sufficient to maintain lactogenesis. Prolactin responses, as judged by the amount of hormone release after mammary gland stimulation, decrease as the lactation period progresses.

Another major hormone required for milk production in ruminants is *growth hormone* (GH). There is now considerable interest in the use of GH to promote additional milk production from cows through exogenous administration of the hormone.

The Release of Fat Into Milk From the Alveolar Cell Involves Constriction of the Plasma Membrane Around the Fat Droplet; Fats Are Dispersed in Milk in Droplet Form

The synthesis and release of milk by alveolar epithelial cells is a remarkable physiological process (Fig. 39.7). Alveolar cells synthesize fats, proteins, and carbohydrates, and extrude the products into the lumen of the alveolus. Fat droplets first accumulate in the basal cytoplasm of the cell and then move to the apex, where the droplet protrudes into the alveolar lumen. The cell membrane constricts about the base of the fat droplet, so fat is dispersed in milk in small droplets, surrounded by cell membranes; the droplet often contains portions of cell cytoplasm.

Milk Proteins and Lactose Are Released From Alveolar Cells by the Process of Exocytosis

Milk proteins are synthesized on the endoplasmic reticulum; the casein molecules pass to the Golgi apparatus, where they are phosphorylated and formed into micelles within the Golgi vesicles. Lactose is also synthesized within the Golgi vesicles and is released in conjunction with milk proteins. The process of extrusion of proteins and carbohydrates is different from that of fat; the Golgi

TABLE 39.1 Amounts of Selected Components of Bovine Colostrum as Percentage of Level in Normal Milk

	Day(s) After Parturition		
Constituent	0	3	4
Dry matter	220	100	100
Lactose	45	90	100
Lipids	150	90	100
Minerals	120	100	100
Proteins			
Casein	210	110	110
Albumin	500	120	105
Globulin	3500	300	200
Vitamins			
A	600	120	100
Carotene	1200	250	125
E	500	200	125
Thiamine	150	150	150
Riboflavin	320	130	110
Pantothenic acid	45	110	105

From Jacobson NL, McGilland AD. The mammary gland and lactation. In: Swenson MJ, ed. *Dukes' Physiology of Domestic Animals.* 10th ed. Ithaca, NY: Cornell University Press; 1984.

vesicles fuse with the cell membrane, and the release of proteins and carbohydrates occurs by *exocytosis.* Although it is not certain how often cells go through a synthesis and extrusion cycle, it may occur twice daily, particularly in dairy cows that are milked two times per day.

Milk Removal

For lactogenesis to be maintained, milk must be removed from the mammary gland by suckling or milking. If milk is not removed within about 16 hours in dairy cows, the synthesis of milk begins to be suppressed. As indicated previously, most of the milk in the udder of a dairy cow at the time of milking is located in the ducts and alveoli. The movement of milk into the gland cistern at suckling or milking would be slow, and less milk would be obtained during the milking of a cow if the drainage of milk were a passive process.

Efficient Milk Removal Requires the Release of Oxytocin, Which Causes Contraction of Muscle Cells That Surround the Alveoli (Myoepithelial Cells), and Movement of Milk Into the Ducts and Cisterns

To facilitate the process of milk removal, *myoepithelial cells* surround the alveoli and ducts (see Figs. 39.1 and 39.7). The myoepithelial cells are particularly responsive to oxytocin and, in fact, contract when exposed to the hormone. The synthesis and release of oxytocin from the posterior pituitary is elicited by a neuroendocrine reflex involving tactile stimulation of the udder by suckling by the young or the manual stimulation of washing before milking. The sensory stimuli from the udder are carried through the spinal cord into the hypothalamus. Neurons in the paraventricular and supraoptic nuclei are stimulated to synthesize oxytocin and release it from nerve terminals that impinge on the median eminence (Fig. 39.8). Other sensory stimuli that elicit oxytocin release include auditory, visual, and olfactory stimuli that occur near or within the kennel, cattery, or milking parlor. Past societies used various deceptions to have earlier breeds of cattle release their milk. They often allowed the calf to suckle one teat while they milked the other glands. They also knew about the *Ferguson reflex,* if not in name, in which stimulation of the cervix (and release of oxytocin) was elicited by blowing air into the vagina using hollow tubes.

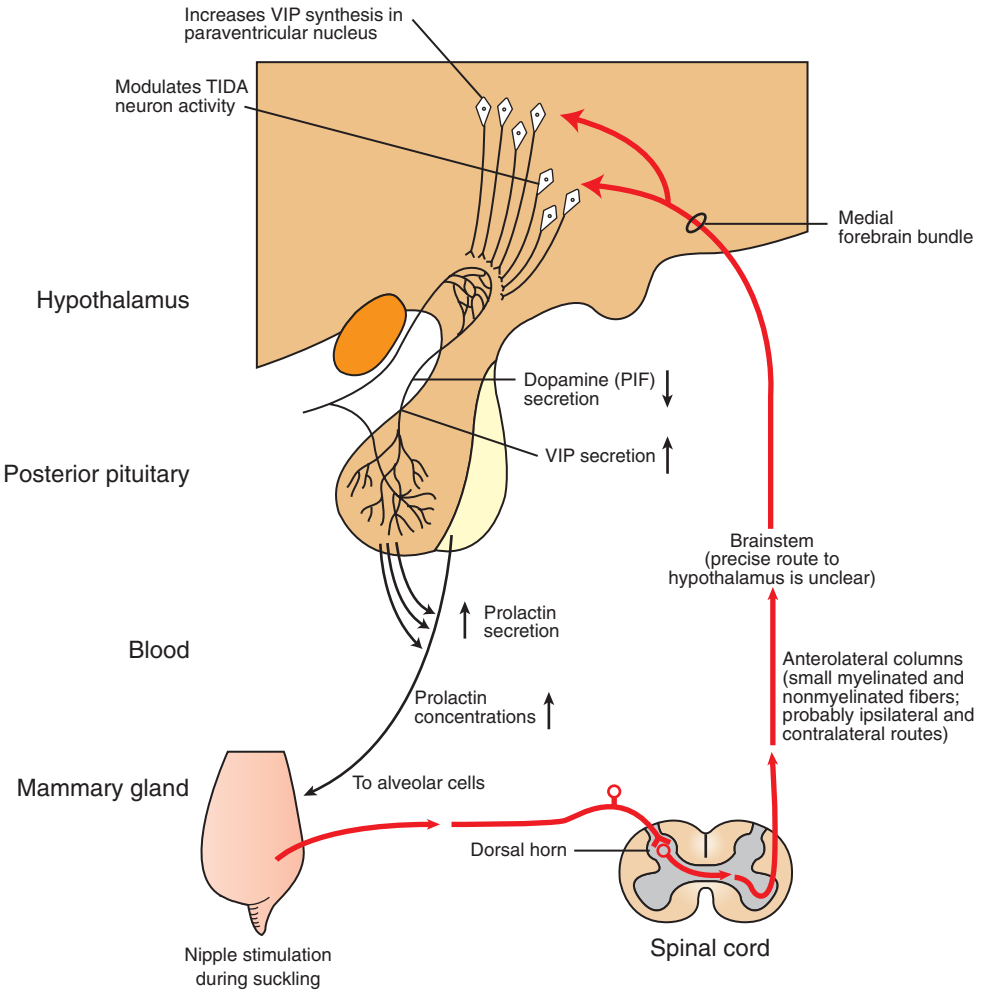
The release of oxytocin occurs within seconds after the stimulus arrives in the hypothalamus; increased pressure within the mammary gland is evident within a minute of stimulation as milk is forced out of the alveoli and ducts because of contraction of the myoepithelial cells. The term used in mammals to describe this phenomenon is *milk letdown.* Increased pressure within the udder is often obvious within a minute of the stimulation. The release of oxytocin lasts only a few minutes, and it is important that the milking process begin soon after milk letdown is complete (Fig. 39.9). The milking process, as done by machine or by hand in earlier times, is often completed within 4 to 5 minutes.

It is interesting to compare stimuli that release oxytocin, which initiates the passive part of lactogenesis, with stimuli that release prolactin, which directly influences lactogenesis. Any sensory stimulus that a female associates with nursing/milking has the potential for releasing oxytocin. The neuroendocrine reflex is elicited in the expectation of milk removal because of the environment (kennel, cattery, or milking parlor) to which the animal is exposed. Prolactin, on the other hand, is released only by tactile stimulation of the udder. The latter makes sense, because there is no need to stimulate milk synthesis and release unless the evidence for milk removal (udder stimulation) is strong. Milk removed during hand milking is trapped in the teat and forced out, whereas milk removed by milking machines moves by suction.

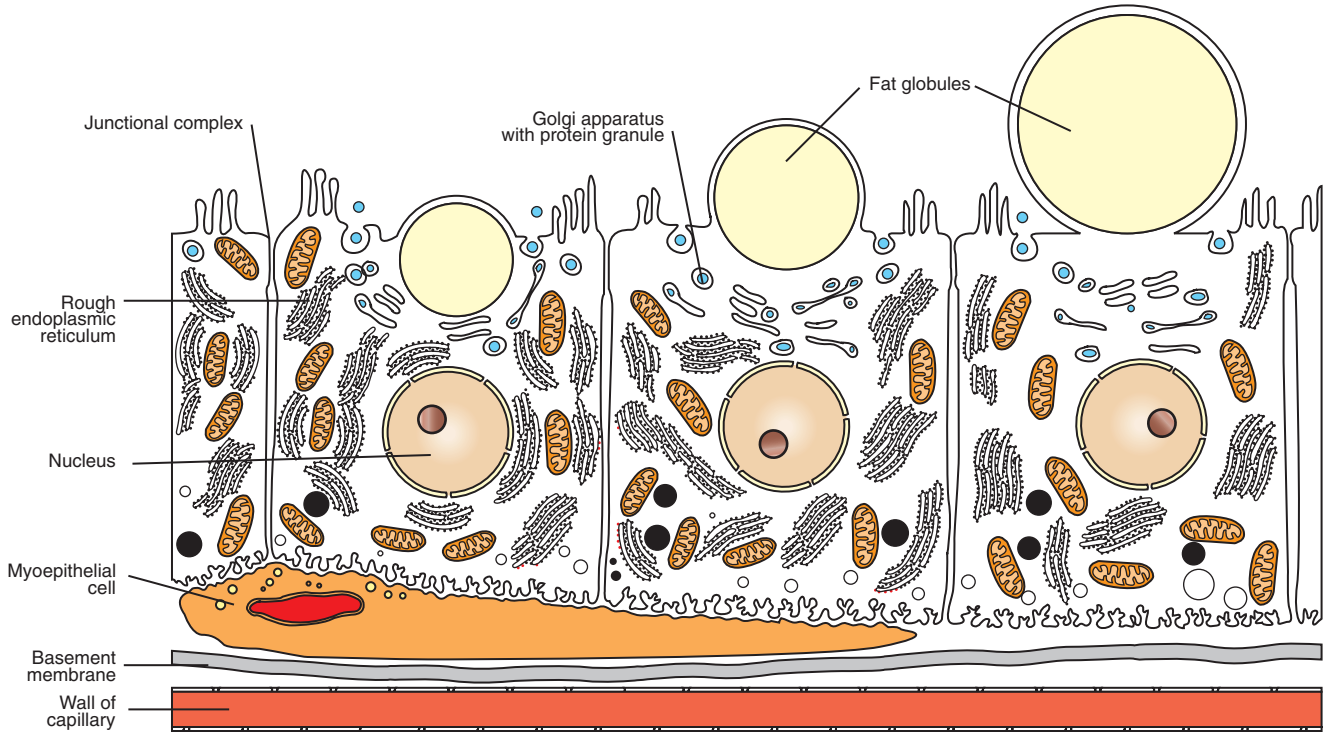
First Nursing

Carbohydrate Stores Are Good in Neonates Born as Singles or Twins, Whereas Carbohydrate Stores Are Low in Neonates Born in Litters; the Former Can Stand a Longer Interval to First Suckling Than Can the Latter

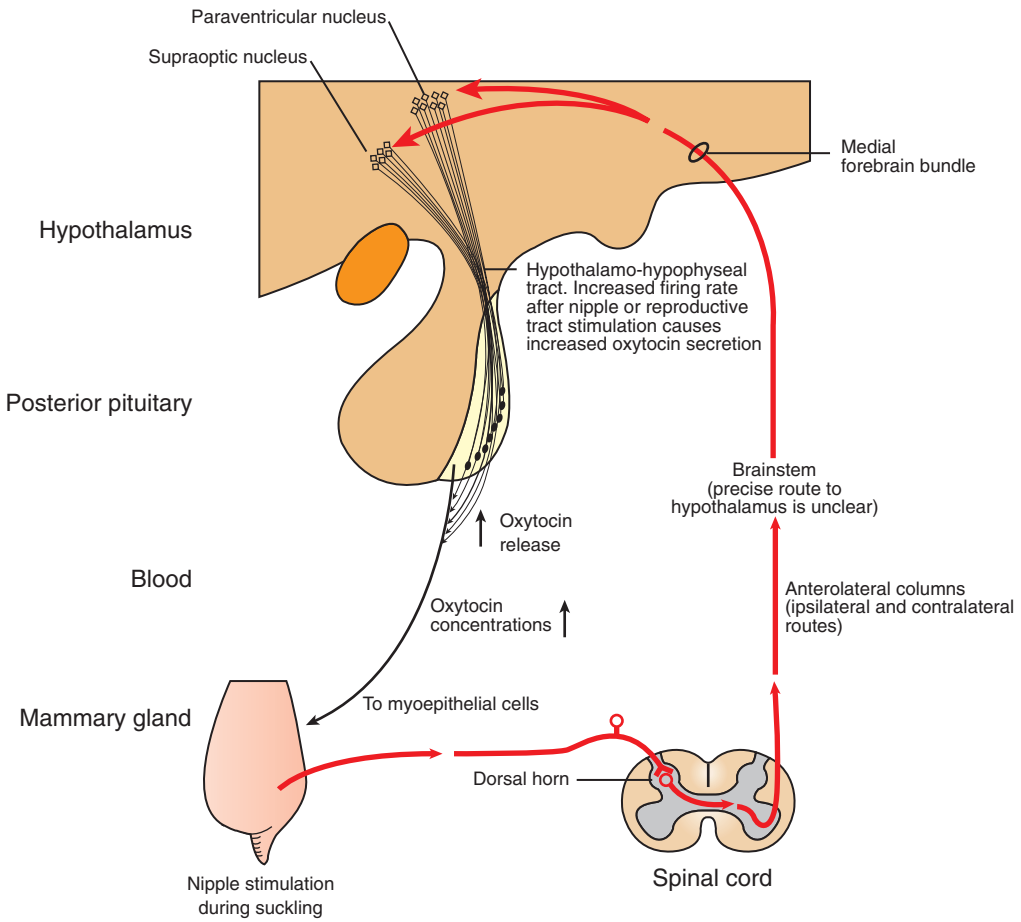
In domestic animals that have one or two offspring, such as cattle, horses, sheep, and goats, the young have to be able to stand to suckle. In this situation, neonates have reasonably good carbohydrate stores, and suckling may not occur for 1 to 2 hours without adverse effect as the young gain the ability to stand and locate the mammary gland. Young that are part of litters (cats, dogs, and pigs) are usually immediately nestled toward the mammary glands and often will be sucking in less than 10 minutes. Suckling continues interspersed with the delivery of additional littermates. This is important for animals born in litters because they tend to be immature at birth and susceptible to hypoglycemia, and suckling delays are often detrimental to their survival. Hypoglycemia results in stasis of the bowel (ileus) and can promote neonatal sepsis (overwhelming infection).



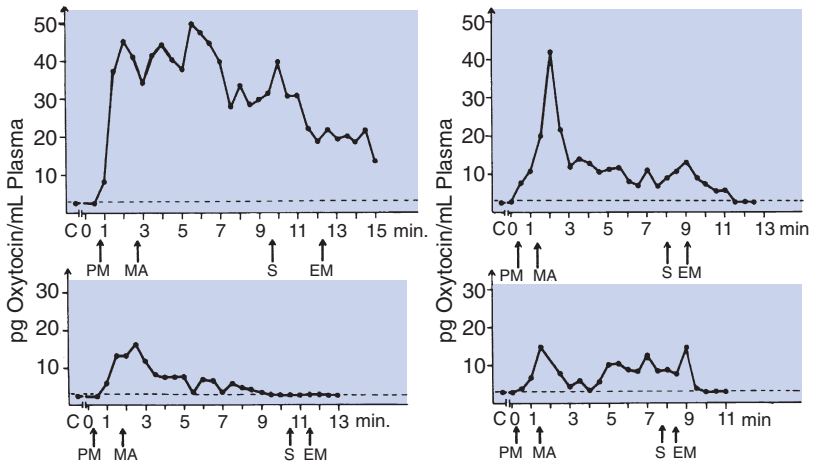
• **Fig. 39.6** Somatosensory pathways in the suckling-induced reflex release of prolactin. The exact route taken by sensory information between brainstem and hypothalamus is speculative. Although tuberoinfundibular (*TIDA*) dopamine neuron activity is modulated as a result of the arrival of this somatosensory-derived input (prolactin inhibiting factor, *PIF*), the increased secretory activity of neurons containing vasoactive intestinal peptide (*VIP*) in the paraventricular nucleus is probably also crucial in driving prolactin secretion during suckling. (Modified from Johnson M, Everitt B. *Essential Reproduction.* 3rd ed. London: Blackwell Scientific; 1988.)



• **Fig. 39.7** Diagram of the ultrastructure of three alveolar cells and a myoepithelial cell. (From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)



• **Fig. 39.8** Somatosensory pathways in the suckling-induced reflex release of oxytocin. The actual pathway of sensory input in the hypothalamus is unknown, but it probably involves the medial forebrain bundle. (Modified from Johnson M, Everitt B. *Essential Reproduction.* 3rd ed. London: Blackwell Scientific; 1988.)



• **Fig. 39.9** Oxytocin in the blood of cows before, during, and after milking. *Abscissae* show time in minutes. *C,* Control level; *EM,* end of machine milking; *PM,* preparation for milking; *MA,* application of teat cups; *S,* stripping. (From Schams, et al. *Acta Endocrinol.* 1979;92:258–270.)

The suckling interval during the neonatal period varies considerably among domestic animals. Species nursing litters, such as cats, dogs, and pigs, often nurse at intervals of 1 hour or less, especially during the first 2 weeks. Goats, horses, and sheep nurse at slightly longer intervals, often up to 2 hours. Rabbits are an exception regarding the time between suckling periods; their young nurse at 24-hour intervals. As can be imagined, baby rabbits are engorged after each suckling period.

Composition of Milk

Fats Are the Most Important Energy Source in Milk

Of the components of milk, fat is the most important energy source. Milk fat is composed of a number of lipids, including monoglycerides, diglycerides, triglycerides, free fatty acids, phospholipids, and steroids. Triglycerides are the main component of milk fat. The types of lipid synthesized are complex, with great variations in both chain length and saturation of fatty acids observed on the basis of species. The amount of fat produced varies greatly both within and among species (Table 39.2). Marine mammal milk has a high fat content, with values of about 40% to 50% in seals, 40% in dolphins, and 30% in whales. In these species, the high energy content of the milk through fat helps offset the heat loss of the young by promoting rapid adipose development.

In domestic animals, sheep, swine, dogs, and cats have milk that ranges from 7% to 10% in fat content. Dairy cattle have values that range from 3.5% to 5.5%; goats are similar to cows (3.5%); and mares have lower values (1.6%). In the past, milk was sold on a butterfat basis, and breeds that had a relatively high butterfat content of milk (e.g., the Jersey with 5% butterfat) found more acceptance in dairy operations than is currently the case. Small farms produced mainly cream (for butter manufacture); the fat-concentrated portion of milk was produced by use of a separator that separated cream on the basis of specific gravity and centrifugal force. Because milk is now sold on a solids, not fat, basis, breeds that produce more milk (and protein) are favored, even though the fat content of the favored breed, Holstein-Friesian, is lower (3.5%).

Lactose, Composed of Glucose and Galactose, Is the Main Carbohydrate of Mammalian Milk

Lactose is the main carbohydrate of most mammals. It is composed of glucose and galactose. Blood glucose is the main precursor molecule for lactose, with propionate an important precursor for glucose in ruminants. Lactose is formed under the direction of *lactose synthetase,* an enzyme composed of *α-lactalbumin* (a milk protein) and *galactosyl transferase.* Lactose synthesis is held in abeyance until immediately before term because progesterone is inhibitory for the formation of α-lactalbumin. Prolactin, on the other hand, is stimulatory for the formation of lactose synthetase. Animals must have the enzyme *lactase* present in the jejunum for lactose to be cleaved (to glucose and galactose) and used. Lactase is present in most mammalian young but is sometimes not present in adult animals, including humans. In the absence of lactase, lactose can have an osmotic effect in the gastrointestinal tract, which can lead to diarrhea.

TABLE 39.2 Composition of Milk From Various Species (Percentage)

Species	Fat	Protein	Lactose	Ash
Cat	7.1	10.1	4.2	0.5
Cow	3.5	3.1	4.9	0.7
Dog	9.5	9.3	3.1	1.2
Goat	3.5	3.2	4.6	0.8
Horse	1.6	2.4	6.1	0.5

From Jacobson NL, McGilland AD. The mammary gland and lactation. In: Swenson MJ, ed. *Dukes' Physiology of Domestic Animals.* 10th ed. Ithaca, NY: Cornell University Press; 1984.

The Main Proteins in Milk Are Called *Caseins* and Are Found in Curd

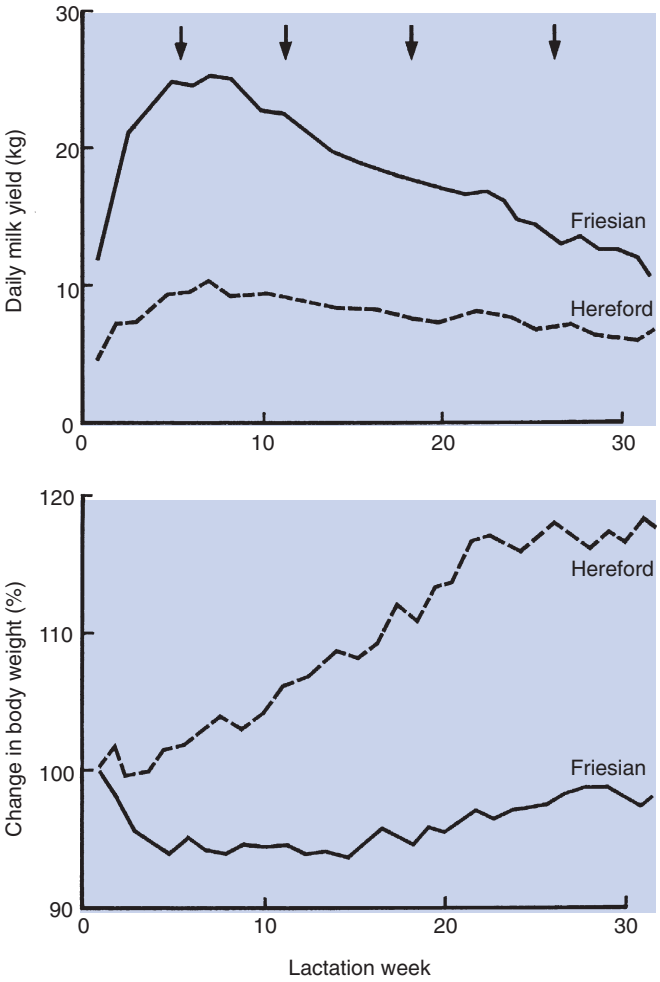
The main proteins produced by the alveolar cells are called *caseins.* Caseins can be removed (as a curd) from milk through a process called *curdling* or *coagulation,* with other milk proteins, such as albumins and globulins, remaining in the fluid part of the milk (whey).

The Lactation Cycle

The time required for changeover from colostrum to normal milk secretion varies with each species. In cattle, colostral milk tends to be stringy and yellow for several days postpartum. The complex bovine udder needs time for all areas to be flushed of colostrum. The milk of cattle is withheld from the milk supply for several days because of its unacceptable esthetic quality, not because of the basic quality of the milk. The bitch transitions to milk on day 2 to 3 postpartum.

Milk Production Peaks at 1 Month Postpartum in Dairy Cattle, Followed by a Slow Decline in Production; Milking Usually Stops at 305 Days of Lactation so That the Animal Can Prepare the Mammary Gland for the Next Lactation

Milk production tends to increase for the first 3 to 4 weeks of lactation and then begins to slowly decline through the end of lactation (Fig. 39.10). Cows are usually "dried up" after 305-day lactational periods; pounds of milk and butterfat production rates are calculated on this basis. Dairy animals are forced to stop lactating to prepare for the next lactation. The usual procedure is to stop milking. Nondairy animals wean neonates at the time deciduous dentition erupts, usually 4 to 5 weeks postpartum. The back pressure of milk within the alveoli gradually inhibits the secretion of milk by the alveolar epithelial cells, with a resultant regression of the alveolar cells and small ducts. The process, called *involution,* often requires at least a month in dairy animals, with a 6-week period usually desired as the minimal interval from drying off to the onset of the next lactational period. Within 1 to 2 months, the secretory (alveoli) and excretory (duct) systems regress and are once again replaced. The process by which epithelial structures regress, yet retain coding for the renewal of duct and alveolar systems, is truly remarkable.

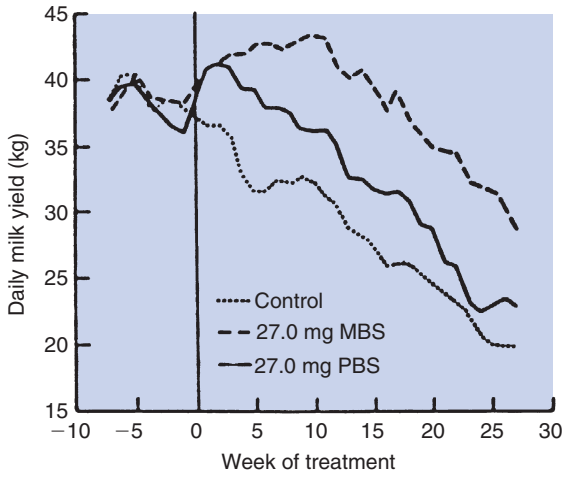


• **Fig. 39.10** Average daily milk yield (*top*) and average percentage change in body weight (*bottom*) in seven low-yielding cows (*broken line*) and eight high-yielding cows (*solid line*). *Arrows* indicate times of blood sampling. (Courtesy Dr IC Hart. From Cowie AT. Lactation. In: Austin CR, Short RV, eds. *Reproduction in Mammals.* 2nd ed. Vol 3. Hormonal control of reproduction. Cambridge, UK: Cambridge University Press; 1984.)

Lactation Can Be Induced by Hormone Administration (Estrogen and Progesterone) and Enhanced by Growth Hormone and Increased Photoperiod Exposure

The *induction of lactation* by hormone treatment is sometimes desired, especially in dairy animals with high-lactation records but poor reproductive performance. The use of a combined treatment of estrogen and progesterone over a relatively short period (1 week) has induced alveolar development sufficiently to result in milk production. Although the amount of milk produced is less than normal, the cows can be maintained in the milking string while efforts to impregnate them continue. To induce lactogenesis by hormonal means, animals should not be lactating at treatment and should have mammary glands free of infection.

GH, which is important to the normal lactational process, can be used for the enhancement of lactation when administered over a rather wide range of concentrations (Fig. 39.11). The ability to synthesize GH is relatively recent; its availability has expanded interest in its use for increasing the amount of milk produced by dairy cows. In general, GH acts on the post absorptive use of



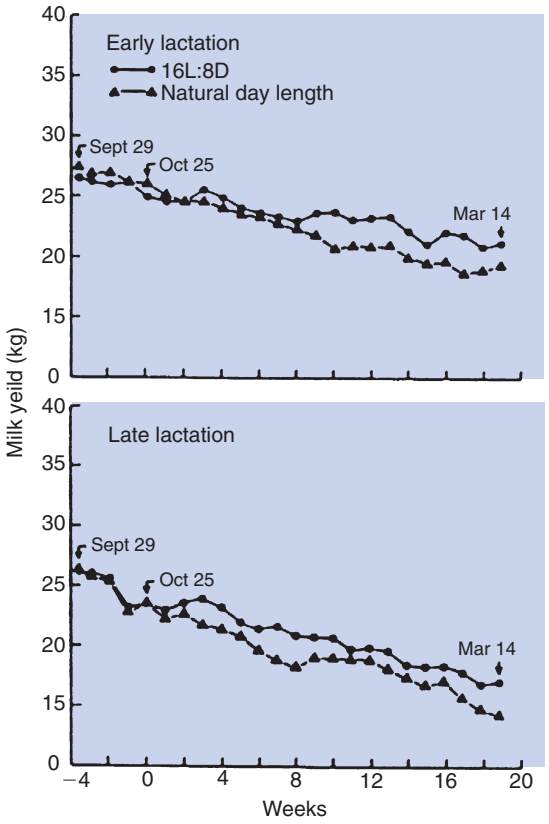
• **Fig. 39.11** Average weekly milk yield of cows injected daily with diluent (control), 27 mg of methionyl bovine somatotropin (*MBS*), or 27 mg of pituitary bovine somatotropin (*PBS*). Treatments began at week 0 at an average of 84 ±10 days after parturition. (From Tucker HA. Lactation and its hormonal control. In: Knobil E, Neill J, Ewing LL, et al., eds. *The Physiology of Reproduction.* Vol 2. New York: Raven Press; 1988.)

nutrients so that protein, fat, and carbohydrate metabolism in the whole body are changed, and the nutrients are directed toward milk synthesis. If cows are in early lactation and in a negative energy balance, GH administration results in the mobilization of body fats that are used for milk formation. If cows are in positive energy balance, GH has no effect on the metabolism of body fat. Initially, GH treatment decreases the energy balance of cows; however, this is adjusted by a voluntary increase in feed consumption. Despite the increased feed consumption, GH administration increases the gross efficiency of lactation by as much as 19%. In essence, the effects of exogenous GH do not depend on gross alterations in nutrient digestibility or on body maintenance requirements. The use of GH may be economically viable, with the increased milk production justifying the expense of the hormone.

An interesting controversy has arisen from the fact that cows treated with GH do not produce "organically" derived milk, despite that synthetic GH is almost identical to endogenously derived GH. Although there is no evidence that increased concentrations of GH occur in the milk as a result of its administration, some consider the resultant milk to be abnormal.

The results with GH are in contrast to the studies in which thyroid hormone administration, in the form of *iodinated casein* (thyroprotein), was used to increase lactation in cows. Although the administration of thyroprotein increased lactation, extra feed was necessary to prevent excessive body weight loss, and milk production declined abruptly when thyroprotein was removed from the diet. In essence, the use of thyroprotein does not affect the efficiency of the lactational process as GH does. In dogs, one differential for gynecomastia (mammary enlargement) is profound hypothyroidism, causing elevated thyrotropin-releasing hormone (TRH) levels, which in turn stimulate prolactin secretion.

Another interesting finding concerning the manipulation of lactation has been that the milk yield in cows can be increased by exposing them to increased light (called a *photoperiod*). Cows under a photoperiod regimen of 16 hours of light (8 hours of dark) produced 6% to 10% more milk than animals under the reverse photoperiod (8 hours of light and 16 hours of dark) (Fig. 39.12). Although the mechanism by which light affects lactation is not known, it likely involves prolactin secretion, at least to some extent,



• **Fig. 39.12** Influence of day length on milk production of Holstein cows. Between September 29 and October 24, cows at 37 to 74 days (early lactation) or 94 to 204 days (late lactation) after parturition were exposed to natural photoperiods of 12 hours of light per day and standardized diets. Between October 25 and March 14, cows were exposed to natural photoperiod (9–12 hours of light daily) or to 16 hours of fluorescent lighting superimposed on the natural photoperiod. *D,* Dark; *L,* light. (From Tucker HA. Lactation and its hormonal control. In: Knobil E, Neill J, Ewing LL, et al., eds. *The Physiology of Reproduction.* Vol 2. New York: Raven Press; 1988.)

in that increased light exposure results in increased prolactin secretion. There is evidence that entrainment of physiological processes to photoperiod often involves the suprachiasmatic nucleus of the hypothalamus, often called the body's "master clock nucleus". The queen's estrous cycle is affected by photoperiod, mediated by melatonin and prolactin levels. Melatonin and prolactin secretion may play a role in ovarian function in the cat, with lower levels of both hormones present during estrus than during the interestrous period. Protocols exist for inducing improved lactation in postpartum bitches using low-dose oxytocin and metoclopramide (a dopamine D-2 receptor antagonist).

Diseases Associated With the Mammary Gland

The Main Diseases That Affect the Mammary Gland Directly Are Mastitis (Prevalent in Dairy Cattle and Dogs) and Neoplasia (Prevalent in Intact Dogs and Cats)

The most important problems involved in the production of milk are those caused by inflammation of the gland (*mastitis*). One fundamental cause of mastitis is injury to the teat canal from the repeated stretching that occurs with the milking process. Organisms that ordinarily would be excluded from the gland are able to make their way past the barrier located within the teat canal; with repeated microorganism exposure, an infection is established.

One of the adverse consequences of mastitis is the formation of connective tissue within the udder as a result of the attempt of the gland to wall off the infection. The presence of connective tissue limits the area into which ducts and alveoli can proliferate, thus reducing the milk-producing potential of the gland. The mammary gland is an example of an organ (the eye is another example) in which the elicitation of an inflammatory response is often detrimental to the function of the organ. Thus therapies directed toward the treatment of mastitis often combine antiinflammatory and antibacterial agents.

Another process that disturbs the structure of the mammary gland is *neoplasia.* In domestic animals, the dog is most susceptible to the occurrence of mammary tumors. The exposure of the mammae to the ovarian hormones estrogen and progesterone greatly increases the chance of neoplasia. The incidence of mammary tumors is very low if the dog is ovariectomized before the first estrous cycle, but it increases progressively through exposure to more than two ovarian cycles; ovariectomy has little effect on neoplasia if done after the third or fourth cycle. Some owners want their dogs to have one or two cycles before they are ovariectomized. It is important for veterinarians to point out the beneficial aspects of ovariectomy before the second heat cycle because of the incidence of mammary neoplasia and the usual benefits of fertility and behavioral control.

The Main Conditions That Involve the Mammary Gland /Indirectly Are Passive Transfer of Red Blood Cell Agglutinating Antibodies by the Ingestion of Colostrum (Mare, Queen) and Hypocalcemia Caused by the Transient Drain of Calcium That Occurs With Initiation of Lactation (Dairy Cattle) or During the Perinatal Period (Dog)

An immunological disease associated with the mammary gland involves the transfer of red blood cell (RBC) agglutinating antibodies to the neonate through the milk. The situation is most common in the horse, in which fetal RBCs pass into the maternal system and elicit antibody formation against the fetal RBCs. These antibodies tend to be concentrated in the colostrum along with other immunoglobulins. At birth, the foal is able to absorb the RBC agglutinating antibodies (and other beneficial immunoglobulins) for up to 48 hours. Foals often go into a hemolytic crisis between 24 and 48 hours after delivery and can die unless given vigorous therapy, including blood transfusions. If fetal RBC antibody formation is suspected in a mare, the disease can be handled by muzzling the foal at birth through 48 hours and feeding with colostrum saved (frozen) from other preparturient mares. A similar condition has been reported in group A blood type kittens (usually purebred) born to group B blood type queens bred to toms with type A blood type. Unfortunately, no feline colostrum is commercially available, but kittens can be given serum or plasma from another blood type A queen for immunoglobulin transfer, while being prevented from nursing from their dam.

Diseases associated with the mammary gland and life-threatening to the dam are *hypocalcemia, puerperal tetany,* or *eclampsia.* At

parturition, the acceleration of lactogenesis causes a great increase in the movement of calcium from the blood into the milk. Both cows and dogs are particularly susceptible, with some dams unable to respond immediately to the calcium drain from the blood by the mobilization of calcium. As a result, the animals lose their ability to maintain normal muscle activity; cows are often unable to stand and become prostrate with the appearance of being comatose. Bitches develop tremors progressing to tetany. The syndrome occurs in cows at parturition and in dogs during the last weeks of pregnancy or the first few weeks postpartum, when lactation reaches its peak. Inappropriate prenatal nutrition, often with calcium supplementation, sets bitches up for this condition by inhibiting normal parathyroid gland development necessary to meet the demands for mobilization of calcium by lactation. The systemic administration of calcium to hypocalcemic cows or bitches often produces a dramatic recovery in 10 to 20 minutes.

CLINICAL CORRELATIONS

Pregnant Mare That Does Not Have Sufficient Milk or Udder Development

History

A 13-year-old pregnant Arabian mare is due to foal in the next week, based on breeding dates. This is her second foal; the first foal did well. She has had limited udder development. The mare is current on vaccinations and deworming; she has no previous medical problems. She has been on pasture with all the other horses. She is supplemented with 2 pounds of 14% mare and foal feed per day. She is given grass hay, and she is in good body condition. The farm is located in eastern Tennessee.

Clinical Examination

The mare is pregnant. The foal is viable based on movement and heartbeat. The foal is in the pelvic inlet, indicating that the mare will foal soon. All other parameters of the examination are normal. The mare is in good body condition.

Comment

With a mare close to parturition that has had limited udder development, one concern is whether the mare has been fed fescue. The owners are asked about this and say that she has been receiving a grass hay with fescue in it. They did not know that fescue could cause problems. The clients are informed that fescue can contain an endophytic fungus, *Neotyphodium coemophialum,* which produces alkaloid toxins. These toxins are dopaminergic and inhibit prolactin. Additionally, the increase in dopamine activity directly decreases prolactin. Normally, neurons in the paraventricular nucleus would induce prolactin release, which would stimulate lactogenesis. In some animals, progesterone levels are also decreased.

Treatment

Domperidone can be given because it inhibits the dopaminergic effects. If domperidone is started 5 to 10 days before parturition, this is often sufficient time for the mare to develop adequate milk for the foal. If domperidone is not started until up to 24 hours after foaling, it must be given for 10 to 14 days. Some mares will respond and will increase the milk production, whereas some mares will not respond. Although this treatment is often useful in stimulating milk production by the mare, it does not necessarily reverse other complications associated with fescue. Mares that have been given fescue may have foals that appear dysmature, are weak, or have prolonged gestation. The placenta from these affected mares is often thickened. To prevent these effects of fescue, clients are encouraged to keep the mares off fescue-containing feeds for at least 30 to 60 days before foaling. In addition, endophyte-free fescue grasses are available, but expensive.

Neonatal Isoerythrolysis

History

You are called to examine a mare, 7 months pregnant, that has a previous history of having conceived and delivered a normal foal after her first pregnancy; the foal was subsequently suckled and was sold as a weanling. The mare had no trouble conceiving and carrying the next two pregnancies, but the foals died within 2 to 4 days of birth, even though they were healthy and vigorous at birth and the mare had colostrum and milk. The previous owner became discouraged because of these deaths and sold the mare to the current owner at a bargain price.

Clinical Examination

You perform a general physical examination of the mare and find all organ systems to be normal. Palpation of the uterus per rectum reveals the presence of a viable fetus that appears to be of the correct size for a pregnancy of the purported duration. Both the external genitalia and the mammary glands are normal in appearance.

Comment

From the history, and because the mare appears to be undergoing a normal pregnancy, you conclude that there is nothing wrong with the reproductive process. The fact that the previous two foals were healthy at birth and yet weakened rapidly and died within 4 days indicates that something likely happened to them after delivery. If the deaths were caused by an issue associated with the mare, the most likely cause of these deaths would be *neonatal isoerythrolysis.* In this situation, the mare becomes exposed to the RBCs (erythrocytes) of the fetus during pregnancy, or the mare could have been exposed to RBCs from a stallion whose erythrocytes are recognized as foreign. If the mare is exposed to the RBCs from the stallion or the fetus, those RBCs enter the circulation of the dam. She responds by making antibodies to the RBCs because of the presence of foreign antigen on the fetal erythrocytes that were inherited from the sire. In the mare, these antibodies do not pass through the placental barrier, so the fetus is protected from these antibodies during pregnancy. The antibodies do pass into the colostrum and are concentrated during the process of colostrum formation. Therefore when the foal suckles the colostrum, it acquires the antibodies that will react to its own RBCs. The foal develops a type II hypersensitivity reaction in which the antibodies destroy the foal's RBCs through different mechanisms.

Treatment

The foal needs to be prevented from suckling the mare for the first 2 to 3 days of life. During the first 1 to 2 days, the foal is able to absorb large protein molecules, including the important immunoglobulins that enable the foal to ward off infections, and in this case, antibodies against fetal RBC antigens. The gut epithelium closes to the passage of large protein molecules by 36 to 48 hours of life; at this time, or shortly thereafter, the foal can be allowed to suckle without risk for absorbing the antibodies. The key is to prevent the foal from suckling during the first 2 to 3 days of life to prevent absorption of the antibodies that will react to its own RBCs. The mare needs to be monitored closely before parturition so that the foal can be muzzled shortly after delivery. The foal does need nourishment during the first 2 to 3 days of life; thus it is important that the foal be fed colostrum obtained from other mares (usually maintained frozen). If the foal does receive the mare's colostrum containing antibodies that will react to its own RBCs, the foal can still be treated. The foal should not nurse the mare for the first 3 to 5 days of life, and the foal can be given packed RBCs from a donor if necessary. To anticipate the potential risk, blood typing of stallions and mares is now available.

Practice Questions

1. The development of the duct system in the mammary gland is under the control of estrogens, growth hormone, and adrenal steroids. If the duct system is to develop functional milk-secreting units, called *alveoli,* which of the following hormone(s) are essential to this development?
a. Progesterone
b. Prolactin
c. Relaxin
d. Prolactin and progesterone
e. Prolactin and relaxin
f. Progesterone and relaxin

2. The hormone that is most important for the maintenance of lactation (lactogenesis) is:
a. Estrogen
b. Oxytocin
c. Progesterone
d. Prolactin
e. Relaxin

3. Sensory inputs (including sound, sight, and smell, but not necessarily touch) elicit the release of what important hormone required for the lactation process in the cow?
a. Estrogen
b. Oxytocin
c. Progesterone
d. Prolactin
e. Relaxin

4. The contraction of what anatomical structure is of fundamental importance for the release of milk from the udder of the cow?
a. Alveoli
b. Duct
c. Myoepithelial cell
d. Duct cistern
e. Teat cistern

5. The most important energy source in milk is:
a. Carbohydrates
b. Lactose
c. Lipids
d. Proteins

Bibliography

Bogaerts P. *Clinical approach to genital and mammary pathologies in cats*. Presented at 5th Biannual Congress, *European Veterinary Society for Small Animal Reproduction (EVSSAR),* Budapest, Hungary, 2006.

Capuco AV, Wood DL, Baldwin R, Mcleod K, Paape MJ. Mammary cell number, proliferation, and apoptosis during a bovine lactation: relation to milk production and effect of bST1. *J Dairy Sci*. 2001;84(10): 2177–2187.

Claus MA, Levy JK, MacDonald K, Tucker SJ, Crawford PC. Immunoglobulin concentrations in feline colostrum and milk, and the requirement of colostrum for passive transfer of immunity to neonatal kittens. *J Feline Med Surg*. 2006;8(3):184–191.

England GC, Heimendahl AV. *BSAVA Manual of Canine and Feline Reproduction and Neonatology*. British Small Animal Veterinary Association; 2010.

Park CS, Lindberg GL. The mammary gland and lactation. In: Reece WO, eds. *Dukes' Physiology of Domestic Animals*. 12th ed. Ithaca, NY: Comstock Publishing; 2004.

Peterson ME, Kutzler MI, eds. *Small Animal Pediatrics: the First 12 Months of Life*. Philadelphia: Saunders; 2011.

Plant TM, Zeleznik AJ, eds. *Knobil and Neill's Physiology of Reproduction*. Vol. 1 and 2. 4th ed. London: Academic Press; 2015.

Scantlebury M, Butterwick R, Speakman JR. Energetics of lactation in domestic dog (Canis familiaris) breeds of two sizes. *Comp Biochem Physiol A Mol Integr Physiol*. 2000;125(2):197–210.

40

Reproductive Physiology of the Male

JUAN E. ROMANO AND STEVEN P. BRINSKO

KEY POINTS

Functional Anatomy

1. The male reproductive system consists of many individual organs acting in concert to produce spermatozoa and deliver them to the female's reproductive tract.
2. Normal spermatogenesis requires maintenance of uniform testicular temperature 2°C to 6°C lower than core body temperature.
3. Emission is the release of spermatozoa and accessory gland fluids into the pelvic urethra, whereas ejaculation is the forceful expulsion of semen from the urethra.

Spermatogenesis

1. Spermatogenesis is a lengthy orchestrated process in which diploid stem cells divide by mitosis to maintain their own numbers and cyclically produce progeny that undergo meiotic division and differentiation into haploid germ cells.
2. Testicular size can predict daily sperm production.

Hypothalamic-Pituitary-Testicular Axis

1. The reproductive system of the male is regulated by the hypothalamus, which is hormonally linked to the anterior pituitary and testes by luteinizing hormone and follicle-stimulating hormone.

Puberty

1. Puberty is not synonymous with sexual maturity.
2. Puberty results from a continuous process of endocrine changes initiated shortly after birth.

Anabolic Steroids

1. Anabolic steroids are androgen derivatives that exert negative feedback on the hypothalamic-pituitary-testicular axis.

Functional Anatomy

The Male Reproductive System Consists of Many Individual Organs Acting in Concert to Produce Spermatozoa and Deliver Them to the Female's Reproductive Tract

The male reproductive system is made up of a number of individual organs acting in concert to produce spermatozoa and deliver them to the reproductive tract of the female. This concerted effort involves both the neuroendocrine (hypothalamus and anterior pituitary glands) and the genital system. The genital organs consist of two testes, each suspended within the scrotum by a spermatic cord and external cremaster muscle; two epididymides; two deferent ducts; accessory sex glands; and the penis. The accessory sex glands include paired ampullae, paired seminal vesicles (vesicular glands), a prostate gland, and paired bulbourethral glands (Cowper glands). The presence of individual accessory glands, the testicular orientation, the type of penis, and the site of semen deposition in the female are dependent on the species (Table 40.1).

Normal Spermatogenesis Requires Maintenance of Uniform Testicular Temperature 2°C to 6°C Lower Than Core Body Temperature

Normal *spermatogenesis* in most mammals is dependent upon maintenance of uniform testicular temperature between 2°C to 6°C lower than core body temperature. Elevated testicular temperature reduces the numbers of live, normal spermatozoa. Thermoregulation of the testes in domestic animals is maintained by the pendulous scrotal sac, the testicular vasculature, the dartos and cremaster muscles, and the scrotal skin. A pendulous scrotum facilitates thermoregulation by using several mechanisms such as conduction, convection, and evaporation. The internal spermatic artery of many mammals is highly convoluted and, in farm animals, the coiling is so extensive that the artery forms a vascular cone on the dorsal pole of the testis. The testicular vascular cone is composed of a venous pampiniform plexus network surrounding the highly coiled testicular artery. This ramification reduces the mean blood pressure and permits transfer of heat from the testicular arteries (high temperature) to the veins (low temperature) by a countercurrent heat-exchange system. This mechanism of heat exchange is possible because the spermatic artery is extensively coiled and in close proximity to the venous pampiniform plexus. In addition, there are periarterial veins and arteriovenous shunts that facilitate heat transfer as well as the transfer of hormones such as testosterone from the veins to the arteries. The dartos and cremaster muscles can increase or reduce the exposure surface area of the scrotum and move the testes closer to or farther from the abdomen, depending on their state of contracture. The scrotal skin is usually thin, generally lacks subcutaneous fat, has relatively little hair or wool, and contains numerous sebaceous and sweat glands. The blood and lymphatic system in the scrotal skin is very extensive, with blood vessels near the skin surface, facilitating radiation of heat. In hot environments, the blood flow in the scrotal skin increases and the evaporation per unit area of scrotal skin is greater than the evaporation from the general body surface. In the scrotum, the number and volume of the sweat glands per unit skin surface

TABLE 40.1 Male Reproductive Parameters

	Bull, Buck, and Ram	Stallion	Boar	Dog	Tom	Llama/Alpaca
Testis orientation	Vertical cauda down	Horizontal	Perineal cauda up	Horizontal	Perineal cauda up	Perineal cauda up
Ampullae	+	+	–	+	–	+
Vesicular gland	+	+	+	–	–	–
Bulbourethral gland	+	+	++	–	+	+
Prostate	+	+	+	+	+	+
Penis type	Fibroelastic sigmoid	Vascular	Fibroelastic sigmoid	Vascular	Vascular	Fibroelastic sigmoid
Semen deposition	Vagina	Uterus	Cervix/uterus	Vagina	Vagina	Uterus

is greater than other body regions. In addition, the scrotal skin has thermoreceptors that trigger a local and a systemic response in the presence of an increase in local temperature. Locally, the blood flow and scrotal sweating will increase. The systemic response will increase the number of breaths per minute (polypnea).

As previously mentioned, in domestic mammals, normal testicular function, especially normal spermatogenesis, is temperature dependent and requires an environment lower than core body temperature. Therefore in normal domestic males, the testes are located outside the abdominal cavity, in the scrotum. Failure of one or both of the testes to descend into the scrotum is known as *cryptorchidism.* Although the cryptorchid testis is still hormonally active and thus capable of producing *androgens,* it is incapable of producing normal spermatozoa. Consequently, a bilaterally cryptorchid male would be sterile. The cryptorchid testis is more prone to torsion of the spermatic cord and 10 times more likely to be neoplastic. Cryptorchidism appears to be genetic, although the exact mechanism is not completely understood and may vary among species. It is most common in boars, dogs, and stallions and least common in bulls, rams, and bucks. Descent of the testes into the scrotum normally occurs in domestic animals during the following time periods:

- *Horse*, 9 to 11 months of gestation
- *Cattle,* 3.5 to 4 months of gestation
- *Sheep,* 80 days of gestation
- *Pig,* 90 days of gestation
- *Dog,* 5 days after birth
- *Cat,* 2 to 5 days after birth
- *Llama/alpaca,* Usually present at birth

For the majority of domestic species, passage of the testes through the internal rings by 2 weeks after birth is necessary for a final scrotal position to occur. Many animals may have testes in the inguinal region at birth, and the testes may remain there for weeks or months before descending into the scrotum. In the dog, testicular descent is uncommon after 14 weeks of age and does not occur after 6 months of age. In the stallion, although it is considered abnormal, descent of inguinally retained testes has been known to occur as late as 2 to 3 years of age. Although considered unethical, or at the very least, poor practice, sometimes the cryptorchid testis is not removed during routine castration, resulting in objectionable behavior after the animal reaches puberty. The clinical test for suspected cryptorchidism typically involves acquisition of two blood samples, one before and one after the administration of human chorionic gonadotropin (hCG), to measure the change in circulating testosterone levels. Recently, a single blood sample to measure circulating levels of anti-müllerian hormone (AMH) has been shown to be useful in the diagnosis of cryptorchidism. Also known as müllerian-inhibiting substance, AMH is a homodimeric disulfide-linked glycoprotein hormone that belongs to the transforming growth factor-β superfamily produced exclusively by Sertoli cells. Measurable amounts of AMH are indicative of the presence of testicular tissue. Therefore serum concentrations of AMH are significantly higher in cryptorchid than in castrated animals, in which AMH concentrations are low to undetectable.

The testis is the pivotal organ of the male reproductive system. It must be remembered, however, that all testicular functions are profoundly influenced by the neuroendocrine system. The testis is responsible for steroidogenesis, primarily the production of androgens, as well as the generation of haploid germ cells by spermatogenesis. These two functions occur in the Leydig cells and the seminiferous tubules, respectively.

Functionally, the testis is considered to have three compartments. The interstitial tissue compartment, containing the Leydig cells, surrounds the seminiferous tubules and bathes them with testosterone-rich fluid. The other two compartments reside within the seminiferous tubules. The basal compartment contains spermatogonia, which divide through mitosis, whereas the adluminal compartment represents a special environment where spermatocytes undergo meiosis and continue their meiotic divisions to differentiate into spermatids and finally into spermatozoa. Within the seminiferous tubules, the Sertoli cells, which provide support and nourishment to the developing germ cells, extend from the basal compartment into the adluminal compartment. Tight-junctional complexes between the Sertoli cells separate the basal and adluminal compartments and form the major component of the blood-testis barrier, which functionally prevents many compounds found in the blood and interstitial fluid from entering the adluminal compartment.

The seminiferous tubules empty their contents into the rete testis, which subsequently transports the spermatozoa and seminiferous tubular fluid into the epididymis. The epididymis is a single tortuous duct of considerable length (from 2 m in the cat to 80 m in the stallion) anatomically divided into three segments: head, or caput; body, or corpus; and tail, or cauda. The epididymis not only is a conduit for spermatozoa but also provides a special environment in which spermatozoa are concentrated, undergo maturation, and acquire fertilizing capacity. Spermatozoa that enter the caput from the rete testis are immotile and incapable of fertilization. Only after they undergo migration and maturation through the caput and corpus are both motility and the capacity for fertilization achieved. The cauda epididymis and the deferent duct, into which the cauda empties, serve as a storage depot for mature spermatozoa; together, these are known as the extragonadal sperm reserves. The spermatozoal transit time through the caput and

corpus epididymis is not altered by *ejaculation* and is similar (2 to 5 days) for domestic species. Storage time in the cauda epididymis is more variable among species (3 to 13 days) and can be reduced by several days in sexually active males. Animals that rest sexually for 7 to 10 days have a maximum number of spermatozoa in the cauda epididymis, and this reserve is reduced by at least 25% with daily or every-other-day ejaculation.

The deferent ducts, or *vasa deferentia,* pass through the inguinal rings into the abdomen and connect the cauda epididymis with the pelvic urethra. In most species, the terminal portion of the deferent ducts enlarges to form prominent ampullae such as those found in ruminants and the stallion. In other species, the ampullae either are absent or are anatomically indistinct from the vasa deferentia. The ampullae serve as an additional storage depot for spermatozoa, and in some species, such as the bull, stallion, and dog, ampullary glands add to the ejaculate. Along with spermatozoa, ejaculated semen is composed primarily of accessory gland secretions that add volume, nutrients, buffers, and a number of other substances whose exact functions are unknown. The contribution to the ejaculate by each of the accessory glands varies with the species and is responsible for the variation in concentration, volume, and character between ejaculates. The vesicular glands lie lateral to the ampullae near the neck of the bladder. In the bull, ram, and buck, these organs are firm and lobulated with a narrow lumen, whereas in the stallion and boar, they are more saclike. The dog and tom lack vesicular glands but have relatively prominent prostate glands, especially the dog. The prostate gland is present in all domestic males and is intimately associated with the pelvic urethra, but it varies in size and appearance among species. The bulbourethral glands of the tom are almost as large as the prostate, but these glands are absent in the dog. In the stallion and bull, the bulbourethral glands are small, round to ovoid structures that lie adjacent to the pelvic urethra near the ischial arch, whereas those of the boar are large and cylindrical. The male llama/alpaca lacks vesicular glands, and the bulbourethral and prostate glands are present.

The copulatory organ of the male is the penis. It is more or less cylindrical in all species and extends from the ischial arch to near the umbilicus on the ventral abdominal wall, except in the tom and the llama/alpaca, in which the penis points posteriorly in the relaxed state. The body of the penis is surrounded by a thick fibrous capsule (the tunica albuginea) that encloses numerous cavernous spaces (the corpus cavernosum penis) as well as the corpus spongiosum penis, which immediately surrounds the urethra. Erection is a psychosomatic event that involves mutually occurring actions of the vascular, neurological, and endocrine systems. Contraction of the ischiocavernosus muscle during erection results in occlusion of venous outflow. At the same time, the parasympathetically mediated relaxation of corpus cavernosum and corpus spongiosum results in these cavernous spaces becoming engorged with blood, and the penis becomes elongated and turgid.

Emission Is the Release of Spermatozoa and Accessory Gland Fluids Into the Pelvic Urethra, Whereas Ejaculation Is the Forceful Expulsion of Semen From the Urethra

Emission is the release of spermatozoa and accessory gland fluids into the pelvic urethra as a result of sympathetically mediated thoracolumbar reflex contraction of the smooth muscle in the ductus deferens and accessory glands. Ejaculation is the forceful expulsion of semen from the urethra and is prompted by a parasympathetically mediated sacral reflex that induces rhythmic contractions of the bulbospongiosus, ischiocavernosus, and urethralis muscles. After ejaculation, a sacral sympathetically mediated increase in the smooth muscle tone of the cavernous spaces increases the outflow of blood, and contraction of the retractor penis muscle withdraws the penis into the prepuce. The seminal characteristics of the different species are listed in Table 40.2.

Spermatogenesis

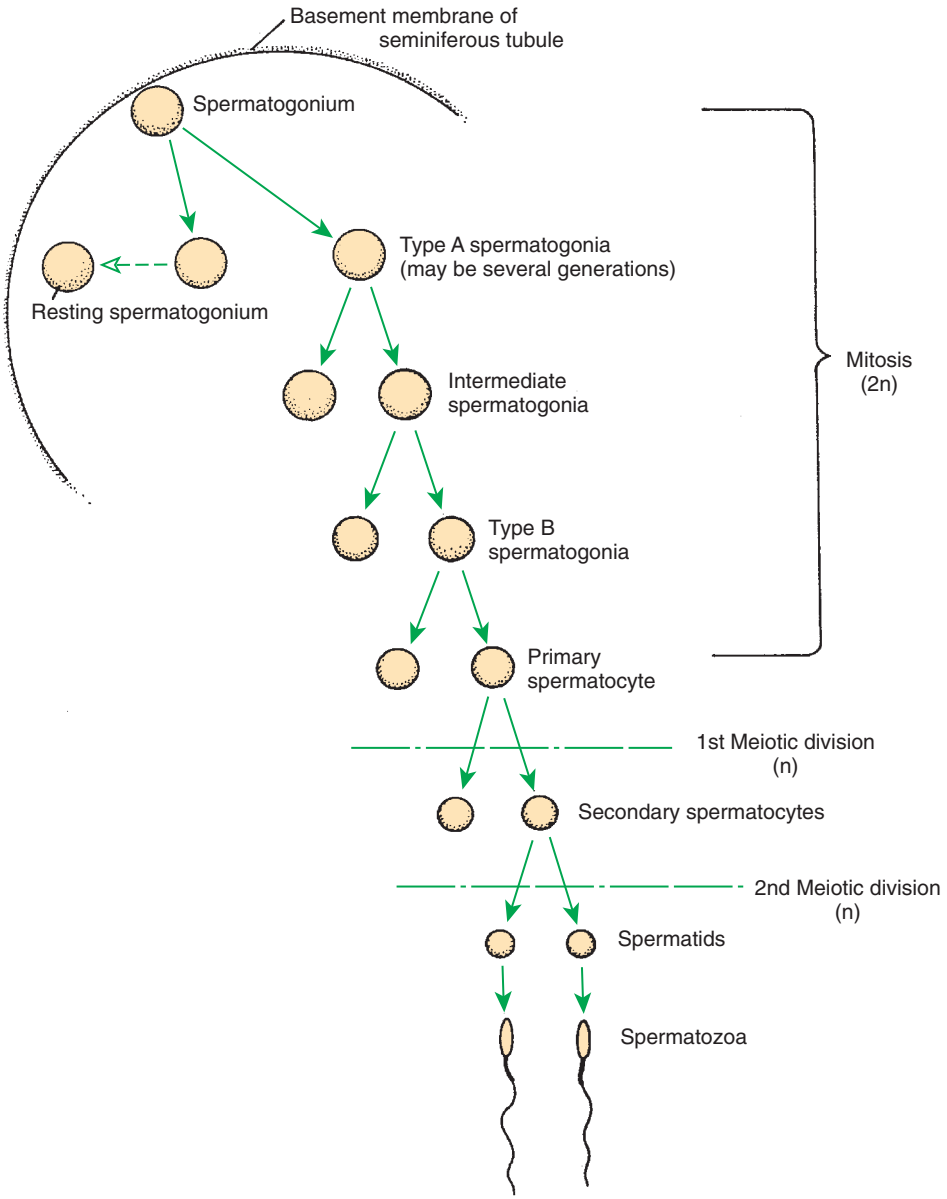
Spermatogenesis Is a Lengthy Orchestrated Process in Which Diploid Stem Cells Divide by Mitosis to Maintain Their Own Numbers and Cyclically Produce Progeny That Undergo Meiotic Division and Differentiation Into Haploid Germ Cells

Spermatogenesis is a lengthy orchestrated process in which diploid stem cells at the base of the seminiferous tubules (spermatogonia) divide through mitosis to maintain their own numbers. These cells also cyclically produce progeny that undergo further meiotic division and differentiation into haploid spermatids, which are released as spermatozoa (Fig. 40.1). Spermatogenesis is generally divided into three major events: spermatocytogenesis, meiosis, and spermiogenesis. Spermatocytogenesis accomplishes two important functions. First, the mitotic divisions of type A spermatogonia produce other spermatogonia that are not yet committed to the immediate spermatozoal production process, thus maintaining a population of stem cells. These stem cell divisions are responsible for the ability of the male to continuously produce spermatozoa throughout his adult life. Second, type A spermatogonia become type B spermatogonia, which further divide through mitosis to produce primary spermatocytes. The primary spermatocytes enter into the pool of meiotically dividing cells and ultimately produce spermatozoa.

TABLE 40.2 Seminal Characteristics From Domestic Animals

Parameter	Bull	Ram	Buck	Boar	Stallion	Alpaca/Llama	Dog[a]	Tom
Ejaculate volume (mL)	5–8	0.7–1.3	0.7–1.4	150–250	50–100	0.7–3.0	2.0–25	0.03–0.3
Spermatozoa concentration (millions/mL)	800–2000	2000–3500	2000–4500	200–300	150–300	80–250	60–500	1700–2900
Motile spermatozoa (%)	40–75	60–80	60–85	50–80	40–75	40–70	50–90	40–90
Normal spermatozoa (%)	65–95	80–95	75–95	70–90	60–90	55–85	50–90	50–90

[a]The ejaculate of the dog consists of three fractions.

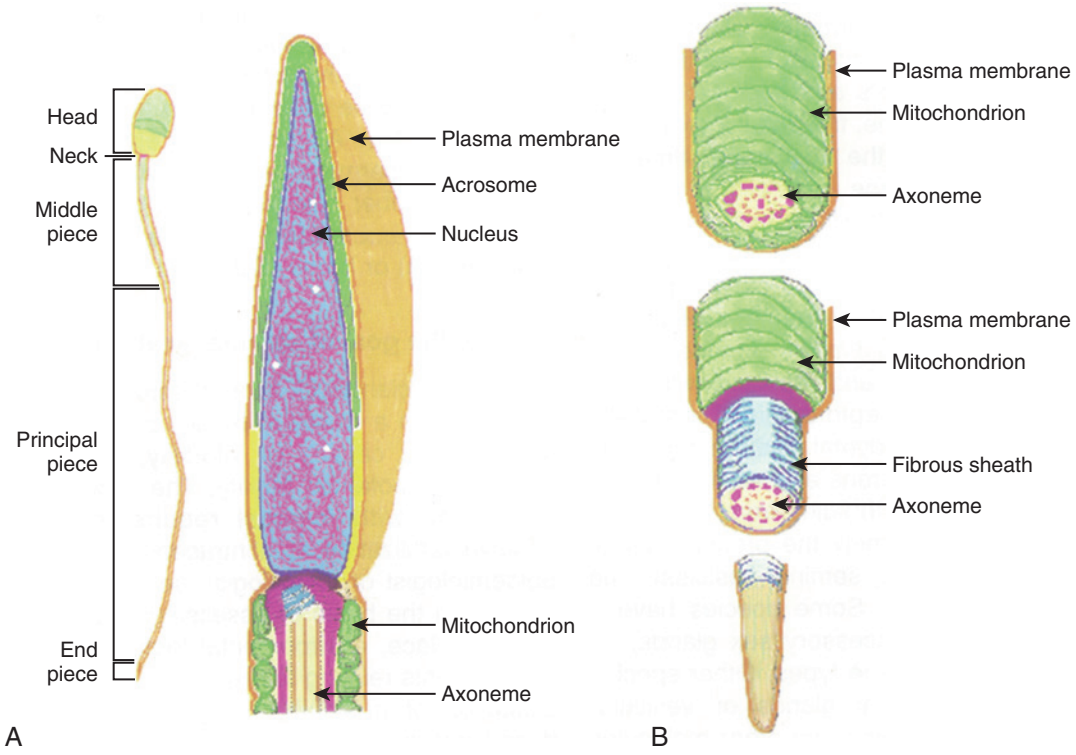


• **Fig. 40.1** Diagram of spermatogenesis. (From McDonald LE, Pineda MH, ed. *Veterinary Endocrinology and Reproduction.* Philadelphia: Lea & Febiger; 1989.)

Meiosis occurs only during the processes of oogenesis and spermatogenesis, in which the haploid condition results after two cell divisions with only one chromosomal duplication. During meiosis, homologous chromosomes pair up, and this facilitates the exchange of genetic material between chromosomes. At the first meiotic division, the homologous chromosomes segregate into the two resulting cells, creating a haploid condition. In the male, the resultant haploid cells are the secondary spermatocytes with duplicated chromatids. In less than 1 day after their formation, secondary spermatocytes divide to form spermatids that contain one chromatid from each of the haploid chromosomes.

The newly formed spermatids continue to differentiate without dividing to form mature spermatids through the process of spermiogenesis. Spermiogenesis occurs just before spermatids are released as spermatozoa at the luminal surface of the seminiferous tubule (spermiation). The major features of spermiogenesis include formation of the acrosome from the Golgi apparatus, condensation and elongation of the nucleus, formation of the flagellum, and extensive shedding of cytoplasm. The spermiated *spermatozoon* consists of a head, middle piece, and tail (Fig. 40.2). The head contains the genetic material to be combined with that of the oocyte during fertilization. Overlying the head is the acrosome, which contains hydrolytic enzymes necessary for penetration of the oocyte. The middle piece contains mitochondria, which provide the energy for microtubules extending into the tail to slide back and forth, past each other, thus producing tail movement.

Taking epididymal transit time into account, the interval from type A spermatogonia to ejaculated spermatozoa is approximately 60 to 70 days for the ram and the bull and 50 to 60 days for the boar, the dog, and the stallion. Therefore the interval from an event that may adversely affect the testis or epididymis to a decline in seminal quality may be as short as a few days to as long as 2 months. Similarly, at least 60 days would likely be required for an ejaculate to return to normal after a toxic insult to the testis.



• **Fig. 40.2** (A) Major elements of the mammalian spermatozoa. (B) Middle piece (*top*), principal (*middle*), and end piece (*bottom*) of a spermatozoon viewed in cross-section. (From Robaire B, Pryor JL, Trasler JM. *Handbook of Andrology.* Lawrence, KS: Allen Press; 1995.)

In theory, 16 primary spermatocytes and 64 spermatozoa develop from one type A spermatogonium in the bull and the ram. However, a percentage of potential sperm production is lost to degeneration during the normal course of spermatogenesis. In humans, approximately 40% of the sperm production potential is lost during the latter stages of meiosis. Daily sperm production is the number of spermatozoa produced per day by the testes. It is highly correlated with testicular size and is not affected by frequency of use for breeding.

Testicular Size Can Predict Daily Sperm Production

Testicular size is an important trait of medium to high heritability that provides an accurate estimate of the amount of sperm-producing parenchyma in the testis. Because of the influence of testicular size, there is a wide range in daily sperm production among domestic species. For example, daily sperm production has been calculated to be 0.37×10^9 in the dog and 16.2×10^9 in the boar. Within a species, both individual and breed variation in testicular size can also influence daily sperm production. Testicular size cannot be measured directly; therefore an indirect measure commonly used in ruminants is *scrotal circumference* (Fig. 40.3). In other species with more horizontally oriented testes, total scrotal width (Fig. 40.4) or testicular volume as determined by ultrasonographic measurements is used. Testicular size is influenced by species, breed, age, and body condition score. Each gram of normal testicular parenchyma produces the same quantity of spermatozoa according to the species but differs among species (Table 40.3). Therefore males with larger testes will produce more spermatozoa than males with smaller testes for the same age and species. In ruminants, scrotal circumference is also an accurate predictor of the age of puberty onset and of the percentage of normal seminiferous tubules. In cattle, there is a negative correlation between scrotal circumferences and age of puberty in the female offspring; this means bulls with larger scrotal circumferences will produce females that attain puberty at an earlier age.



• **Fig. 40.3** Measurement of scrotal circumference in a bull by using a scrotal tape.



• **Fig. 40.4** Measurement of total scrotal width in the stallion using calipers. (From Brinsko SP, Blanchard TL, Varner DD, et al. *Manual of Equine Reproduction*. 3rd ed. St. Louis: Mosby; 2010.)

TABLE 40.3 Body Weight, Testicular Weight, Spermatogenic Efficiency, and Daily Spermatozoa Production

	Body Weight (kg)	Pair Testes Weight (Grams)	Spermatogenic Efficiency[a]	Daily Spermatozoa Production ($\times 10^9$)[b]
Alpaca	65	20	NA	NA
Boar	150	750	23	17.3
Bulls	600	600	11	6.6
Dog	15	30	17	0.5
Lama	115	30	NA	NA
Ram	100	550	21	11.6
Stallion	500	350	16	5.6
Tom	5	20	16	0.3

[a]Spermatozoa produced per gram of testicular parenchyma ($\times 10^6$).
[b]Spermatozoa produced daily by the two testes.
NA, Not available.

Hypothalamic-Pituitary-Testicular Axis

The Reproductive System of the Male Is Regulated by the Hypothalamus, Which Is Hormonally Linked to the Anterior Pituitary and Testes by Luteinizing Hormone and Follicle-Stimulating Hormone

The reproductive system of male mammals is regulated by intricate feedback mechanisms involving the hypothalamus, anterior pituitary, and testes (Fig. 40.5). The hypothalamus synthesizes and secretes the *decapeptide gonadotropin-releasing hormone* (GnRH). Secreted in a pulsatile manner, GnRH acts directly on gonadotropic cells in the anterior pituitary. On stimulation by GnRH, these gonadotropes synthesize and secrete the gonadotropins follicle-stimulating hormone (FSH) and luteinizing hormone (LH). Both FSH and LH are heterodimeric glycoproteins made up of two noncovalently linked polypeptides. The alpha (α) subunit protein is common to both FSH and LH, whereas the beta (β) subunit is specific for each. Individual gonadotropes have the ability to synthesize and secrete FSH, LH, or both. The release of FSH and LH depends on the pulsatile pattern of GnRH secretion. Irregular, low-amplitude GnRH pulses result in FSH release, whereas high-frequency GnRH pulses induce the release of LH.

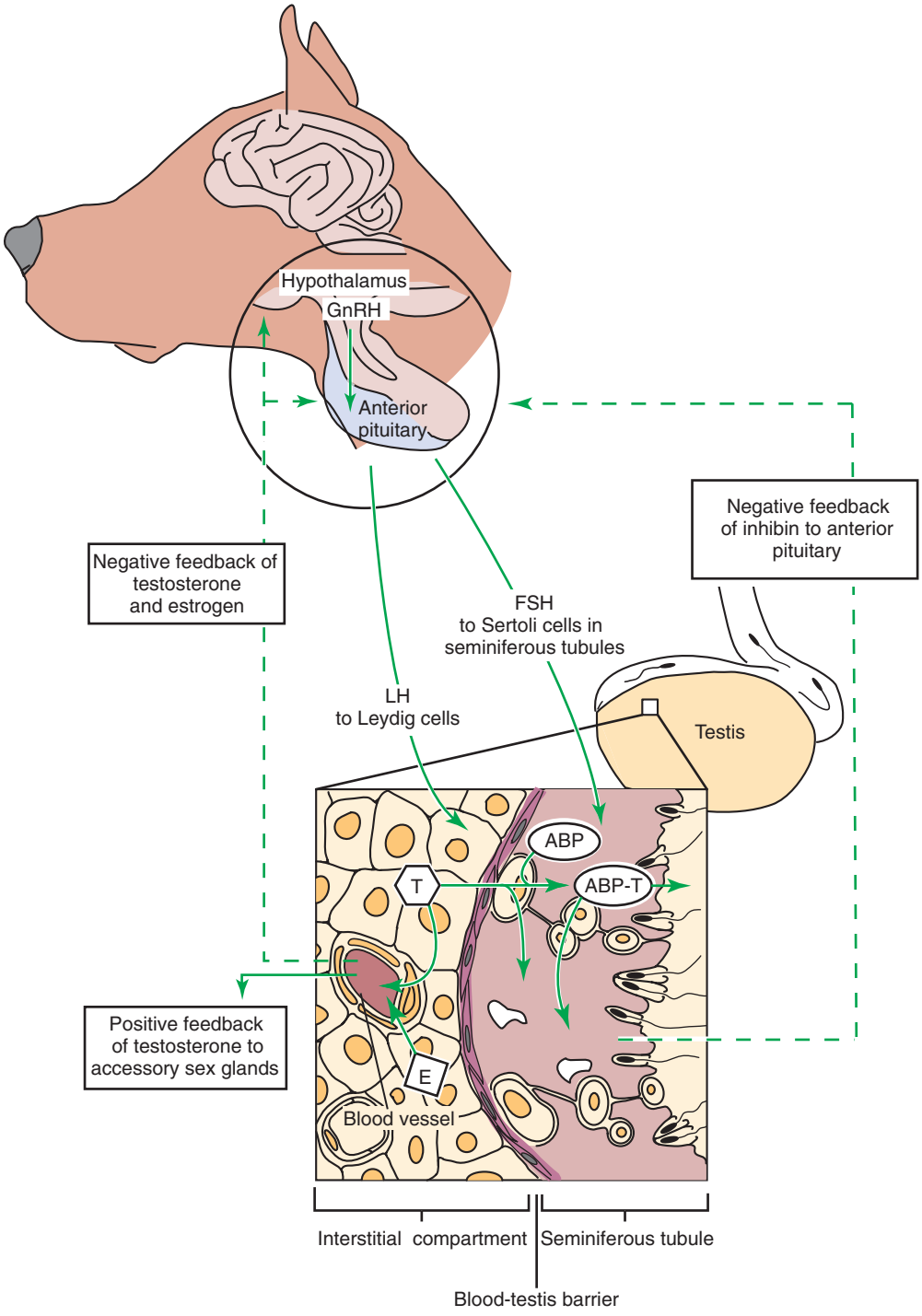
Within the testis, LH binds to membrane receptors on the Leydig cells and stimulates them to convert cholesterol to *testosterone.* Synthesized androgens diffuse into blood and lymph, where they are bound to androgen-binding protein (ABP) produced by the Sertoli cells. High local concentrations of androgens within the testis are considered essential for normal spermatogenesis to occur. ABP enhances the accumulation of testosterone and dihydrotestosterone in high concentrations within the seminiferous tubules and the interstitium of the testis. Within the testis, the target cells for testosterone are the peritubular myoid cells and the Sertoli cells, which envelop and support the developing sperm cells. ABP also facilitates the transport of androgens from the testis to the epididymis, where these hormones influence epididymal transit and the further maturation of spermatozoa.

Studies have demonstrated that FSH specifically targets receptors on the Sertoli cells within the seminiferous tubules. FSH and testosterone stimulate a variety of Sertoli cell functions, including the synthesis and secretion of ABP, inhibin, activin, estrogen, and several products (e.g., transferrin) involved in the transfer of nutrients to germ cells; meiosis; spermatocyte maturation; spermiation; and Leydig cell function. Sertoli and Leydig cells appear to interact in a paracrine fashion. Steroid production of Leydig cells can be stimulated by a product released by Sertoli cells, the secretion of which is enhanced by FSH. A potential candidate for such a substance is inhibin, which is produced by Sertoli cells in response to FSH and stimulates steroidogenesis in Leydig cells. Inhibin, a glycoprotein hormone, along with testosterone, is involved in the complex feedback regulation of pituitary function. Gonadal steroids are known to suppress FSH release, but inhibin appears to be the most potent inhibitor of FSH secretion from the pituitary. Testosterone, dihydrotestosterone, and estrogen regulate LH synthesis and secretion through negative feedback exerted at the level of the hypothalamus or the anterior pituitary gland. Because FSH and LH are necessary for high testicular concentrations of substances responsible for normal spermatogenesis, exogenous administration of testosterone or inhibin to enhance fertility would be contraindicated since they would impede the secretion of those factors responsible for maintaining an optimal spermatogenic environment.

Puberty

Puberty Is Not Synonymous With Sexual Maturity

Puberty in the male is when he is first able to produce sufficient numbers of sperm to impregnate a female. For practical reasons, for bulls, boars, rams, and stallions, this could be defined as the age when the ejaculate contains 50×10^6 spermatozoa, of which 10% or more are motile. It must be remembered that puberty is not synonymous with sexual maturity, which can occur months to years later, depending on the species.



• **Fig. 40.5** The reproductive system of male mammals is regulated by intricate feedback mechanisms that involve the hypothalamus, anterior pituitary, and testes. *ABP,* Androgen-binding protein; *ABP-T,* androgen-binding protein–testosterone; *E,* estrogen; *FSH,* follicle-stimulating hormone; *GnRH,* gonadotropin-releasing hormone; *LH,* luteinizing hormone; *T,* testosterone.

Puberty Results From a Continuous Process of Endocrine Changes That Are Initiated Shortly After Birth

The pituitary gland, gonads, and steroid-dependent target tissues are capable of responding to stimulatory hormones before puberty; therefore the hypothalamus is considered to play a pivotal role in the initiation of puberty. Puberty appears to be the end result of a continuous process of endocrine changes initiated shortly after birth. Some investigators theorize that puberty occurs when the animal's hypothalamic-pituitary complex becomes desensitized to the feedback inhibition of gonadal steroids. This desensitization would apparently allow increased discharge of GnRH from the hypothalamus and a greater response of the pituitary to GnRH. Although numerous factors can influence the modulation by the central nervous system of the endocrine system, the major factors that affect age at puberty in domestic animals are breed, energy intake, and season of birth.

The hypothalamic-pituitary-gonadal system in humans differentiates and functions during fetal life and briefly during infancy; it then is suppressed during childhood and is reactivated during puberty after almost a decade of low activity. Inhibition of the

hypothalamic-pituitary-gonadal system in prepubertal children is mediated through the suppression of GnRH synthesis and pulsatile secretion. Progressive pulsatile stimulation of the pituitary by GnRH and the gonads by LH and FSH is required for the initiation and progression of puberty. Prepubertal children secrete small amounts of FSH and LH from the pituitary, indicating that the hypothalamic-pituitary-gonadal axis is functional, but at a low level. This low level of gonadotropin secretion rapidly decreases when sex steroids are administered. Therefore a highly sensitive negative-feedback mechanism appears to exist in young prepubertal children, and a similar mechanism is likely to exist in prepubertal domestic animals.

Anabolic Steroids

Anabolic Steroids Are Androgen Derivatives That Exert Negative Feedback on the Hypothalamic-Pituitary-Testicular Axis

The use of *anabolic* steroids has become widespread in human and animal athletes as many attempt to increase performance. Testimonials from veterinarians, physicians, human athletes, and trainers indicate that improvements in the athlete's mental attitude, stamina, and physical strength are attained after anabolic steroids are administered. The major concern is that many individuals receiving anabolic steroids are peripubescent or prepubescent. Anabolic steroids are androgen derivatives that have been altered to maximize their anabolic action and minimize their androgenic side effects. It is not yet possible, however, to produce anabolic steroids devoid of androgenic activity, and many of the undesirable side effects of these drugs are caused by their androgenic activity. The adverse reproductive side effects observed with anabolic steroid use are similar to those associated with testosterone administration. Sustained testosterone or anabolic steroid administration affects pituitary function and leads to long-lasting impairment of testicular endocrine function. Potential side effects of the use of anabolic steroids in young animals may lead to incomplete development of the hypothalamic-pituitary-gonadal axis. The long-term side effects of anabolic steroid use on reproductive parameters in sexually immature animals are not yet known.

A high percentage of colts and stallions in training or racing receive androgenic drugs, including anabolic steroids, and these horses have smaller testicles than similar horses not receiving such drugs. The administration of anabolic steroids to stallions has been shown to reduce seminal quality, daily sperm output, daily sperm production, and testicular size. These effects most likely result from a negative-feedback mechanism on gonadotropin release from the pituitary. Alterations in seminal parameters are also observed, including depression of sperm concentration, sperm motility, and total number of sperm per ejaculate. Histological examination of the testes demonstrates a reduction in the number of developing germ cells other than type A spermatogonia. In addition, the mean diameter of Leydig cells is decreased, and changes indicative of testicular degeneration (including marked cytoplasmic vacuolization, shrunken tubules, and Leydig cells) and phagocytosis of spermatids by multinucleated giant cells have been observed. These adverse effects on the testes tend to be more severe in younger stallions. Repeated implantation of anabolic steroids in prepubertal bulls also results in decreased testicular size. Effects on testicular growth depend on the type of anabolic steroid administered, the age of the patient, and the dosage and duration of therapy.

Human studies, which did not directly investigate the effects of anabolic steroids on spermatogenesis, also found reduced levels of circulating gonadotropins, testosterone, or both. Individuals using high doses of testosterone and anabolic steroids for only 3 months still had hypogonadotropic hypogonadism 3 weeks after the cessation of drug use. The presence of atrophic testicles and low LH, FSH, and testosterone levels after drug withdrawal indicates that long-term androgenic or anabolic steroid use affects pituitary function and leads to long-lasting impairment of testicular endocrine function.

Questions regarding the possibility of permanent sterility or testicular atrophy with long-term anabolic steroid use have not been answered for adults, and even less is known about the effects in prepubertal or peripubertal individuals. Indirect evidence suggests that prepubertal individuals may be at a higher risk for permanent physiological derangements from anabolic steroid use than adults. Therefore, the indiscriminant use of anabolic steroids in males intended for breeding should be strongly discouraged.

CLINICAL CORRELATIONS

Infertility in a Stallion

History

You are asked to perform a "breeding soundness" examination on a 3-year-old Quarter Horse stallion that was mated with 10 mares last year and impregnated only one mare. The stallion is in demand because of his bloodlines and because his muscular and mature appearance contributed to winning a number of shows as a yearling. All the mares mated to this young stallion have been shown to be free of reproductive abnormalities. You ask whether the horse has had any illness or febrile episodes or has received any medication recently. The answer to all these queries is "no."

Clinical Examination

The stallion demonstrates normal *libido* when exposed to a mare in estrus, and two ejaculates are obtained 1 hour apart by means of an artificial vagina. Examination of both semen samples reveals poor spermatozoal concentration, low sperm numbers, and a high percentage of morphologically abnormal spermatozoa and immature germ cells. The stallion has a normal penis and prepuce, but his testes are small and soft. You ask if the stallion has ever received anabolic steroids, and with some reluctance the owner admits that the trainer did give anabolic steroids to the horse in preparation for the yearling futurity and for shows thereafter.

Comment

The use of anabolic steroids in performance animals is not uncommon. Even so, many owners and trainers are reluctant to admit to their use. Many colts are administered these drugs to give them a competitive edge in the show ring or on the racetrack so that they will later be in demand as sires. Unfortunately, because they are testosterone derivatives, the negative-feedback effects of anabolic steroids adversely affect the fertility of these animals, sometimes permanently. It is not known how severe or long-standing the adverse effects will be if anabolic steroids are given in the peripubertal period. Because this animal's testes are so small and soft, he apparently received high doses of anabolic steroids for a prolonged period during the development of the hypothalamic-pituitary-testicular axis, and the effects are most likely irreversible. It must be noted that as a 3-year-old, the horse is not yet sexually mature, and in the future he may still be able to produce sufficient numbers of normal spermatozoa to impregnate a small number of mares per season but certainly not a "full book."

CLINICAL CORRELATIONS—cont'd

Treatment

Other than time, no known treatment will reverse the detrimental side effects caused by anabolic steroid use in adult males. Even less is known about the long-term effects in young animals. Depending on the percentage of normal, progressively motile sperm, this animal may be able to breed a limited number of mares, most likely through artificial insemination. The owners could also have the stallion reexamined in several months or longer to see if there is any improvement in sperm morphology.

Infertility in a Bull

History

You are asked to perform a "breeding soundness" examination at your veterinary clinic on a 5-year-old Brangus bull before the breeding season that will start in 1 month. This bull was used in a single sire mating system in the last breeding season with 20 females (75 days exposure) producing 90% of calf crop. The bull was vaccinated against *Clostridium* and respiratory virus with killed products and dewormed 3 months ago. Forty days ago, this bull had a short but intense febrile respiratory illness, which resulted in recumbency and lasted 3 days. The bull was treated with antibiotics and antiinflammatory drugs for 5 days, and the problem resolved promptly.

Clinical Examination

This bull is in good body condition score (6; scale between 1 and 9) without foot/leg problems detectable when walking, and the general physical examination was unremarkable. The accessory sexual organs examined by palpation per rectum were normal. The scrotal contents were within normal limits, and the scrotal circumference was 42 cm. Semen was collected via *electroejaculation,* and normal penile protrusion and erection was observed. No abnormalities in the penis were detected. Semen evaluation revealed 20% of the spermatozoa to be motile with 50% of the spermatozoa having abnormal morphology. Most of the morphologic abnormalities consisted of detached heads, bent tails, and distal and proximal droplets among other defects.

Comment

Testicular function requires a low testicular temperature compared to systemic temperature in order to allow normal spermatogenesis. This bull had two instances that affected the testicular thermoregulatory function: fever and recumbency. The abnormalities noted in the semen are in agreement with the process of testicular degeneration. The spermatogenesis requires approximately 60 days. The degree of testicular compromise depends on the degree and duration of the insult as well as the inherent natural susceptibility of the male. Remember every ejaculate that you collect today is a snapshot of a process that began at least 60 days ago. Therefore the recommendation would be to re-evaluate this animal at least 60 days from the last day of the illness. It is also important to remark that there is variability among males in the degree of response to this insult.

Treatment

No known treatment, other than time, will reverse the detrimental effects caused by the fever. Due to the fact that spermatogenesis will require approximately 60 days in bulls (54 days plus epididymis transit time), the recommendation would be to not use this bull for breeding until the next evaluation is performed. Moreover, because this bull is used in single sire breeding season, the owner needs to find a new satisfactory potential breeder for this upcoming breeding season. Based on your recommendation, you can educate your client that the best time to perform a breeding soundness examination is at least 2 months before the breeding season.

Practice Questions

1. For most domestic species, the duration of spermatogenesis is approximately:
- **a.** 120 days
- **b.** 10 days
- **c.** 60 days
- **d.** 6 months
- **e.** 21 days

2. Normal spermatogenesis in domestic mammals requires a testicular temperature that is:
- **a.** Higher than core body temperature
- **b.** Lower than core body temperature
- **c.** The same as core body temperature
- **d.** Above freezing but below boiling
- **e.** Conducive to testosterone metabolism

3. Normal spermatogenesis requires an intratesticular testosterone concentration that is:
- **a.** The same as circulating levels
- **b.** Lower than circulating levels
- **c.** Static and unchanging
- **d.** Much higher than circulating levels
- **e.** Able to change rapidly with the maturational stage of the spermatozoon

4. Puberty in the male:
- **a.** Occurs at about the same time for all species.
- **b.** Is influenced only by the age of the animal.
- **c.** Is synonymous with sexual maturity.
- **d.** Is defined as when he is first able to produce sufficient numbers of sperm to impregnate a female.
- **e.** Is independent of GnRH secretion.

5. Anabolic steroids are testosterone derivatives and therefore:
- **a.** Should be helpful in treating infertile males.
- **b.** Have no effect on male fertility.
- **c.** Enhance testicular function.
- **d.** Should not be used in males intended for breeding because of negative-feedback effects.

Bibliography

Amann RP, Schanbacher BD. Physiology of male reproduction. *J Anim Sci*. 1983;57(suppl 2):380–403.

Coulter GH, Kastelic JP. Testicular thermoregulation in bulls. *Proc 15th Conf AI Reproduction N.A.A.B*; 1994;28–34.

Robaire B, Chan P. *Handbook of Andrology*. 2nd ed. Lawrence, Kan: Allen Press; 2010.

Roberts SJ. Veterinary obstetrics and genital diseases. In: *Theriogenology*. 3rd ed. Woodstock, Vt: David & Charles; 1986.

Strauss JF, Barbieri RL. *Yen and Jaffe's Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management*. 7th ed. Philadelphia: Saunders; 2014.

Zeleznik A, Plant T, eds. *Knobil and Neill's Physiology of Reproduction*. Vol. 1 and 2. 4th ed. London: Academic Press; 2015.

41 Glomerular Filtration

JILL W. VERLANDER

KEY POINTS

1. Introduction to the physiology of the kidney.
2. The glomerulus filters the blood.
3. The structure of the glomerulus allows efficient, selective filtration.
4. Glomerular filtration rate is determined by the mean net filtration pressure, permeability of the filtration barrier, and area available for filtration.
5. The filtration barrier is selectively permeable.
6. Glomerular filtration rate is regulated by systemic and intrinsic factors.
7. Glomerular filtration rate is measured by determining the plasma clearance rate of a substance.

Introduction to the Physiology of the Kidney

The kidney has diverse roles in maintaining homeostasis. In mammals, the two kidneys normally receive approximately 25% of the cardiac output. The kidneys filter the blood; retrieve substances removed from the blood during the filtration step but are needed by the body, including water, glucose, electrolytes, and low–molecular-weight proteins; and excrete metabolic waste. The kidneys respond to water, electrolyte, and acid-base disturbances by specifically altering the rate of reabsorption or secretion of these substances. The kidneys also produce hormones that regulate systemic blood pressure and red blood cell production.

These myriad functions are accomplished by an extensive variety of cell types, each with specific responses to direct and indirect signals, arranged in a particular pattern to form the functional unit of the kidney, the *nephron.* The nephron consists of the *glomerulus,* where the blood is filtered, and its associated renal tubule segments, where filtered substances are reabsorbed into the blood and plasma components are secreted into the tubule fluid. In the renal cortex, the nephrons merge into the collecting duct system, which traverses the kidney and empties into the renal pelvis. Fig. 41.1 provides an overview of the anatomical arrangement of nephrons in the kidney and the major functions of the nephron and collecting duct segments.

Most of our knowledge of renal physiology comes from experimental evidence from the mouse, rat, and rabbit. Our understanding of renal physiology continually evolves as more information is gathered.

The Glomerulus Filters the Blood

The first step in renal function is filtration of the blood by the glomerulus. The glomerulus is a compact network of capillaries that retains cellular components and medium– to high–molecular-weight proteins within the vessels while extruding a fluid nearly identical to plasma in its electrolyte and water composition. This fluid is the *glomerular filtrate;* the process of its formation is glomerular filtration.

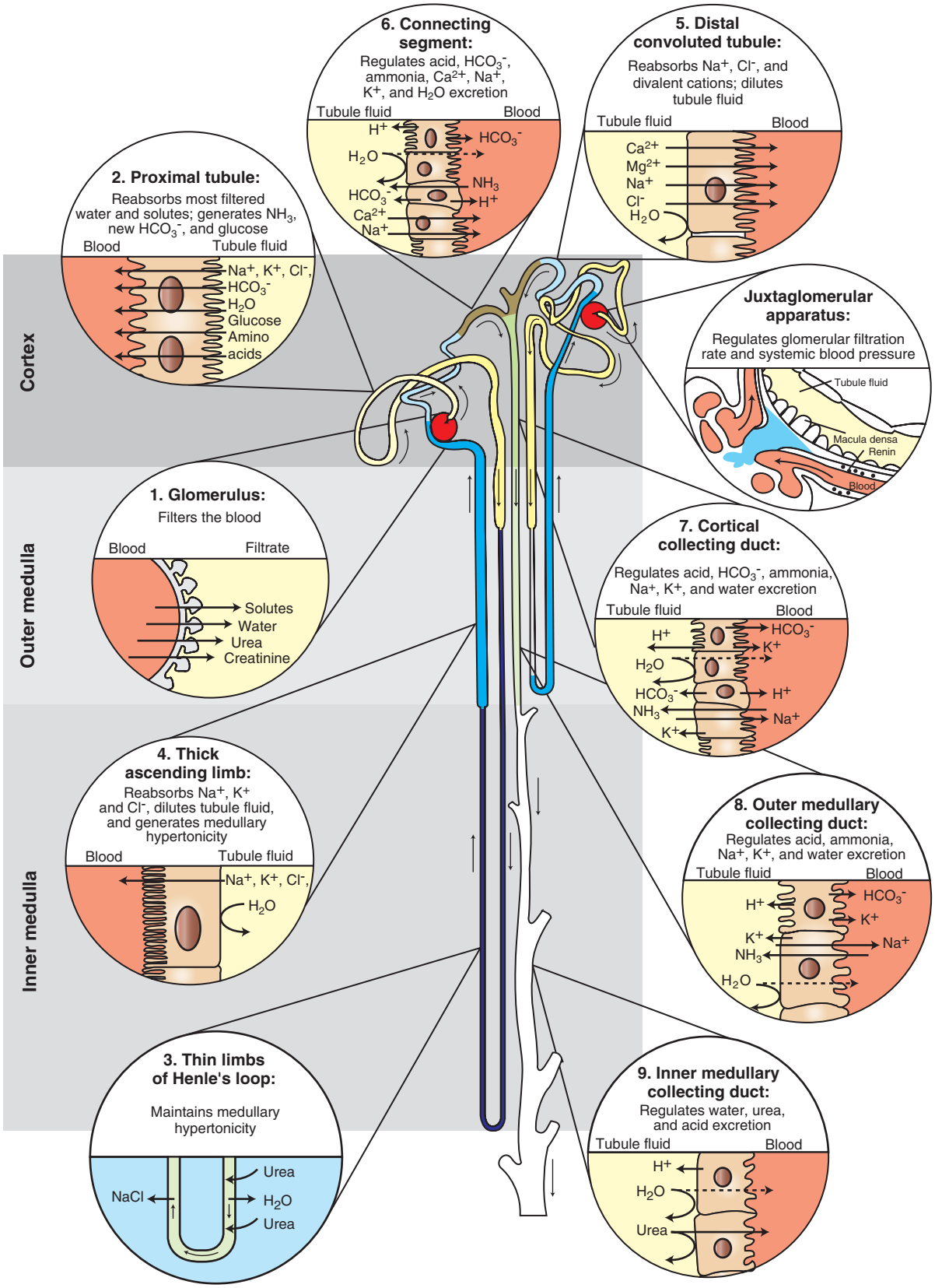
The *glomerular filtration rate* (GFR) is a clinically useful measure of renal function, expressed as milliliters of glomerular filtrate formed per minute per kilogram of body weight (mL/min/kg). To grasp the magnitude of GFR in more tangible terms, consider that an average-size beagle of 10-kg body weight with a typical GFR of 3.7 mL/min/kg would produce approximately 37 mL of glomerular filtrate per minute, or 53.3 L (~14 gallons) of glomerular filtrate per day, almost 27 times the beagle's extracellular fluid volume!

The Structure of the Glomerulus Allows Efficient, Selective Filtration

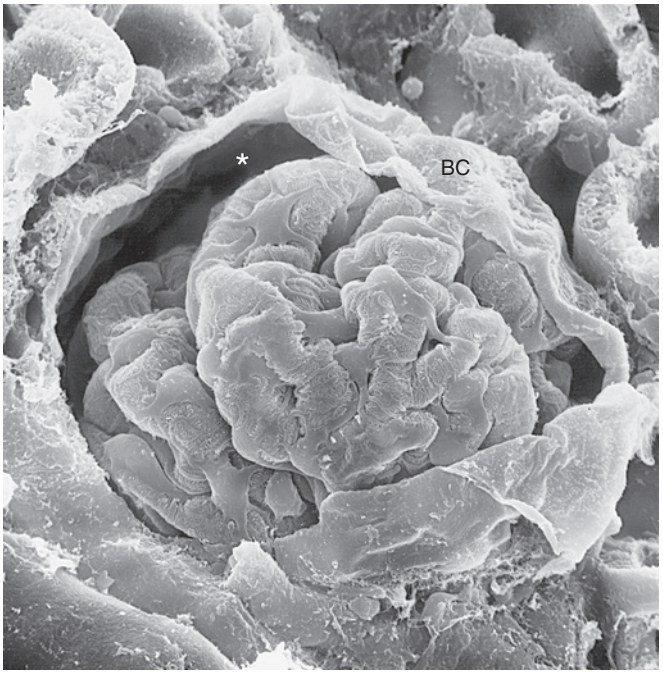
The *glomerular tuft* is composed of a network of capillaries (Fig. 41.2). In mammals, renal arterial blood flows ultimately to the afferent arteriole, which divides into numerous glomerular capillaries, which then anastomose downstream to form the efferent arteriole, which conducts the filtered blood away from the glomerulus (Fig. 41.3). Avian kidneys contain both mammalian-type and reptilian-type nephrons; in glomeruli of reptilian-type nephrons, the capillaries have few branches.

The glomerular tuft is encased by *Bowman's capsule,* which is lined with a single layer of cells, the parietal epithelium, transitioning to proximal tubule epithelial cells at or near the urinary pole. Between the glomerular capillary tuft and Bowman's capsule is *Bowman's space*, where the glomerular filtrate first appears. From here, the glomerular filtrate enters the lumen of the first segment of the proximal tubule.

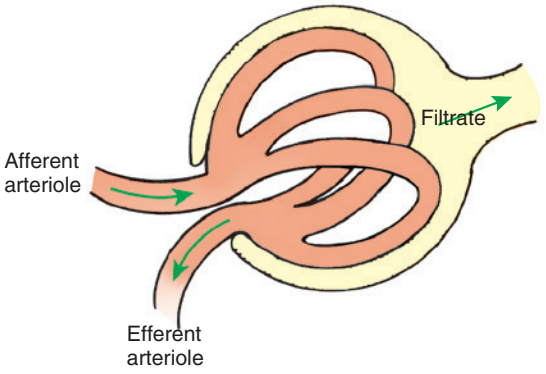
The structure of the glomerular capillaries is important in determining the rate and selectivity of glomerular filtration. The glomerular capillary wall consists of three layers: the capillary endothelium, the basement membrane, and the visceral epithelium (Fig. 41.4). The *capillary endothelium* is a single layer of very thin cells facing the blood in the capillary lumen. Endothelial fenestrae (Latin for "windows") are transcellular pores that conduct water and noncellular blood components to the *glomerular basement*



• **Fig. 41.1** Schematic illustration of juxtamedullary and superficial nephrons, listing functions of the segments of nephron and collecting duct. The glomerulus of a juxtamedullary nephron is located deep in the cortex near the corticomedullary junction. The thin limb of Henle's loop extends deep into the inner medulla. The glomerulus of a superficial nephron is located in the outer cortex, and Henle's loop extends only into the outer medulla. *Arrows* indicate the direction of tubule fluid flow. Segments are numbered in sequential order of modification of the tubule fluid, beginning with the glomerulus. (Modified from Madsen KM, Verlander JW. Anatomy of the kidney. In Tisher CC, Wilcox CS, ed. *Nephrology for the House Officer*. Baltimore: Williams & Wilkins; 2006.)



• **Fig. 41.2** Scanning electron micrograph of rat glomerulus. The glomerular tuft is a complex network of capillaries that is encased in visceral epithelial cells and Bowman's capsule (*BC*). Between the visceral epithelial cells and BC is Bowman's space (*asterisk*), where the glomerular filtrate appears.



• **Fig. 41.3** Schematic illustration of glomerulus. The afferent arteriole carries blood to the glomerulus and subdivides into numerous glomerular capillaries. Water and solutes cross the glomerular capillary wall into Bowman's space, forming the glomerular filtrate (*yellow area*), which flows into the proximal tubule. The glomerular capillaries coalesce, and the filtered blood leaves the glomerulus through the efferent arteriole.

membrane (GBM). The GBM is acellular and composed of various glycoproteins, including collagens (primarily type IV), laminins, nidogen-1, the heparin sulfate proteoglycans, agrin in mature animals and perlecan in developing glomeruli, and tubulointerstitial nephritis antigen-like protein. Compared with other basement membranes, the GBM is thicker and contains distinct glycoprotein isoforms. The GBM has three layers, created during development by the fusion of the basement membranes of the endothelial and epithelial cell layers and named according to their electron density and relative position. As shown in Fig. 41.4, the *lamina densa* (dense layer) is relatively dark because it is relatively resistant to the passage of electrons when viewed with a transmission electron



• **Fig. 41.4** Transmission electron micrograph of rat glomerular capillary wall. The three main layers of the capillary wall are shown in cross-section. A single layer of glomerular capillary endothelial cells lines the capillary lumen (*CL*). Numerous fenestrae (*arrowheads*) pierce the endothelial cells. On the outside of the capillary is a single layer of visceral epithelial cells. At the top of the micrograph is a portion of the cell body of a visceral epithelial cell. The secondary foot processes are aligned along the capillary wall, and the spaces between them are spanned by the slit diaphragm (*arrow*). Between the endothelial and epithelial cell layers is the glomerular basement membrane, consisting of the electron-lucent lamina rara interna adjacent to the endothelial cells, the lamina densa (*asterisk*), and the lamina rara externa adjacent to the visceral epithelial cells.

microscope. The lamina densa is composed of tightly packed glycoprotein fibrils. It is sandwiched between the *lamina rara interna* (inside thin, or loose knit, layer) on the endothelial side of the GBM and the *lamina rara externa* (outside thin, or loose knit, layer) on the epithelial side of the GBM. The laminae rarae are composed of a loose network of glycoprotein fibrils.

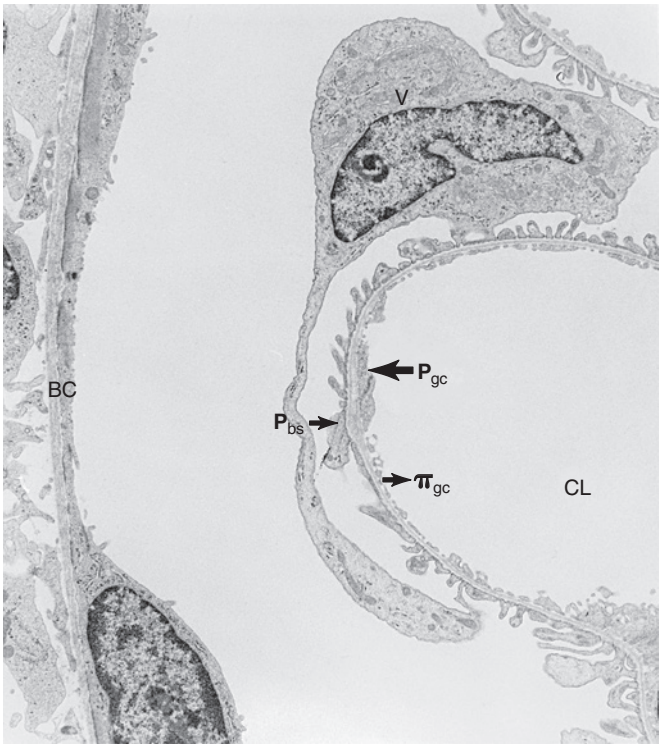
The third compartment of the glomerular capillary wall is the *visceral epithelium,* which is a layer of intricate, interlocking cells called *podocytes.* Numerous long, narrow extensions, the *primary* and *secondary foot processes,* interdigitate with foot processes from other podocytes and wrap around the individual capillaries (Fig. 41.5). The *epithelial slit diaphragm* spans between adjacent foot processes (see Fig. 41.4). The transmembrane proteins, nephrin and podocin, are critical to stabilizing other molecular components of the slit diaphragm, maintaining normal foot process structure and preventing filtration of plasma proteins.

Glomerular Filtration Rate Is Determined by the Mean Net Filtration Pressure, Permeability of the Filtration Barrier, and Area Available for Filtration

The glomerular capillary wall creates a barrier to the forces favoring and opposing filtration of the blood. The forces favoring filtration—that is, movement of water and solutes across the glomerular capillary wall—are the hydrostatic pressure of the blood within the capillary



• **Fig. 41.5** Scanning electron micrograph of the surface of rat glomerular capillaries viewed from Bowman's space. The cell bodies (*P*) of visceral epithelial cells, or podocytes, nestle between the capillary loops. The primary foot processes (*arrowheads*) radiate outward and wrap around the capillaries. Secondary foot processes extend from the primary foot processes and interdigitate with secondary foot processes from other podocytes.



• **Fig. 41.6** Transmission electron micrograph of rat glomerular capillary and Bowman's capsule (*BC*) illustrating the forces favoring and opposing filtration. The main force favoring filtration is the hydrostatic pressure of the glomerular capillary (P_{gc}). The forces opposing filtration are the hydrostatic pressure of Bowman's space (P_{bs}) and the oncotic pressure of the blood (π_{gc}). *CL,* Capillary lumen; *V,* visceral epithelial cell.

and the oncotic pressure of the fluid in Bowman's space (the ultrafiltrate). Normally, the oncotic pressure of the ultrafiltrate is inconsequential because high–molecular-weight proteins are not filtered. Therefore the main driving force for filtration is the glomerular capillary hydrostatic pressure. Forces opposing filtration are the plasma oncotic pressure within the glomerular capillary and the hydrostatic pressure in Bowman's space. Fig. 41.6 illustrates the direction and magnitude of these forces under normal conditions.

The net ultrafiltration filtration pressure (P_{uf}) at any point along the glomerular capillary is the difference between the capillary hydrostatic pressure (P_{gc}) and oncotic pressure of the ultrafiltrate in Bowman's space (π_{bs}) favoring filtration and the combined forces opposing filtration, the capillary oncotic pressure (π_{gc}) plus the hydrostatic pressure of the ultrafiltrate (P_{bs}). This relationship is expressed mathematically as follows:

$$P_{uf} = (P_{gc} + \pi_{bs}) - (P_{bs} + \pi_{gc})$$

Because the oncotic pressure of the ultrafiltrate in Bowman's space is essentially zero, the relevant forces contributing to the ultrafiltration pressure are represented by the simpler equation:

$$P_{uf} = P_{gc} - (P_{bs} + \pi_{gc})$$

As blood travels through the glomerular capillary, a large proportion of the fluid component of the plasma is forced across the capillary wall, whereas the plasma proteins are largely retained in the capillary lumen. Therefore the plasma oncotic pressure increases significantly along the capillary bed. At the same time, the loss of plasma volume along the capillary bed causes a decrease in the hydrostatic pressure in the capillary, although this change is small because of resistance in the efferent arteriole. The result is that the net filtration pressure tends to decrease along the capillary bed. However, during conditions that increase blood flow through the glomerular capillaries, the increase in capillary oncotic pressure is blunted and consequently, filtration in the distal portions of the glomerular capillaries is increased.

The GFR is the product of the mean net ultrafiltration pressure ($\overline{P_{uf}}$), the permeability of the filtration barrier, and the surface area available for filtration. The permeability of the filtration barrier is determined by the structural and chemical characteristics of the glomerular capillary wall. The product of the filtration barrier permeability and its surface area is the ultrafiltration coefficient (K_f). Thus the combined effects of the determinants of GFR are mathematically represented by the following equation:

$$GFR = \overline{P_{uf}} \times K_f$$

The Filtration Barrier Is Selectively Permeable

In addition to determining the hydraulic permeability, the structural and chemical characteristics of the glomerular capillary wall establish the selective permeability (*permselectivity*) of the filtration barrier. The permselectivity is responsible for differences in the rate of filtration of blood components. Normally, essentially all cellular components and plasma proteins the size of albumin or larger are retained in the bloodstream, whereas water and solutes are freely filtered. In general, substances with a molecular radius of 4 nm or more are not filtered, whereas molecules with a radius of 2 nm or less are filtered without restriction.

However, characteristics other than size affect the ability of blood components to cross the filtration barrier. The net electrical charge of a molecule has a dramatic effect on its rate of filtration. The *cationic* (positively charged) form of many substances is more freely filtered than the neutral form, which is more freely filtered than the *anionic* (negatively charged) form of the same molecule. For example, the cationic form of albumin is excreted at a rate approximately 300 times that of native albumin, which has a net negative charge. These differences are caused by a charge-selective barrier in the glomerular capillary wall that is created largely by the strongly negatively charged endothelial surface layer (ESL) coating the glomerular capillary endothelial cells, with a secondary contribution by negatively charged glycoproteins incorporated in the GBM. These fixed negative charges repel negatively charged plasma proteins and thus inhibit their passage across the filtration barrier. The shape and deformability of the molecule also affect its ability to cross the filtration barrier. Neutral dextran, a long, flexible molecule, crosses the filtration barrier approximately seven times as easily as horseradish peroxidase, a globular protein with a similar molecular radius and net charge.

Glomerular Filtration Rate Is Regulated by Systemic and Intrinsic Factors

The kidney normally maintains the GFR at a relatively constant level despite changes in systemic blood pressure and renal blood flow. The GFR stays within the physiological range due to renal modulation of systemic blood pressure and intravascular volume and intrinsic control of renal blood flow, glomerular capillary pressure, and K_f. Renal effects on systemic blood pressure and volume are mediated primarily through humoral factors, particularly the renin-angiotensin-aldosterone system. Intrinsic control of glomerular capillary perfusion is mediated by two autoregulatory systems that control the resistance to flow in the afferent and efferent arterioles: the myogenic reflex and tubuloglomerular feedback.

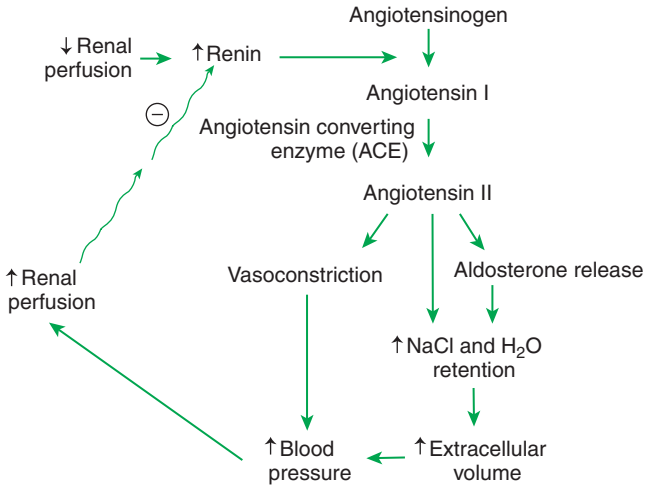
The *renin-angiotensin-aldosterone* system is an important regulator of GFR and renal blood flow. *Renin* is a hormone produced primarily by cells located in the wall of the afferent arteriole, the *granular extraglomerular mesangial cells,* which are specialized *juxtaglomerular cells.* A decrease in renal perfusion pressure, most often caused by systemic hypotension, is detected by baroreceptors in the afferent arteriole and stimulates renin release. Renin release is also modulated by variations in sodium chloride delivery to the distal nephron and in parallel with changes in sympathetic nerve activity. Renin catalyzes the transformation of *angiotensinogen,* which is produced primarily by the liver, to *angiotensin I.* Angiotensin I is converted to the more active *angiotensin II* by *angiotensin-converting enzyme* (ACE), which is located primarily in the vascular endothelium of the lung.

It is now well established that a local renin-angiotensin system (RAS) in the kidney also regulates intrarenal hemodynamics, glomerular filtration, and glomerular permeability in both physiologic and pathologic conditions. Components of the intrarenal RAS system include local production not only of renin but also angiotensinogen; prorenin, which can be activated locally to convert angiotensinogen to angiotensin I; and angiotensin converting enzyme, located primarily in the interstitial capillary endothelium and proximal tubule. Depending on the condition, the intrarenal RAS may operate in concert with the systemic renin-angiotensin system but also may function independently of the systemic system and may be activated inappropriately and contribute to pathologic conditions.

Angiotensin II is a potent vasoconstrictor and thus directly increases systemic blood pressure and renal perfusion pressure and decreases renal blood flow as well as the ultrafiltration coefficient, K_f. Angiotensin II activates sodium uptake from the tubule fluid into the blood in several renal tubules, including the proximal tubule, thick ascending limb, distal convoluted tubule, and collecting duct. Furthermore, circulating angiotensin II stimulates the release of *aldosterone* from the adrenal gland and *vasopressin* from the pituitary, other hormones that enhance renal sodium, chloride and water reabsorption into the blood. Thus angiotensin II increases salt and water retention, intravascular volume, and vascular resistance, all of which contribute to increased systemic blood pressure and renal perfusion pressure. Renin release is suppressed by the resulting increased extracellular volume and increased renal perfusion pressure as well as by a direct effect of the elevated plasma angiotensin II, which suppresses renin release by activation of AT1 receptors. This contributes to a negative-feedback system that helps to maintain renal perfusion and GFR within the physiological range (Fig. 41.7).

Within the kidney, there is direct control of glomerular capillary perfusion by two systems previously mentioned: the *myogenic reflex* and *tubuloglomerular feedback.* The myogenic reflex is an autoregulatory mechanism triggered by changes in glomerular perfusion pressure, whereas tubuloglomerular feedback is an autoregulatory mechanism triggered by changes in sodium chloride delivery to the distal nephron. These mechanisms protect the glomerulus from excessive perfusion pressure, stabilize GFR, and prevent filtration rates that exceed the transport capacity of the renal tubules.

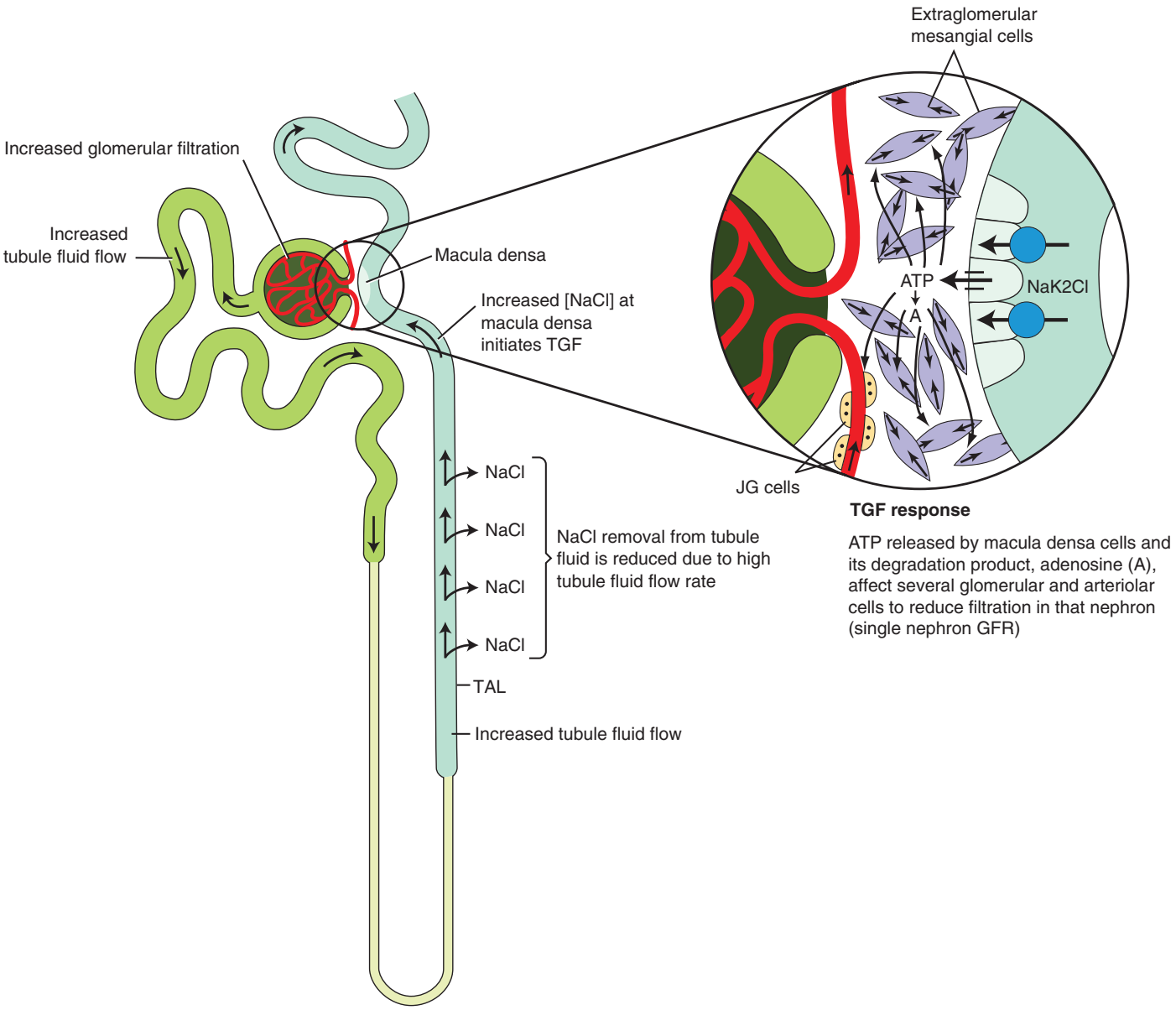
The *myogenic reflex* regulates renal blood flow and GFR by almost immediate constriction of the afferent arteriole and other preglomerular resistance vessels when increased perfusion pressure causes vessel distention and activation of stretch receptors in the arteriolar wall, thus increasing resistance to blood flow. Conversely, arteriolar dilation occurs almost immediately after a decrease in preglomerular arteriolar wall tension, thus reducing resistance to flow when vascular perfusion pressure decreases. These changes in vascular resistance contribute to maintenance of GFR and renal blood flow at a constant level despite marked changes in the renal arterial blood pressure. The myogenic response is independent of renal innervation but may be influenced by chemical mediators, such as nitric oxide and reactive oxygen species.



• **Fig. 41.7** Schematic illustration of renin-angiotensin-aldosterone system. *Wavy arrows* and *circled minus sign* represent inhibition of renin release.

The second intrinsic control mechanism, *tubuloglomerular feedback* (TGF), coordinates single nephron GFR with tubule transport capacity. The anatomical arrangement of individual nephrons, such that the distal nephron is closely associated with the glomerulus of the same nephron, is fundamental to this feedback system (see Figs. 41.1 and 41.8). An anatomically distinct cluster of epithelial cells, the *macula densa,* is located in the distal portion of the thick ascending limb of the loop of Henle and is situated between the afferent and efferent arterioles, adjacent to the extraglomerular mesangial region. These four structures (macula densa, afferent and efferent arterioles, and extraglomerular mesangium) together comprise the *juxtaglomerular apparatus.*

The elements of TGF are summarized here and illustrated in Fig. 41.8. An increase in glomerular filtration in a single nephron increases tubule fluid flow and sodium chloride concentration at the macula densa. Apical uptake of NaCl via the Na^+, K^+, $2Cl^-$ cotransporter (NKCC2) in macula densa cells leads to depolarization of the cells and basolateral release of adenosine triphosphate (ATP), which is degraded to adenosine. By signaling mediated through the extraglomerular mesangium, these substances cause afferent arteriolar vasoconstriction, thereby increasing resistance in the afferent arteriole, decreasing glomerular capillary perfusion pressure, and reducing glomerular filtration. These responses lead to reduced GFR in the individual nephron (single-nephron GFR),



• **Fig. 41.8** Schematic illustration of tubuloglomerular feedback (*TGF*) mechanism and juxtaglomerular (*JG*) apparatus. Increased glomerular filtration rate (*GFR*) increases tubule fluid flow rate. The increased flow in the thick ascending limb (*TAL*) decreases NaCl removal from the tubule fluid and thus increases NaCl delivery to the macula densa. Increased NaCl uptake at the macula densa suppresses renin release from juxtaglomerular (*JG*) cells in the afferent arteriole and stimulates basal release of ATP from the macula densa cells, which is degraded to adenosine. Binding of ATP and/or adenosine (A) to specific receptors causes afferent arteriolar constriction and decreased glomerular perfusion pressure. The result is decreased single-nephron GFR, thereby balancing GFR with the tubule reabsorptive capacity.

which prevents tubule fluid flow rates that exceed the tubule's reabsorptive capacity and thus prevents excessive fluid and solute loss. Conversely, increased NaCl delivery to the distal nephron also stimulates production of vasodilatory agents by macula densa cells; including carbon monoxide (CO), nitric oxide (NO) via nitric oxide synthase, and prostaglandin E_2 (PGE_2) via cyclooxygenase-2 (COX-2). NO and PGE_2 release modulates arteriolar constriction and blunts the TGF response, serving as a brake to prevent excessive reductions in single-nephron GFR. Macula densa activation also has systemic effects on intravascular volume by suppressing renin release and the salt and water retention caused by angiotensin II and aldosterone.

In addition to traditional macula densa signaling and TGF, epithelial-vascular crosstalk has recently been identified between the connecting segment, which is downstream of the macula densa, and preglomerular arterioles. This connecting tubule-glomerular feedback system is also triggered by increased luminal sodium chloride delivery and acts to reduce afferent arteriolar tone by release of the vasodilators, PGE_2 and epoxyeicosatrienoic acids. The specific role of this mechanism is presently unclear but likely serves as a modulator of the vasoconstriction induced by macula densa-activated TGF.

The endothelium also contributes to local control of renal vascular tone by producing potent vasoconstrictors and vasodilators. Endothelium-derived constricting factors include the vasoconstrictors endothelin, thromboxane A_2 (a metabolite of arachidonic acid), and angiotensin II. Angiotensin II and endothelin isoforms activate specific receptors in vascular smooth muscle cells of afferent and efferent arterioles in vitro and in general cause vasoconstriction and thereby regulate glomerular perfusion pressure.

Endothelium-derived relaxing factors include nitric oxide (NO), prostacyclin (prostaglandin I_2), and PGE_2. NO is produced in the kidney by oxidation of L-arginine, catalyzed by isoforms of nitric oxide synthase, and has important protective effects on the kidney. NO prevents renal damage by quenching reactive oxygen species, thus inhibiting intrarenal vasoconstriction, glomerular hypertension, mesangial cell proliferation, and mesangial matrix production. Cyclooxygenase in renal vasculature and tubules mediates PGE_2 production, which generally acts to dilate the glomerular and medullary microcirculation and increase glomerular perfusion. Clinical use of nonsteroidal antiinflammatory drugs (NSAIDs), which are cyclooxygenase inhibitors, commonly causes significant reductions in GFR and even renal failure, particularly in patients who have decreased extracellular fluid volume; these adverse effects are important concerns for long-term or high-dose administration in dogs and horses.

The intrarenal regulation of vascular tone and glomerular filtration is subject to complex interactions among the various regulatory mechanisms. For example, the vasoconstrictor angiotensin II can stimulate endothelin release; endothelin can stimulate release of the vasodilators NO and PGE_2; NO modulates renal COX-2 expression and PGE_2 production. Additional complications arise from counter-regulatory effects of the same agent on specific receptor subtypes; for example, angiotensin II both inhibits and enhances macula densa COX-2 expression and PGE_2 production, depending on which angiotensin receptor subtype is activated. These complicated interactions remain to be fully deciphered, but undoubtedly provide a refined system of checks and balances to preserve renal perfusion and glomerular function in healthy animals.

In addition to intrinsic renal controls, systemic factors can contribute to changes in GFR by regulating blood volume and vessel tone. Many hormones regulate blood volume. Angiotensin II, aldosterone, and vasopressin (antidiuretic hormone) enhance water and sodium chloride reabsorption by the renal tubules and thus increase blood volume. *Atrial natriuretic peptides,* produced in the cardiac atria, cause both natriuresis (sodium wasting) and *diuresis* (water wasting) and thereby reduce blood volume.

Systemic factors that affect vessel tone also affect systemic blood pressure, renal perfusion, and ultrafiltration. Vasopressin, angiotensin II and circulating catecholamines can cause systemic vasoconstriction and increase blood pressure. Beta-adrenergic stimulation can activate the renin-angiotensin system, and alpha-adrenergic stimulation can cause renal vasoconstriction, which can both reduce and redistribute renal blood flow. In addition to altering renal perfusion, vasoconstrictors can affect the other determinant of GFR, the ultrafiltration coefficient K_f. Vasoconstrictors may cause contraction of the mesangial cells within the glomerulus and thus reduce the area available for filtration. Because K_f is the product of the area available for filtration and the hydraulic permeability, mesangial cell contraction in vivo reduces K_f and thus reduces GFR.

Other factors that increase GFR include insulin-like growth factor and high dietary protein. Insulin-like growth factor increases GFR in normal and ischemic kidneys. A single high-protein meal causes transient elevations in renal blood flow and GFR; chronically high dietary protein intake causes sustained increases in renal blood flow and GFR. These observations are clinically relevant in the management of chronic renal insufficiency and renal failure. Although it may seem desirable to increase GFR by any means in patients with chronic renal disease, in fact the increase in GFR resulting from some high-protein diets can cause more rapid progression of glomerular injury and renal failure in animals and humans.

In birds, the GFR is more variable than in mammals, but the regulatory mechanisms are not well understood. Birds, unlike mammals, exhibit intermittent filtration in reptilian-type glomeruli. For example, during dehydration, birds, as well as reptiles, decrease GFR primarily by reducing the number of glomeruli that are actively filtering the blood, rather than by reducing the rate of filtration by the majority of nephrons. This may result from the release of arginine vasotocin, the avian analogue of mammalian arginine vasopressin, which decreases GFR in birds by causing constriction of the afferent arteriole of reptilian-type nephrons. Although some authors report that a juxtaglomerular apparatus is present in avian kidneys, the macula densa is either absent or rudimentary, and TGF has not yet been demonstrated.

Glomerular Filtration Rate Is Measured by Determining the Plasma Clearance Rate of a Substance

In experimental settings and in clinical practice, GFR is one of the most important parameters of renal function. Determination of GFR relies on the concept of clearance, that is, the rate the plasma is cleared of a substance. The rate of clearance is measured by the rate of elimination of a substance divided by its plasma concentration, mathematically expressed as follows:

$$C_X = (U_X V)/P_X$$

where C_X is the volume of plasma cleared of substance X per unit time, U_X is the urine concentration of substance X, V is the volume of urine collected divided by the time period of the collection, and P_X is the plasma concentration of substance X.

The net clearance rate of a substance is the sum of the rates of glomerular filtration and tubular secretion (peritubular capillaries

to tubule lumen) minus the rate of tubular reabsorption (tubule lumen to peritubular capillaries) of the substance. To determine the GFR accurately, the rates of tubular secretion and reabsorption must be determined or excluded from the equation. The latter requirement is satisfied by using inulin as the substance for the measurement of clearance. Inulin is freely filtered by the glomerulus but is neither reabsorbed nor secreted by the renal tubule cells. Because of these properties, and because inulin is not produced by the body, the rate of its disappearance from the blood after intravascular injection is strictly related to the rate of glomerular filtration. Therefore measured GFR can be expressed mathematically by the clearance equation, in which substance X is inulin:

$$\text{GFR} = C_{\text{inulin}} = (U_{\text{inulin}} V) / P_{\text{inulin}}$$

where GFR is in milliliters per minute, C_{inulin} is the rate of clearance of inulin from the plasma in milliliters per minute, U_{inulin} is the inulin concentration in a urine sample collected over a time T, V is the volume of the urine in milliliters collected over time T in minutes, and P_{inulin} is the mean plasma inulin concentration during time T.

Although inulin clearance is the gold standard for determining GFR, other methods of estimating GFR are more commonly used. In clinical settings, the most widely used measure of glomerular filtration is endogenous creatinine clearance. *Creatinine* is a byproduct of muscle metabolism that is freely filtered, is not reabsorbed by the tubule, and, at least in the dog, is not secreted by the tubule. In some species, however, at least 10% of urinary creatinine is secreted by the proximal tubule and thus endogenous creatinine clearance can overestimate GFR. Nevertheless, depending on the accuracy of the assay used for creatinine, the endogenous creatinine clearance test is a good measure of GFR. In practice, a 24-hour urine collection is done, and urine volume, mean urine creatinine, and plasma creatinine are measured. These values are used in the clearance equation as follows:

$$C_{\text{creatinine}} = U_{\text{creatinine}} V / P_{\text{creatinine}}$$

This results in an approximation of GFR in milliliters per minute. In veterinary medicine, the GFR is better expressed on the basis of body weight or body surface area—that is, as milliliters per minute per kilogram or milliliters per minute per square meter—because of the large variation in size within individual species.

In birds, creatinine clearance cannot be used for determination of GFR because avian renal tubules can secrete creatinine when the plasma level is elevated and can reabsorb creatinine when the plasma level is normal.

In both human and veterinary clinical practice, the serum creatinine level alone is frequently used to assess renal function. Increased serum creatinine suggests decreased renal clearance. It must be remembered that a very small increase in serum creatinine correlates with a large reduction in GFR and conversely, that a normal serum creatinine does not necessarily reflect normal renal function. For example, abnormally low muscle mass or low dietary protein intake tends to lower serum creatinine, which could mask or cause underestimation of a deficiency in renal clearance.

Another endogenous marker of renal clearance used in human medicine is serum cystatin-C. Cystatin-C is produced by all nucleated cells and is not significantly altered by relative muscle mass or dietary protein intake. Cystatin-C is freely filtered by the glomerulus and is not secreted by renal tubules. Unlike creatinine, approximately 99% of filtered cystatin-C is taken up by the proximal tubule and metabolized, thus it is not used to directly measure clearance. But like creatinine, abnormally high blood levels can indicate renal dysfunction. Recent studies in dogs suggest that serum cystatin-C may be a more sensitive indicator of kidney disease than serum creatinine, but studies of its usefulness in cats are conflicting.

In human medicine, various equations that use serum creatinine or cystatin-C levels or both, and the age, sex, and body weight or body surface area as surrogates for measuring muscle mass, are entered to estimate GFR. However, similar formulas have not yet been validated for veterinary patients.

CLINICAL CORRELATIONS

Chronic Renal Failure

History

You examine a 15-year-old male Siamese cat. The owner reports that the cat is listless, inappetent, and thin. The cat has been drinking more water than usual lately, urinating large volumes, and vomiting frequently.

Clinical Examination

The cat is very thin and moderately dehydrated. The mucous membranes are pale. Both kidneys are easily palpable and feel small, firm, and slightly irregular. The hematocrit is 22% (normal, 30% to 42%), serum creatinine level is 8.7 mg/dL (normal, 0.5 to 1.2 mg/dL), and urine specific gravity is 1.012. The urine sediment is unremarkable.

Comment

The cat has *chronic renal failure,* which is common in geriatric patients in small animal practice. The serum creatinine level is elevated because progressive loss of glomerular function has severely reduced the glomerular filtration rate (GFR), and creatinine is not cleared from the plasma normally. The urine is not concentrated in response to dehydration because tubular function is also compromised. The small kidney size is an indication of chronicity and is a result of gradual nephron loss and scarring. Anemia is common in chronic renal failure and results from many factors, including decreased production of erythropoietin by the kidney.

Treatment

In veterinary medicine, the treatment of chronic renal failure is usually supportive and symptomatic. This cat would probably benefit initially from rehydration with intravenous fluids and correction of electrolyte and acid-base disturbances as dictated by the serum chemistry profile. Chronic support may greatly improve the cat's quality of life and slow the progression of disease. This should include a diet containing low total protein that is high in bioavailability, low sodium, and low phosphorus. Supplementing with the water-soluble vitamins may be beneficial. Anabolic steroids have been used to improve the anemia, although exogenous erythropoiesis stimulating agents (darbepoetin, most commonly in cats) have become standard treatment in humans with anemia caused by chronic renal failure and are used in veterinary medicine as well.

Glomerulonephritis

History

A client presents his 3-year-old, spayed female springer spaniel. He reports that the dog has not been eating well for several days and seems to tire easily.

Clinical Examination

The dog seems bright and alert and is in good flesh. The only abnormality detected by physical examination is slight pitting edema in the distal

Continued

CLINICAL CORRELATIONS—cont'd

extremities. The left kidney is palpable, and feels smooth and normal in size. A urinalysis yields normal results except for 3+ protein (normal, negative to trace amounts), dipstick 3+ positive for blood, and the presence of a few red blood cell casts. A complete blood cell count is normal, and the only abnormality on a serum chemistry profile is a serum albumin level of 1.5 g/dL (normal, 2.3 to 4.3 g/dL).

Comment

This dog has *acute glomerulonephritis.* Proteinuria is indicative of glomerular disease because normally the filtration barrier established by the glomerular capillary wall prevents the passage of proteins into the tubular fluid. Damage to the glomerular capillary wall increases filtration of plasma proteins and, when the capacity of the proximal tubule to reabsorb the filtered protein is exceeded, protein appears in the urine. The loss of albumin in this case appears to be marked because the serum albumin level has dropped below normal levels.

Acute glomerulonephritis is likely because of the recent onset of clinical signs, the absence of renal failure, and the presence of red blood cell casts in the urine. Additional tests that are helpful in assessing the patient and guiding therapy include monitoring blood pressure, renal ultrasound, and a 24-hour urine collection to measure the severity of the protein loss and an endogenous creatinine clearance test done at the same time to determine if the GFR is reduced. A renal biopsy may be indicated to determine the type and severity of glomerular injury. Potential causes of acute glomerulonephritis, such as recent or concurrent bacterial or viral infections, heartworm disease, or autoimmune diseases, should be explored.

Several factors contribute to the development of peripheral edema: (1) Hypoalbuminemia decreases intracapillary oncotic pressure and allows leakage of fluid into the extravascular space; (2) Decreased intravascular volume activates the renin-angiotensin-aldosterone system and stimulates release of vasopressin, enhancing salt and water retention; and (3) Like the glomerular capillaries, the permeability of peripheral capillaries is increased, allowing leakage of plasma components into the interstitium.

Treatment

The treatment of glomerulonephritis varies. Occasionally, the initiating cause can be determined and removed. Some cases resolve spontaneously; at other times, various combinations of immunosuppressive and antiinflammatory agents are used to combat ongoing damage from immune complex deposition and glomerular inflammation. If pulmonary edema is present, the hypoalbuminemia must be treated with plasma or other colloids and volume overload (increased extracellular fluid) treated with loop diuretics to sustain the animal until the glomerular lesion is resolved or controlled. Blood pressure should be monitored and treated as indicated. Frequent assessment of urine protein/creatinine ratios as well as serum creatinine is warranted to monitor disease progression.

Practice Questions

1. The major force favoring filtration across the glomerular capillary wall is the:
- **a.** Oncotic pressure of the plasma
- **b.** Oncotic pressure of the glomerular filtrate
- **c.** Hydrostatic pressure of the blood
- **d.** Hydrostatic pressure of the glomerular filtrate
- **e.** Ultrafiltration coefficient

2. The glomerular filtration rate (GFR) is the:
- **a.** Volume of blood filtered by the kidneys per minute per kilogram of body weight
- **b.** Volume of plasma filtered by the kidneys per minute per kilogram of body weight
- **c.** Volume of urine produced by the kidneys per minute per kilogram of body weight
- **d.** Volume of glomerular filtrate formed by the kidneys per minute per kilogram of body weight
- **e.** Volume of blood cleared of creatinine by the kidneys per minute per kilogram of body weight

3. In clinical practice, the GFR may be estimated by determining the rate of creatinine clearance. The rate of creatinine clearance is the:
- **a.** Volume of plasma cleared of creatinine per minute per kilogram of body weight
- **b.** Volume of glomerular filtrate formed per minute per kilogram of body weight
- **c.** Weight of creatinine filtered from the blood per minute per kilogram of body weight
- **d.** Weight of creatinine per volume of urine formed per minute per kilogram of body weight
- **e.** Difference between the rate of plasma flow in the afferent and efferent arterioles

4. The two major characteristics that determine whether a blood component is filtered or retained in the capillary lumen are its:
- **a.** Molecular radius and molecular weight
- **b.** Molecular radius and lipid solubility
- **c.** Molecular radius and plasma concentration
- **d.** Molecular radius and electrical charge
- **e.** Molecular weight and length

5. The GFR is increased by:
- **a.** A low-protein meal
- **b.** Afferent arteriolar constriction
- **c.** Tubuloglomerular feedback
- **d.** Release of atrial natriuretic peptide
- **e.** Activation of the renin-angiotensin-aldosterone system

Bibliography

Carlström M, Wilcox CS, Arendshorst WJ. Renal autoregulation in health and disease. *Physiol Rev*. 2015;95(2):405–511.

Dantzler WH. Challenges and intriguing problems in comparative renal physiology. *J Exp Biol*. 2005;208(Pt 4):587–594.

Levey AS, Inker LA. Assessment of glomerular filtration rate in health and disease: a state of the art review. *Clin Pharmacol Ther*. 2017;doi:10.1002/cpt.729. [Epub ahead of print]. PMID: 28474735.

Munger KA, Maddox DA, Brenner BM, Kost CK Jr. The renal circulations and glomerular ultrafiltration. In: Skorecki K, Chertow GM, Marsden PA, et al, eds. *Brenner & Rector's the kidney*. 10th ed. Philadelphia: Elsevier; 2016.

Navar LG. Intrarenal renin-angiotensin system in regulation of glomerular function. *Curr Opin Nephrol Hypertens*. 2014;23(1):38–45.

Scott RP, Quaggin SE. The cell biology of renal filtration. *J Cell Biol*. 2015;209(2):199–210.

42

Solute Reabsorption

JILL W. VERLANDER

KEY POINTS

1. The renal tubule reabsorbs filtered substances.
2. Renal tubule function may be assessed by determining fractional excretion rate.
3. The proximal tubule reabsorbs the bulk of filtered solutes.
4. The proximal tubule secretes organic ions.
5. The thick ascending limb and distal convoluted tubule reabsorb salts and dilute the tubule fluid.
6. The collecting duct reabsorbs sodium chloride.
7. The distal convoluted tubule and collecting duct control renal potassium excretion.
8. Solute transport is regulated by systemic and intrarenal signals.
9. Angiotensin II stimulates sodium uptake in proximal tubule, distal nephron, and collecting duct.
10. Aldosterone enhances sodium reabsorption and potassium secretion.
11. Other hormones and ligands that regulate sodium transport include antidiuretic hormone, nitric oxide, endothelin-1, and atrial natriuretic peptide.
12. Phosphate uptake in the proximal tubule is decreased by parathyroid hormone.
13. Calcium reabsorption in the distal nephron and connecting segment is enhanced by parathyroid hormone, vitamin D_3, and calcitonin.

The Renal Tubule Reabsorbs Filtered Substances

It is vital that the renal tubules reabsorb the bulk of the glomerular ultrafiltrate and that they do so in a regulated manner to adapt to changes in intake and variations in losses from other organs. To understand the importance of tubule reabsorption of filtered substances, consider the 10-kg beagle that forms 53.3 L of glomerular filtrate each day. The ultrafiltrate contains virtually the same concentration of salts and glucose as plasma; without tubular reabsorption, the urinary loss of sodium, chloride, potassium, bicarbonate, and glucose alone would total more than 500 g of solute. Without tubular reabsorption, the beagle would need to replace these losses constantly throughout the day by eating more than a pound of salts and drinking more than 50 L of water at the same rate as the urinary loss to maintain fluid and electrolyte balance.

Fortunately, the renal tubule efficiently retrieves these and other constituents of the ultrafiltrate. Fig. 42.1 illustrates the percentages of various filtered substances that remain in the tubule fluid at different points along the tubule. Normally, 100% of the filtered glucose is reabsorbed by the proximal tubule; by the time the final urine is formed in the terminal collecting duct, approximately 99% of the filtered water and sodium has been retrieved.

Renal Tubule Function May Be Assessed by Determining Fractional Excretion Rate

The net rate of tubular reabsorption and secretion of a filtered substance is expressed as the *fractional excretion rate.* The fractional excretion rate is determined by comparing the urine-to-plasma ratio of a specific solute to the urine-to-plasma ratio of a reference substance to eliminate the confounding effect of water reabsorption on the urinary concentration of the solute of interest. Thus the fractional excretion rate of a substance X is the ratio of the urinary concentration of X (U_X) to the plasma concentration of X (P_X) divided by the urinary/plasma (U/P) ratio of a reference substance that is neither secreted nor reabsorbed. In experimental settings, the urinary and plasma concentrations of inulin during a constant inulin infusion may be used for reference. However, it is more practical in clinical situations, and many experimental settings, to use creatinine as the reference substance. Therefore the fractional excretion rate of X (FE_X) is determined by the following equation:

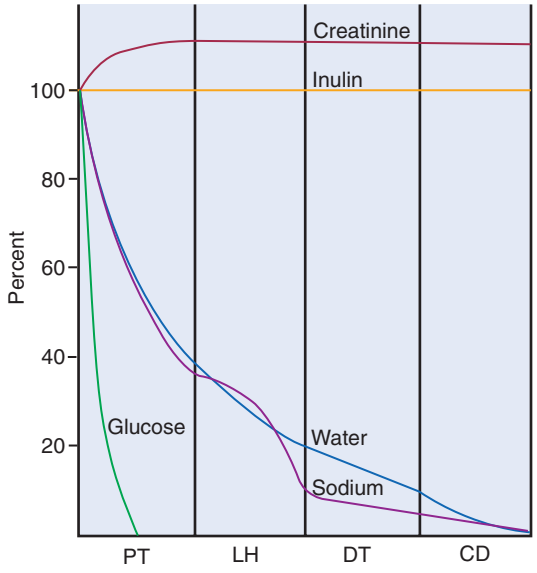
$$FE_X = U_X / P_X \div U_{creatinine} / P_{creatinine}$$

where $U_{creatinine}$ and $P_{creatinine}$ are the urinary and plasma concentrations of creatinine. By multiplying FE_X by 100, the fractional excretion rate is expressed as the percentage of filtered X that is excreted. The fractional excretion rate, such as of sodium (FE_{Na}) or of urea (FE_{urea}), can be used in clinical cases to assess the functional integrity of the renal tubules and thereby help distinguish between prerenal causes of reduced glomerular filtration rate (GFR) and intrinsic renal disease causing acute kidney injury and tubular necrosis.

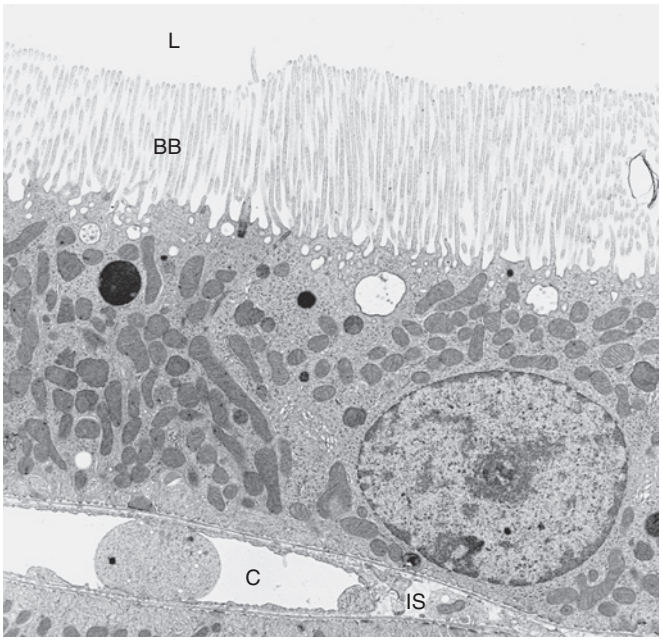
The Proximal Tubule Reabsorbs the Bulk of Filtered Solutes

The rate of reabsorption and secretion of filtered substances varies among segments of the renal tubule. In general, the proximal tubule reabsorbs more of the ultrafiltrate than the other tubule segments combined, at least 60% of most filtered substances.

The structure of the proximal tubule and its proximity to the peritubular capillary facilitate the movement of tubule fluid components into the blood through two pathways: the *transcellular*



• **Fig. 42.1** Illustration of the percentage of filtered substances [$(U_x/P_x) \times 100/(U_{inulin}/P_{inulin})$] remaining in the tubule fluid in various tubule segments. In some species, creatinine is secreted by the proximal tubule and is excreted at a greater rate than the reference substance, inulin. *CD,* Collecting duct; *DT,* distal tubule; *LH,* loop of Henle; *PT,* proximal tubule. (Modified from Sullivan LP, Grantham JJ, eds. *Physiology of the Kidney.* 2nd ed. Philadelphia: Lea & Febiger; 1982.)

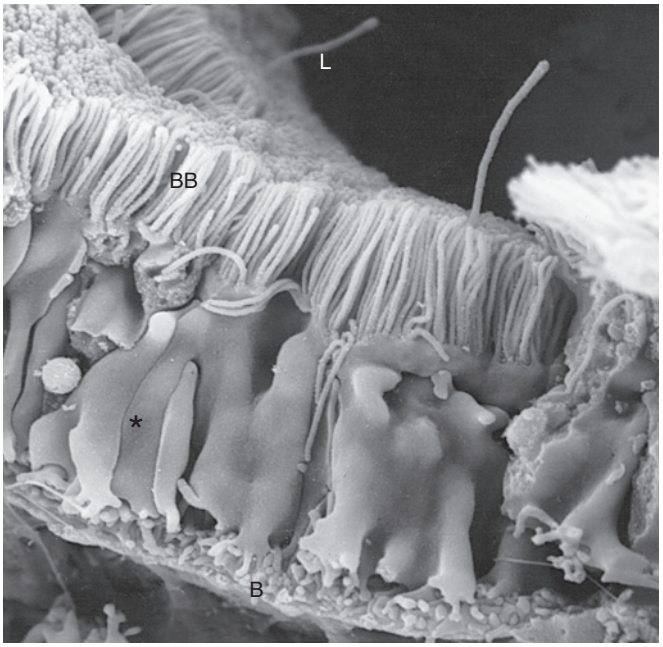


• **Fig. 42.2** Transmission electron micrograph of cross section of rat proximal tubule. The brush border (*BB*) of the apical plasma membrane extends from the epithelial cells into the tubule lumen (*L*), where it contacts tubule fluid. On the basal side of the cell are the interstitial space (*IS*) and the peritubular capillary (*C*).

pathway and the *paracellular pathway.* The tubule fluid flows over the apical (luminal) surface of the proximal tubule epithelial cell. Substances transported through the transcellular pathway cross the apical plasma membrane, cytoplasm, and the basolateral plasma membrane into the interstitial fluid that lies between the capillary and the epithelial cells. Movement across the apical and basolateral plasma membranes occurs largely by carrier-mediated transport. The vast plasma membrane surface area of the proximal tubule contributes to transcellular transport. The apical plasma membrane has extensive microprojections, called *microvilli,* which collectively create the *brush border* (Figs. 42.2 and 42.3) and increase the apical surface area by 36-fold in measurements of isolated rabbit proximal convoluted tubules. On the blood (basal) side of the cell, the plasma membrane has complex infoldings that can enhance the surface area such that the basal surface area equals that of the apical surface area in cortical segments of the proximal tubule. The benefits of the enhanced plasma membrane surface area include increased capacity for the multitude of solute transporters and increased exposure to the luminal and interstitial fluids.

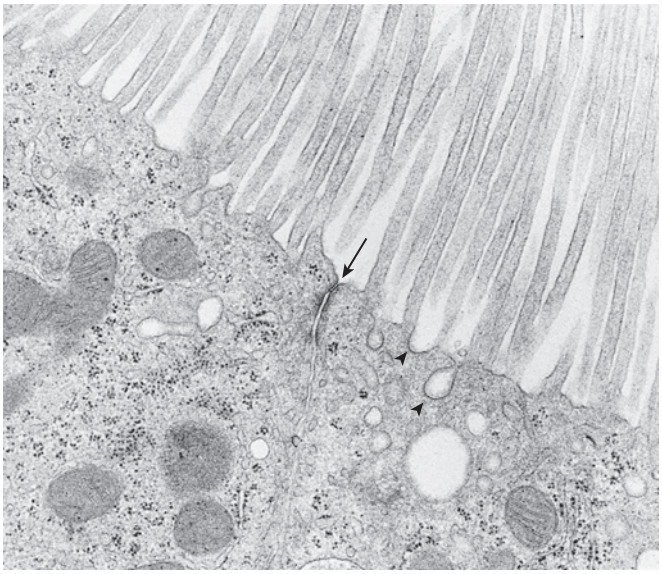
The second route of transport in the proximal tubule is the *paracellular pathway.* Substances pass through the paracellular pathway from the tubule fluid across the *zonula occludens,* a permeable structure that attaches the proximal tubule cells to each other at the junction of the apical and basolateral plasma membrane domains (Fig. 42.4). Along the proximal tubule, various *claudin* proteins form channels in the zonula occludens that confer permeability to specific solutes and water. Paracellular transport occurs by passive diffusion or by *solvent drag,* which is the entrainment of solute by the flow of water. Substances crossing the zonula occludens enter the lateral intercellular space, which is thought to communicate freely with the interstitial fluid, where reabsorbed substances can be taken up by the peritubular capillary.

Movement of water and solute from the interstitial fluid into the bloodstream is driven by Starling's forces (see Chapter 23) and aided by the proximity of the *peritubular capillary.* In

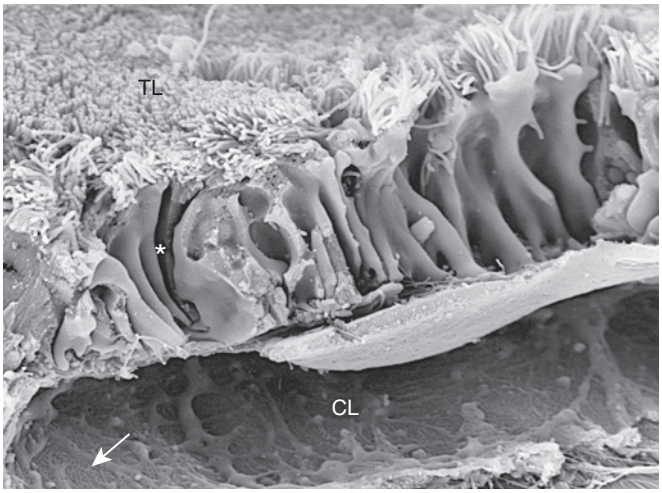


• **Fig. 42.3** Scanning electron micrograph of rat proximal tubule, viewed from the lateral intercellular space. The lush brush border (*BB*) carpets the luminal aspect (*L*). Lateral cellular processes (*asterisk*) interdigitate with those of neighboring cells. The surface of the basal plasma membrane (*B*) is amplified by extensive membrane infoldings, creating numerous processes called *micropedici* ("tiny feet").

mammals, the peritubular capillary originates at the glomerular efferent arteriole, subdivides, and wraps closely around the basal aspect of the proximal tubule (Fig. 42.5). The plasma leaving the glomerulus has a high oncotic pressure because water and salts are filtered, but proteins are retained in the capillary. The peritubular capillary has low resistance, and thus the hydrostatic pressure



• **Fig. 42.4** Transmission electron micrograph of apical region of rat proximal tubule viewed in cross-section. The zonula occludens (*arrow*) joins adjacent proximal tubule cells. The zonula occludens divides the apical plasma membrane from the basolateral plasma membrane and separates the tubule fluid from the fluid of the lateral intercellular space. Also seen are coated pits (*arrowheads*) that contain the binding sites for substances reabsorbed by receptor-mediated endocytosis.



• **Fig. 42.5** Scanning electron micrograph of rat proximal tubule and peritubular capillary. The peritubular capillary wraps around the basal aspect of the proximal tubule cells. Substances retrieved from the tubule lumen (*TL*) are delivered through either the transcellular pathway or the paracellular pathway into the fluid bathing the basolateral aspect of the epithelial cells. Water and solutes enter the interstitial space and diffuse across the peritubular capillary wall into the capillary lumen (*CL*). *Asterisk* represents lateral intercellular space; *arrow* represents fenestrae of peritubular capillary endothelium.

in the capillary is low. Both these conditions—high peritubular plasma oncotic pressure and low peritubular capillary hydrostatic pressure—favor fluid and solute uptake from the interstitium into the bloodstream.

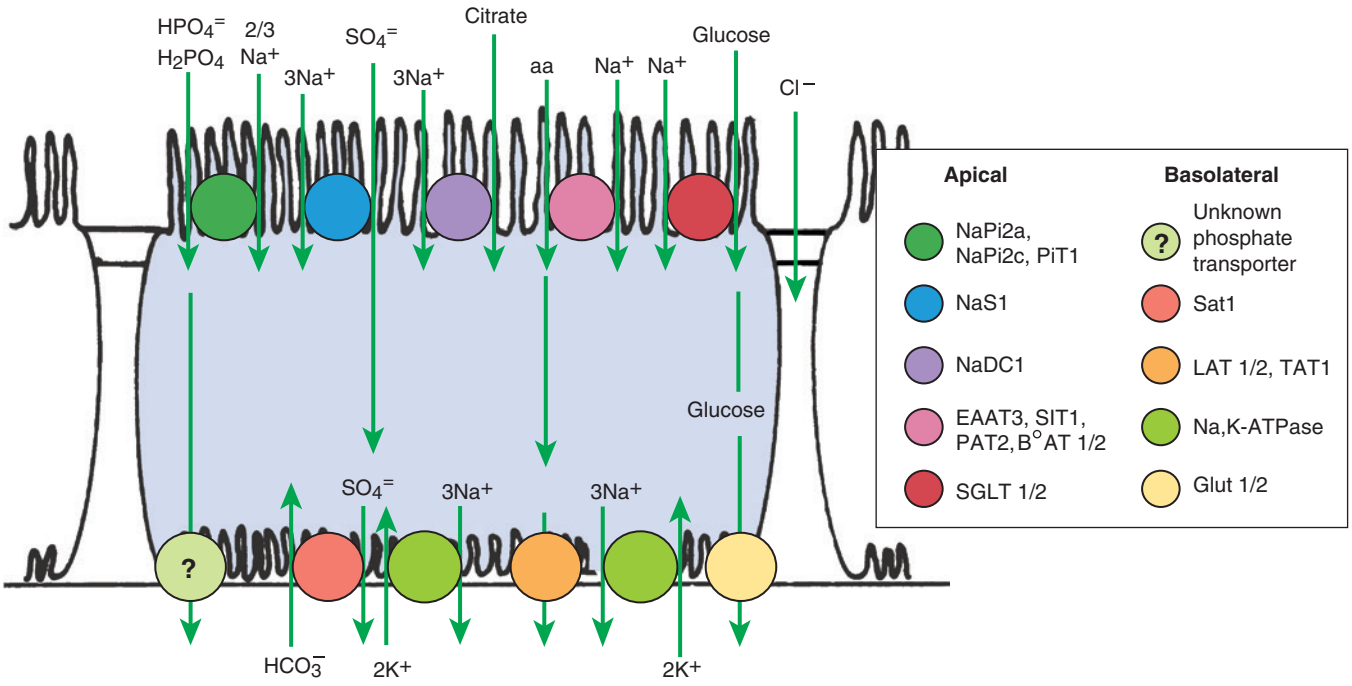
In birds, the effect of the peritubular blood supply on tubular reabsorption and secretion is complicated by the presence of a renal portal circulation. Renal portal veins anastomose with the efferent glomerular arterioles and supply peritubular blood to the reptilian-type nephrons and the proximal and distal tubules of mammalian-type nephrons; thus these tubules, but not the loops of Henle of the mammalian-type nephrons, receive a mixture of portal venous and arterial blood. The rate of flow to the renal portal supply varies and is controlled by a smooth muscle valve.

Reabsorption of solutes by renal tubule epithelia takes place by a number of mechanisms, including primary active transport, carrier-mediated secondary active transport, solvent drag, and passive diffusion. (Transport mechanisms are described in Chapter 1.) In the proximal tubule, most solute reabsorption is driven by the active transport of sodium ions (Na^+) by the sodium-potassium–adenosinetriphosphatase (Na^+,K^+-ATPase) pump, which is located in the basolateral plasma membrane. The Na^+,K^+-ATPase extrudes 3 Na^+ ions from the cell and takes up 2 K^+ ions on each turnover of the pump (Fig. 42.6).

Na^+,K^+-ATPase activity reduces the intracellular Na^+ concentration and increases intracellular K^+ concentration. Outward diffusion of K^+ down its chemical gradient through K^+ channels makes the cell interior electrically negative relative to the exterior. These two factors create an electrochemical gradient for Na^+ across the apical plasma membrane, favoring Na^+ uptake from the tubule fluid into the cell. Na^+ uptake across the apical plasma membrane is facilitated by specific transporters in the membrane that couple the movement of other solutes either in the same direction as Na^+ (*cotransport*) or in the opposite direction (*countertransport*). Specific Na^+-dependent transporters for glucose (SGLT1, SGLT2), amino acids (EAAT3, TauT, B^0AT1, and several more), phosphate (NaPi2a, NaPi2c, PiT-1), sulfate (NaS1), and citrate (NaDC1) mediate their uptake from the proximal tubule fluid by this mechanism of *secondary active transport.* The uptake of these substances increases their intracellular concentration, and they move across the basolateral plasma membrane and into the blood down their electrical or chemical gradient, primarily via solute specific transporters and partly by passive diffusion. The list of solute transporters in the apical and basolateral plasma membranes continues to grow as more are discovered. Several of the apical Na^+-coupled solute cotransporters and basolateral exit mechanisms are illustrated in Fig. 42.6.

Bicarbonate (HCO_3^-) reabsorption in the proximal tubule is also driven by the Na^+ gradient, although indirectly. The chemical gradient for Na^+ drives Na^+ and proton (hydrogen ion, H^+) countertransport across the apical plasma membrane through a Na^+/H^+ exchanger (NHE3). Secreted H^+ combines with filtered HCO_3^- in the tubule fluid to form water (H_2O) and carbon dioxide (CO_2), catalyzed by the enzyme carbonic anhydrase tethered to the apical plasma membrane of proximal tubule cells. CO_2 enters the cell across the apical plasma membrane, possibly facilitated by the integral membrane protein, aquaporin 1 (AQP1). Cytoplasmic carbonic anhydrase catalyzes the hydroxylation of CO_2 with OH^- donated from H_2O, forming H^+ and HCO_3^- in the cell. HCO_3^- crosses the basolateral plasma membrane primarily through a splice variant of the Na^+,3-(HCO_3^-) cotransporter (NBCe1-A). The majority of H^+ is transported into the tubule fluid through the Na^+/H^+ antiporter (NHE3); the electrogenic proton pump, H^+ATPase, also contributes to proton secretion. By this complex mechanism, illustrated in Fig. 42.7, the proximal tubule reabsorbs 60% to 85% of filtered HCO_3^-.

Chloride ion (Cl^-) reabsorption in the proximal tubule is also indirectly powered by the Na^+,K^+-ATPase pump and occurs through both paracellular and transcellular routes. As Na^+, HCO_3^-, glucose, amino acids, and other solutes are selectively reabsorbed and water is taken up along with these solutes, the concentration of Cl^- in the tubule fluid rises, establishing a chemical gradient for



• **Fig. 42.6** Schematic illustration of transport processes in the proximal tubule epithelial cell. The majority of transport is driven by active reabsorption of Na^+ by the Na^+,K^+-ATPase located in the basolateral plasma membrane. Glucose, phosphate, sulfate, citrate, and amino acids (*aa*), and other solutes enter the cell by Na^+-coupled secondary active transport on solute specific transporters, driven by the low intracellular Na^+ concentration resulting from the active transport of Na^+ out of the cell. Cl^- diffuses across the zonula occludens into the lateral intercellular spaces down its electrochemical gradient.

Cl^- movement toward the blood side of the epithelium. In addition, in the early proximal tubule, the selective uptake of Na^+ exceeds that of anions, resulting in a net positive charge on the blood side. This creates a small electrical gradient favoring anion reabsorption. Thus in the early proximal tubule, the chemical and electrical gradients favor Cl^- reabsorption. The zonula occludens is highly permeable to Cl^-; thus passive, paracellular transfer of Cl^- from the tubule lumen to the interstitial fluid occurs. Transcellular Cl^- absorption also occurs in the proximal tubule and is driven by electrical and chemical gradients established by Na^+,K^+-ATPase activity. Cl^- transport across the apical plasma membrane mediated by the Cl^-/anion exchanger, SLC26A6, is one mechanism of luminal uptake. Chloride transport across the basolateral plasma membrane is mediated by K^+, Cl^- cotransporters (KCC), Cl^- channels, and a sodium-dependent Cl^-/HCO_3^- exchanger.

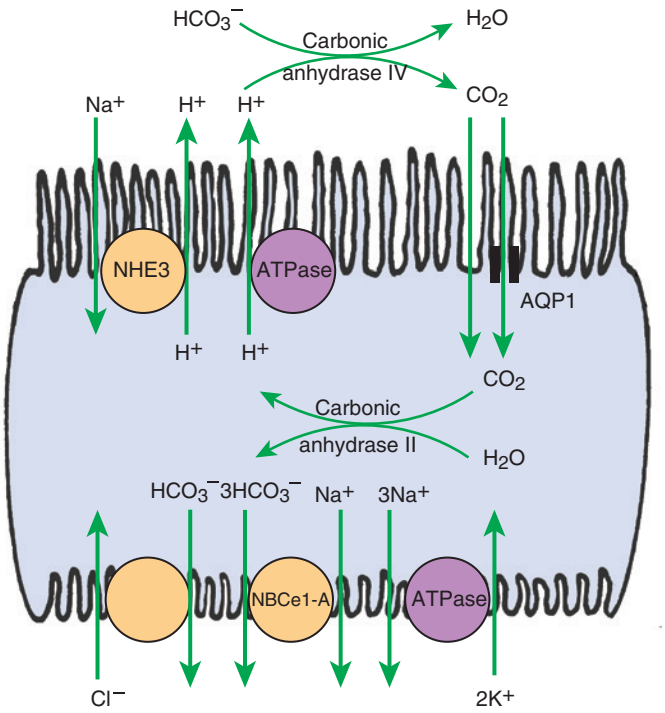
In distal portions of the proximal tubule, Na^+,K^+-ATPase continues to move Na^+ from the cell into the interstitial fluid. However, the luminal fluid is depleted of many of the solutes coupled to Na^+ transport upstream limiting transport by these mechanisms, and transmembrane Na^+ reabsorption occurs predominantly by electrically neutral sodium chloride (NaCl) uptake facilitated by coordinated Na^+- and Cl^--coupled transporters. In particular, in late proximal tubule segments Cl^-/OH^- exchange via SLC26A6, in parallel with Na^+/H^+ exchange via NHE3, yields apical NaCl uptake and H_2O formation in the lumen. Passive reabsorption of Na^+ through the paracellular pathway occurs because of the chemical gradient for Cl^- established by the selective reabsorption of other solutes in the early proximal tubule. As Cl^- moves down its chemical gradient from the tubule lumen to the blood side, it carries Na^+ along with it by electrostatic attraction. The passage of Cl^- down its chemical gradient also abolishes the small, lumen-negative charge and in fact establishes a small, lumen-positive charge in the late proximal tubule, which further favors the passive movement of Na^+ to the blood side.

Other filtered solutes, such as potassium (K^+) and calcium (Ca^{2+}) ions, are present in the tubule fluid in low concentrations and are reabsorbed by the proximal tubule. The proximal tubule reabsorbs approximately 65% of filtered Ca^{2+}. About 90% of the Ca^{2+} uptake in the proximal tubule is paracellular because of a favorable electrochemical gradient in the late proximal tubule and solvent drag. The majority of K^+ reabsorption in the proximal tubule also occurs by passive mechanisms, primarily through the paracellular route.

The proximal tubule also reabsorbs filtered peptides and low–molecular-weight proteins. A large proportion of filtered peptides are degraded to amino acids by peptidases in the proximal tubule brush border and are reabsorbed by cotransport with Na^+ across the apical plasma membrane. Short-chain peptides are themselves transported through cotransport with H^+ on specific transporters (PEPT1 and PEPT2) in the proximal tubule brush border, driven by the proton gradient between the tubule fluid and cytoplasm. Most of these dipeptides and tripeptides are degraded by intracellular peptidases, although some may exit intact to the blood side through another peptide transporter.

Low–molecular-weight proteins are avidly reabsorbed by the proximal tubule but by a different mechanism. Filtered proteins such as insulin, glucagon, parathyroid hormone, and many more are taken up at the apical plasma membrane by receptor-mediated endocytosis (see Fig. 42.4). The proteins bind the receptor proteins, *megalin* or *cubilin,* in the plasma membrane. The bound proteins are endocytosed and delivered by the endocytic vesicles to intracellular organelles called *lysosomes,* whereas the receptors are recycled to the apical plasma membrane (Fig. 42.8). Proteolytic enzymes in the

lysosomes degrade the reabsorbed proteins; the amino acid end products are transported into the interstitial fluid and returned to the blood. Diseased glomeruli often leak protein into the filtrate; in these instances, the proximal tubule endocytic machinery is upregulated and the lysosomal compartment is expanded, often to the extent that the increased number and size of lysosomes in proximal tubules are appreciable in histological sections.

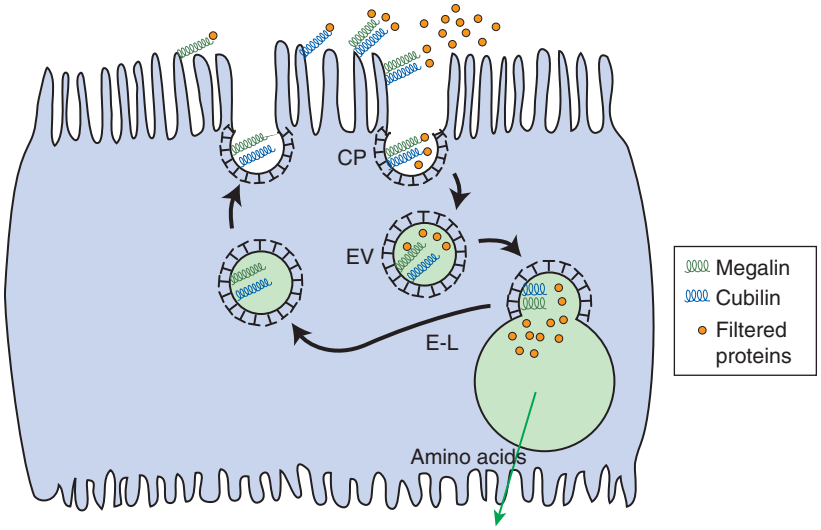


• **Fig. 42.7** Schematic illustration of bicarbonate (HCO_3^-) reabsorption and acid secretion in the proximal tubule. The active reabsorption of Na^+ by the basolateral Na^+,K^+-ATPase pump drives the secretion of H^+ through the Na^+/H^+ exchanger (NHE3) in the apical plasma membrane; apical H^+ATPase also contributes to proximal tubule proton secretion. In the lumen, H^+ and filtered HCO_3^- form H_2O and CO_2, catalyzed by apical membrane–associated carbonic anhydrase. The CO_2 crosses the apical plasma membrane into the cell, possibly facilitated by AQP1 channels. Intracellular CO_2 combines with intracellular H_2O to form H^+ and HCO_3^-, catalyzed by cytoplasmic carbonic anhydrase II. The H^+ is secreted into the tubule fluid, and the HCO_3^- is transported to the blood side of the cell through cotransport with Na^+ (NBCe1-A).

The Proximal Tubule Secretes Organic Ions

The proximal tubule secretes a wide variety of organic ions into the tubule fluid. Many organic ions, including both endogenous waste products and exogenous drugs or toxins, are protein bound in the plasma and thus are poorly filtered by the glomerulus. However, the proximal tubule clears these substances from the blood by basolateral uptake and apical secretion into the tubule fluid by multiple carrier-mediated processes. The carriers include organic anion transporters (basolateral, OAT1, OAT3, OATP4C1, and Na^+-dicarboxylate cotransporter [NaDC3]; apical, OAT4 and URAT1) and organic cation transporters (basolateral, OCT1–3; apical, OCTN1 and 2, P-glycoprotein [Pgp], multidrug resistance transporter 1 [MRP1], and the multidrug and toxin extrusion transporters, MATE1 and MATE2/2K). Endogenous organic compounds secreted by the proximal tubule via these multiple transporters include bile salts, oxalate, urate, creatinine, prostaglandins, epinephrine, and hippurates. Drugs and toxins secreted by the proximal tubule are legion: they include antibiotics (e.g., penicillins, cephalosporins, tetracyclines, sulfonamides), diuretics (e.g., chlorothiazide, furosemide), cardiac glycosides (e.g., digoxin), antiviral agents (e.g., acyclovir, ganciclovir), antihypertensives (e.g., angiotensin-converting enzyme inhibitors and angiotensin receptor blockers), chemotherapeutic agents (methotrexate, azathioprine, cyclophosphamide, 5-fluorouracil), the analgesic morphine and many of its derivatives, the potent herbicide paraquat, and many more.

This aspect of proximal tubule function has broad practical applications. Tubule secretion of endogenous organic ions, drugs, and toxins provides the basis for urine testing for hormones and foreign substances as a reflection of blood levels that may be only transiently elevated. Tubule secretion of exogenous *p*-aminohippurate is used to estimate renal plasma flow. Tubule secretion of certain antibiotics is important in choosing antibiotics that reach high concentrations in the urine for more effective treatment of urinary tract infections. Similarly, secretion of diuretics enhances delivery of these drugs to their site of action, such as furosemide to the thick ascending limb of Henle's loop and thiazides to the distal convoluted tubule. Tubule secretion of drugs determines, in part,



• **Fig. 42.8** Schematic illustration of receptor-mediated endocytosis of filtered proteins in the proximal tubule. Filtered proteins bind with receptors, megalin and cubilin, in the membrane of coated pits (*CP*) in the apical plasma membrane. The coated pits invaginate and form endocytic vesicles (*EV*) that transport the proteins to the endosomal-lysosomal system (*E-L*). The proteins are degraded and the amino acids transported to the interstitium; megalin and cubilin are recycled to the apical plasma membrane.

their excretion rate and affects appropriate dosing, which can be particularly important to consider in patients with compromised renal function. Finally, competitive inhibitors of organic ion secretion raise blood levels and prolong the activity of other drugs excreted by this route and administered simultaneously, which can either create unintended drug toxicities or can be used for therapeutic advantage.

Tubule secretion plays a larger role in birds than in mammals. The end product of protein metabolism in mammals is *urea,* which is excreted primarily through glomerular filtration. In birds, the end product of protein metabolism is *uric acid.* Uric acid is produced by the liver and the kidney in birds and is excreted primarily by tubule secretion. In fact, in starlings the total amount of uric acid excreted by the kidney is more than five times the amount filtered. The principal site of uric acid secretion in the avian kidney is believed to be the proximal portion of reptilian-type nephrons. Blood levels of uric acid are used as a measure of renal function in birds because clearance measurements are impractical or inaccurate measures of GFR.

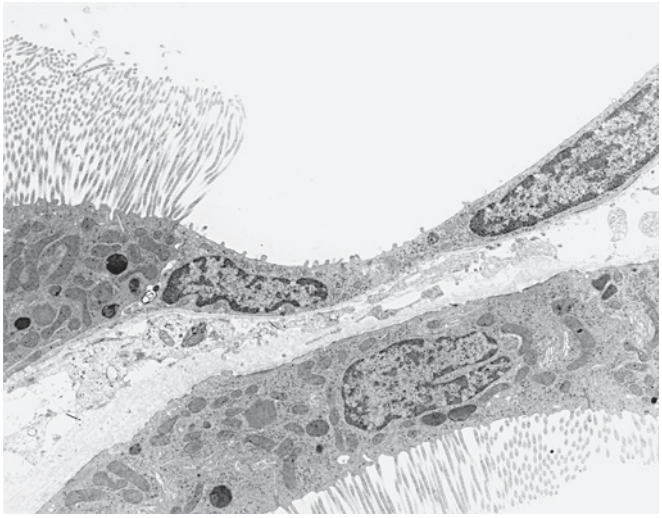
The Thick Ascending Limb and Distal Convoluted Tubule Reabsorb Salts and Dilute the Tubule Fluid

The structure of the tubule epithelium changes abruptly and dramatically at the end of the proximal tubule. The proximal tubule, with its abundant mitochondria, luxuriant brush border, and pronounced basolateral plasma membrane infoldings, is suited for high-volume transport of a large variety of substances by both active and passive mechanisms. The segments that follow the proximal tubule each have a unique structure, in keeping with their specialized functions. Immediately downstream from the straight portion of the proximal tubule is the thin limb of Henle's loop, which is a low, flat epithelium with few mitochondria and few membranous infoldings (Fig. 42.9). As might be expected, active transport of solutes in this segment is virtually nonexistent. The function of the thin limb is determined by its spatial orientation in the medulla and by the segmental distribution of specific solute and water transporters, which convey the passive permeability properties of the thin limb segments. These characteristics are essential to its role in the urine concentrating mechanism and are discussed in Chapter 43.

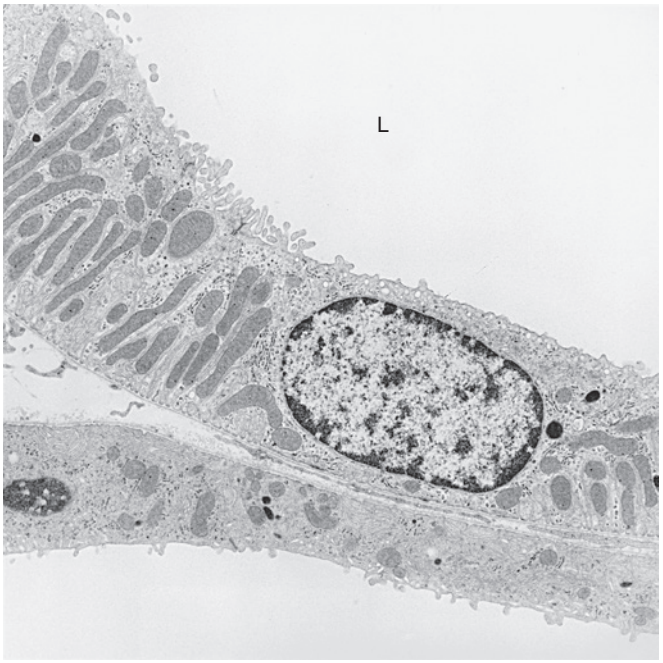
In the ascending limb of Henle's loop, the low, flat epithelium of the thin limb abruptly changes to the taller, cuboidal epithelium of the thick ascending limb (TAL). The TAL has many mitochondria and basolateral plasma membrane infoldings, reflecting its high capacity for active solute transport (Fig. 42.10). The distal convoluted tubule (DCT) follows with an even taller epithelium and a dense array of mitochondria. Next is the connecting segment, a segment with a heterogeneous cell population that connects the nephrons to the collecting duct system.

The TAL of Henle's loop and the DCT reabsorb Na^+, Cl^-, and the divalent cations Ca^{2+} and Mg^{2+}. These segments reabsorb Na^+ and Cl^- against a high gradient, and unlike the proximal tubule, the tight junctions and the plasma membrane are impermeable to water. By the time the tubule fluid leaves the DCT, more than 90% of the filtered salts have been reabsorbed, and the osmolality of the tubule fluid is typically reduced from approximately 300 to 100 mOsm/kg H_2O.

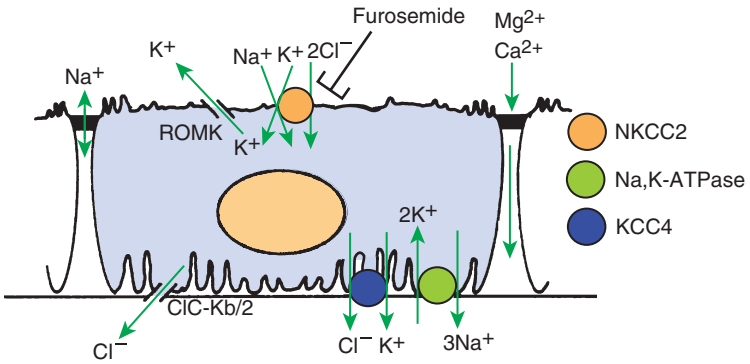
As in the proximal tubule, the driving force for solute reabsorption in the TAL and DCT is Na^+,K^+-ATPase in the basolateral plasma membrane. In the TAL, the electrochemical gradient for Na^+ established by basolateral Na^+,K^+-ATPase activity drives ion uptake through the apical Na^+,K^+,$2Cl^-$ cotransporter (NKCC2) (Fig. 42.11). Intracellular Cl^- moves down its chemical gradient into the interstitial fluid primarily through heteromeric ClC-Kb/2 Cl^-channels in the basolateral plasma membrane. Intracellular K^+ diffuses down its concentration gradient primarily through basolateral KCl cotransport (KCC4), which mediates K^+ reabsorption accounting for up to 30% of the filtered load, depending



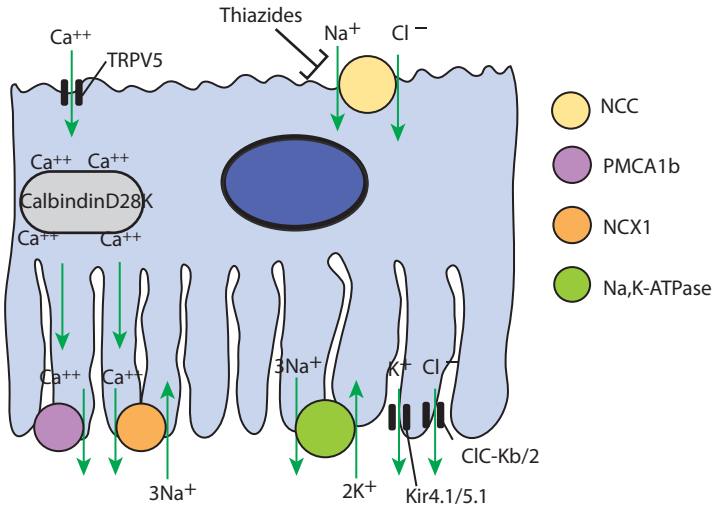
• **Fig. 42.9** Transmission electron micrograph of rat kidney illustrating the transition from the proximal tubule to the thin descending limb of Henle's loop. The tall epithelium of the proximal tubule with the extensive brush border and abundant mitochondria abruptly changes to the low epithelium of the thin limb of Henle's loop. Epithelial cells of the thin limb have a smooth, simple plasma membrane surface and few mitochondria, which is consistent with the apparent absence of significant active transport.



• **Fig. 42.10** Transmission electron micrograph of thick ascending limb of Henle's loop in rat kidney. In accordance with its important role in active Na^+ reabsorption, the thick ascending limb is a tall epithelium, with extensive basolateral plasma membrane infoldings and numerous mitochondria. A collecting duct is adjacent to the basolateral aspect of the thick limb. *L,* Tubule lumen.



• **Fig. 42.11** Schematic illustration of transport functions of the thick ascending limb of Henle's loop. Na^+ is actively reabsorbed through the basolateral Na^+,K^+-ATPase pump. Na^+, K^+, and Cl^- enter the cell from the luminal fluid through secondary active cotransport via the Na^+, K^+, 2 Cl^- cotransporter, NKCC2. Cl^- exits through the basolateral Cl^- channel, ClC-Kb/2. K^+ leaves the cell down its concentration gradient primarily by basolateral KCl cotransport (via KCC4), and also through apical K^+ channels (ROMK), which facilitates continued NaCl transport on NKCC2. A lumen-to-blood gradient for cations is established and drives reabsorption of Ca^{2+} and Mg^{2+} through cation-selective paracellular channels in the tight junction formed by claudins. Na^+ also crosses paracellular channels, initially from lumen to blood, but as the tubule fluid becomes more dilute, paracellular Na^+ back-leak occurs. Loop diuretics, such as furosemide, inhibit NKCC2.



• **Fig. 42.12** Schematic illustration of transport functions of the distal convoluted tubule. Na^+ is actively reabsorbed by the basolateral Na^+,K^+-ATPase pump. Na^+ and Cl^- enter the cell from the luminal fluid through secondary active cotransport via the thiazide-sensitive Na^+, Cl^- cotransporter, NCC. Cl^- exits through basolateral ClC-Kb/2 Cl^- channels. K^+ is recycled to the interstitium through basolateral K^+ channels (Kir 4.1/5.1) and possibly a KCl cotransporter. Calcium uptake is driven by basolateral Ca^{2+}ATPase (PMCA1b) and Na^+,K^+-ATPase, which drive Ca^{2+} uptake through the basolateral Na^+/Ca^{2+} exchanger (NCX1) and apical Ca^{2+} channel (TRPV5). Calbindin D28k binds free Ca^{2+}, thereby preventing increases in cytoplasmic free Ca^{2+} and subsequent suppression of apical TRPV5.

on the systemic K^+ status. Diffusion through apical K^+ channels primarily $K_{ir}1.1$, also known as renal outer medullary potassium channel (ROMK) into the lumen allows continued transport of Na^+ and Cl^- on NKCC2 despite depletion of luminal K^+. The Cl^- absorption and K^+ secretion cause a lumen-positive voltage. The electrical gradient drives lumen-to-interstitium diffusion of the divalent cations, Ca^{2+}, Mg^{2+}, Na^+, and K^+ through cation-selective paracellular channels formed by claudin proteins in the tight junctions. The apical Na^+,K^+,2Cl^- cotransporter in the TAL is inhibited by the *loop* diuretics (named for Henle's loop), such as bumetanide and furosemide, which are often used in clinical veterinary medicine.

The DCT contains an apical NaCl cotransporter (NCC) that mediates Na^+ and Cl^- uptake from the tubule fluid and is inhibited by thiazide diuretics. NaCl cotransport is driven by the transmembrane chemical gradient for Na^+ generated by the basolateral Na^+,K^+-ATPase (Fig. 42.12). Potassium is recycled down its chemical gradient into the interstitial fluid via basolateral K^+ channels (Kir 4.1/5.1) and possibly a KCl cotransporter; in late DCT, apical ROMK mediates potassium secretion. As in the TAL, Cl^- exits to the interstitium largely through basolateral ClC-Kb/2 Cl^- channels, driven by the electrical gradient created by Na^+ and K^+ extrusion across the basolateral plasma membrane.

Unlike the TAL, where calcium and magnesium uptake occurs through the paracellular route, calcium and magnesium reabsorption in the DCT is transcellular and mediated by specific membrane transport proteins; these are discussed in detail in a later section.

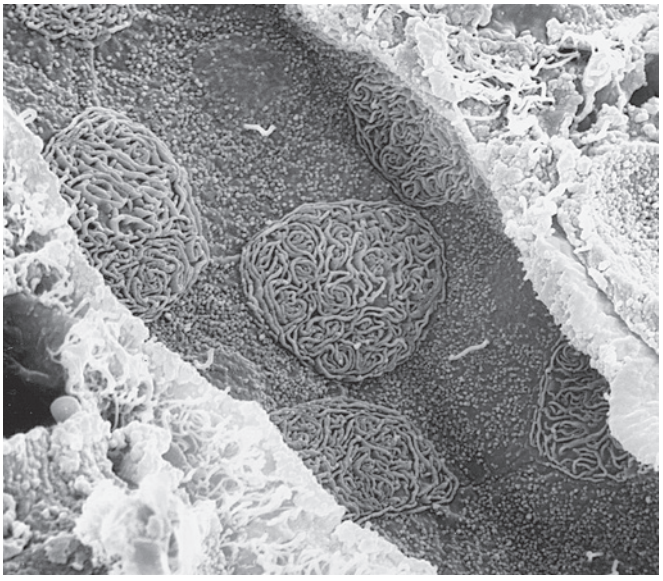
Both the TAL and the DCT are impermeable to water. The avid reabsorption of salt without water produces a hypotonic tubule fluid; thus these segments are sometimes called the *diluting segments.* Dilution of the tubule fluid takes place regardless of the volume status of the animal. It is an important component of fluid volume regulation, allowing the kidney to excrete excess water without salt, thereby preventing plasma hypotonicity when water intake is high, and also generates a hypertonic medullary interstitium, which is necessary for urine concentration and water conservation. The generation of hypotonic tubule fluid and avid NaCl reabsorption in the TAL and DCT set the stage for the final regulation of water excretion in the collecting duct, discussed further in Chapter 43.

The Collecting Duct Reabsorbs Sodium Chloride

The collecting duct system begins with the connecting segment, which follows the DCT. The tubules of individual nephrons begin merging in the connecting segment and the downstream initial collecting tubule. Depending on the species, the connecting segment contains several distinct epithelial cell types, including DCT cells, connecting segment cells, intercalated cells, and principal cells. Each of these structurally distinct cell types has specific physiological functions.

The initial collecting tubules merge into the collecting duct, which traverses the cortex and medulla to the papillary tip, where the tubule fluid (urine) discharges into the renal pelvis. In most of the collecting duct, there are two main cell types: the *intercalated cell,* which has many intracytoplasmic vesicles, mitochondria, and a relatively complex apical surface, and the *principal cell,* which has fewer intracytoplasmic vesicles and mitochondria and a relatively smooth apical surface but generally more extensive basolateral plasma membrane infoldings (Fig. 42.13). The principal cell is the major cell type in the initial collecting tubule and the cortical and outer medullary collecting ducts, accounting for approximately two-thirds of the cells in most regions. Intercalated cells account for the remainder of the cortical and outer medullary collecting duct cells.

Sodium reabsorption in the collecting duct is primarily a function of principal cells and is driven by basolateral Na^+,K^+-ATPase. As in other tubule segments, Na^+ is actively transported by this pump into the interstitial fluid, establishing an electrochemical gradient that promotes Na^+ uptake through apical *epithelial Na^+ channels* (ENaC). The resulting lumen-negative electrical potential drives Cl^- absorption through the paracellular pathway, through channels in the tight junctions formed by claudins; approximately 70% of collecting duct Cl^- reabsorption is through the paracellular pathway. A subpopulation of intercalated cells, *type B intercalated cells,* also contribute to NaCl reabsorption in the collecting duct. Type B intercalated cells mediate Cl^- reabsorption via an apical Cl^-/HCO_3^- exchanger, *pendrin,* and the basolateral ClC-Kb/2 Cl^- channel. In addition to transporting Cl^- from the lumen, pendrin activity in type B intercalated cells enhances ENaC abundance, open probability, and activity in principal cells. The molecular mechanisms by which pendrin activity in type B intercalated cells alters ENaC activity in principal cells have not been fully determined. There also is evidence that pendrin activity coordinates with the Na-dependent Cl^-/ HCO_3^- exchanger (NDCBE) in the apical plasma membrane of the type B intercalated cells. Thus enhanced pendrin activity promotes not only Cl^- reabsorption, but also Na^+ reabsorption via ENaC and possibly NDCBE.



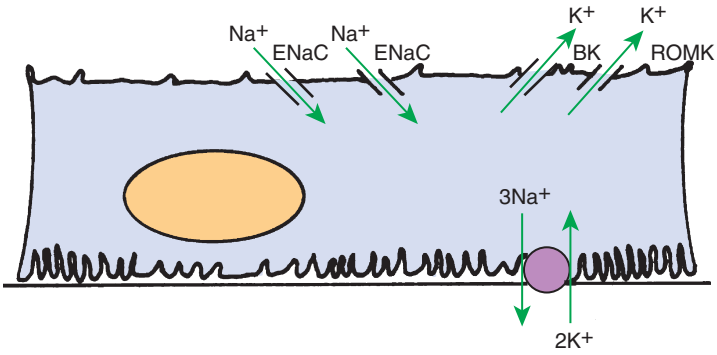
• **Fig. 42.13** Scanning electron micrograph of outer medullary collecting duct in the rat, viewed from the luminal surface. Two cell types are evident: the principal cell, with short, small projections over the apical surface and a single central cilium; and the intercalated cell, with extensive, complex membrane folds (microplicae) over the apical surface.

The Distal Convoluted Tubule and Collecting Duct Control Renal Potassium Excretion

Control of net renal K^+ excretion is another function of the collecting duct. In principal cells, basolateral Na^+,K^+-ATPase actively pumps K^+ into the cell, raising the intracellular K^+ concentration above that of the interstitial and tubule fluid. Intracellular K^+ exits the cell down the chemical gradient through K^+ channels present in the apical (primarily ROMK; also the maxi-K^+ channel, BK) and basolateral (partly Kir4.1/Kir5.1) plasma membranes. Under normal circumstances, however, net K^+ secretion occurs for two reasons: (1) the apical K^+ channel, ROMK, is more permeable than the basolateral K^+ channel(s), and (2) the lumen-negative electrical potential favors K^+ secretion (Fig. 42.14).

The collecting duct can also reabsorb K^+. Intracellular potassium is actively transported in exchange for hydrogen ions in the tubule fluid by apical H^+,K^+-ATPase isoforms similar to that in the gastric parietal cell. H^+,K^+-ATPases are most abundant in intercalated cells, but are also present in principal cells. When dietary potassium is restricted, H^+,K^+-ATPase activity and expression in the collecting duct is upregulated and apical K^+ channel (ROMK in principal cells and BK in intercalated cells and principal cells) activity is inhibited; these effects enhance net K^+ uptake from the lumen and favor K^+ exit through basolateral K^+ channels, such as Kir 4.1/5.1 in principal cells, thus promoting K^+ reabsorption.

Proximal to the collecting duct, the DCT has an important role in regulating renal K^+ excretion. The basolateral Na^+,K^+-ATPase pumps K^+ into the cell and the basolateral potassium channel, Kir



• **Fig. 42.14** Schematic illustration of transport in the principal cell of the collecting duct. Basolateral Na^+,K^+-ATPase actively transports Na^+ and drives passive diffusion of Na^+ from the tubule lumen into the cell through a Na^+-selective channel, ENaC, in the apical plasma membrane. K^+-selective channels (ROMK, BK) in the apical plasma membrane enable K^+ secretion into the tubule fluid. The hormone aldosterone enhances Na^+,K^+-ATPase and ENaC channel activity and increases K^+ permeability of the apical plasma membrane, thus enhancing Na^+ reabsorption and K^+ secretion.

4.1/5.1 not only returns intracellular K^+ to the interstitium, but also senses changes in extracellular K^+ and regulates a signaling cascade that modifies NCC phosphorylation and alters NCC transport activity. Because the resultant change in Na^+ delivery to the collecting duct downstream influences K^+ secretion in that segment, plasma K^+ regulation of NCC activity through Kir 4.1/5.1 affects both K^+ and Na^+ excretion. A slight elevation in plasma K^+ from ingestion of a high K^+ diet increases apical ROMK in the late DCT, enabling increased DCT K^+ secretion, increases plasma aldosterone (discussed later), and reduces NCC phosphorylation and activity, much like a thiazide diuretic. The increased aldosterone plus increased Na^+ delivery to, and fluid flow in, the collecting duct favors K^+ secretion in the collecting duct. Increased aldosterone and increased Na^+ delivery also increase fractional reabsorption of Na^+ in the collecting duct, but the net effect is increased renal Na^+ excretion. Through these mechanisms, the effect of a high K^+ diet is increased renal excretion of Na^+, through inhibition of NCC in the DCT, and increased renal excretion of K^+, through stimulation of K^+ secretion in the DCT and collecting duct. Low K^+ diets have the converse effect.

Solute Transport Is Regulated by Systemic and Intrarenal Signals

In the proximal tubule, filtered solutes and water are reabsorbed regardless of the animal's physiological state, but specific hormones regulate the rate of reabsorption of sodium, chloride, phosphate, and other solutes. The distal tubule and collecting duct control the ultimate rate of excretion of electrolytes and water to maintain homeostasis despite variations in dietary intake and extrarenal losses of salts and water. The specific homeostatic responses are controlled in large part by several hormones, including angiotensin II, aldosterone, antidiuretic hormone (ADH), endothelin-1, atrial natriuretic peptide (ANP), parathyroid hormone (PTH), $1\alpha,25\text{-}(OH)_2$–vitamin D_3, and calcitonin. Many of these hormones are produced exclusively by other organs and delivered to the kidney through the circulation. Others, such as angiotensin II and endothelin-1, are produced at least partly in the kidney and exert local effects on renal transport.

In birds, the relative importance of the reptilian-type and the mammalian-type nephrons in regulation of salt balance has not been established. Furthermore, in many species of birds, particularly marine and desert species, sodium balance is regulated largely by secretion of NaCl by the nasal (supraorbital) gland, rather than by regulation of renal excretion. Finally, ureteral urine delivered to the cloaca moves in a retrograde manner into the digestive tract, where additional salt reabsorption occurs; the importance of this mechanism varies among avian species.

Angiotensin II Stimulates Sodium Uptake in Proximal Tubule, Distal Nephron, and Collecting Duct

Angiotensin II directly enhances sodium reabsorption in the proximal tubule, TAL of Henle's loop, DCT, and the collecting duct. These segments contain specific angiotensin II receptors (AT_1 receptors) that, when activated, increase Na^+ transport. In the proximal tubule, angiotensin II stimulates Na^+ uptake through the apical Na^+/H^+ exchanger and the basolateral $Na^+(HCO_3^-)_3$ cotransporter and Na^+,K^+-ATPase. Angiotensin II also increases expression of the apical Na^+/H^+ exchanger and the $Na^+,K^+,2Cl^-$ transporter in the TAL. In the DCT, it increases the apical plasma membrane expression of the NCC. In the collecting duct, angiotensin II stimulates NaCl reabsorption by enhancing Na^+ transport via ENaC in principal cells and apical Cl^-/HCO_3^- exchange via pendrin in intercalated cells.

Aldosterone Enhances Sodium Reabsorption and Potassium Secretion

Aldosterone is a mineralocorticoid hormone secreted by the adrenal cortex. Dietary sodium chloride restriction and systemic hypotension stimulate aldosterone release by activating the renin-angiotensin system. Aldosterone activates the mineralocorticoid receptor in connecting segment cells and principal cells of the collecting duct, stimulating signal transduction pathways that enhance Na^+ reabsorption, which in turn enhances water reabsorption to increase fluid volume. At the cellular level, aldosterone's major effect on Na^+ transport is to increase the number of ENaC Na^+ channels in the principal cell apical plasma membrane, thereby enhancing Na^+ reabsorption. In addition, aldosterone acutely increases the *open probability* of ENaC and stimulates basolateral Na^+,K^+-ATPase activity. Chronic aldosterone stimulation causes proliferation of the principal cell basolateral plasma membrane, increased Na^+,K^+-ATPase abundance, and increased apical plasma membrane expression of the NCC in the DCT and the Cl^-/HCO_3^- exchanger, pendrin, in intercalated cell subtypes, all of which contribute to enhanced NaCl reabsorption.

Aldosterone release is also stimulated by hyperkalemia (elevated plasma K^+ level) and has an important role in regulating K^+ homeostasis. Aldosterone increases basolateral K^+ entry into principal cells by stimulating Na^+,K^+-ATPase activity. Enhanced apical ENaC activity and luminal Na^+ uptake create a favorable electrical gradient for K^+ secretion through apical K^+ channels. Aldosterone also increases the activity of the apical K^+ channel, ROMK. The result is enhanced urinary K^+ excretion.

Other Hormones and Ligands That Regulate Sodium Transport Include Antidiuretic Hormone, Nitric Oxide, Endothelin-1, and Atrial Natriuretic Peptide

In some species, *antidiuretic hormone* (ADH, vasopressin), which is released when an animal is volume depleted, dehydrated, or hypotensive, enhances salt reabsorption from the TAL and the collecting duct. The increased salt transport partly results from vasopressin-stimulated increases in the activity of the apical Na^+,K^+,$2Cl^-$ cotransporter (NKCC2) in the TAL and ENaC in the collecting duct. Although ADH stimulation of salt reabsorption in the TAL has the seemingly paradoxical effect of enhancing the dilution of the tubule fluid, this in fact allows maximal salt and water conservation because the increased salt uptake contributes to the interstitial osmolarity and enables enhanced water reabsorption in the collecting ducts (see Chapter 43).

Nitric oxide (NO) is a gas produced by the catabolism of L-arginine, catalyzed by nitric oxide synthase (NOS) in renal endothelial and epithelial cells. NO increases renal Na^+ and water excretion by inhibiting Na^+ uptake mechanisms in several renal tubule segments. In the proximal tubule, NO inhibits Na^+ reabsorption by inhibiting apical NHE3 and basolateral Na^+,K^+-ATPase activity. In the thick ascending limb, NO inhibits the apical Na^+ transporters, NKCC2 and NHE3; in the collecting duct, NO inhibits Na^+ uptake via ENaC. NO-mediated inhibition of Na^+ reabsorption in these segments plays an important role in regulating systemic extracellular fluid volume and blood pressure.

Endothelin-1 is a peptide hormone produced in the kidney, primarily in the collecting duct, endothelial cells, and the TAL of Henle's loop. Endothelin-1 binding to ET-B receptors in the proximal tubule, collecting duct, and TAL increases renal NaCl and water excretion by effects on epithelial transport and renal microcirculation, mediated at least in part by stimulation of NO synthesis. The transport mechanisms inhibited include NHE3 in the proximal tubule, NKCC2 in the TAL, ENaC in the collecting duct, and Na^+,K^+-ATPase in all three of these segments.

Atrial natriuretic peptide (ANP) is produced in the cardiac atria. ANP release is stimulated by atrial distention in healthy subjects and plasma ANP levels are elevated in patients with congestive heart failure and other conditions causing extracellular fluid retention. ANP inhibits aldosterone and renin release, and increases renal Na^+ excretion, primarily by inhibition of ENaC and Na^+,K^+-ATPase activity in the medullary collecting duct. ANP also inhibits Na^+ transport by NKCC2 in the TAL and by Na^+,K^+-ATPase in the proximal tubule.

Phosphate Uptake in the Proximal Tubule Is Decreased by Parathyroid Hormone

The majority of filtered phosphate is reabsorbed in the proximal tubule, but the rate of reabsorption is regulated by several factors. Filtered phosphate is reabsorbed via Na-coupled phosphate transporters (NaPi2a, NaPi2c, PiT2) located in the proximal tubule brush border; the identity of the basolateral phosphate transporter is not yet known. Regulation of phosphate uptake is mediated by changes in the apical abundance of these transporters. PTH decreases brush border NaPi2a, NaPi2c, and PiT2, thereby decreasing phosphate uptake and increasing urinary phosphate excretion. PTH administration stimulates removal of NaPi2a from the brush border within minutes and removal of NaPi2c and PiT2 within hours.

Increased serum phosphate stimulates the release of *fibroblast growth factor-23* (FGF23) from osteoblasts. FGF23 and its cofactor, Klotho, activate the FGF receptor in proximal tubule cells and reduce the expression and activity of NaPi2a, NaPi2c, and PiT2, thus reducing phosphate reabsorption and enhancing renal phosphate excretion.

Several other factors regulate phosphate uptake in the proximal tubule. In addition to PTH and FGF23/Klotho, factors that downregulate proximal tubule phosphate uptake include dietary potassium deficiency, metabolic acidosis, hypertension, estrogen, and glucocorticoids. Factors that upregulate proximal tubule phosphate uptake include dietary phosphorus deficiency, thyroid hormone, and vitamin D_3.

Calcium Reabsorption in the Distal Nephron and Connecting Segment Is Enhanced by Parathyroid Hormone, Vitamin D_3, and Calcitonin

The kidney reabsorbs the majority of filtered calcium (Ca^{2+}) and contributes significantly to the regulation of systemic Ca^{2+} balance. Approximately 60% to 70% of filtered Ca^{2+} is absorbed in the proximal tubule; about 80% of Ca^{2+} reabsorption in the proximal tubule is paracellular and passive, driven by electrical and chemical gradients. A small component of proximal tubule Ca^{2+} reabsorption is transcellular, mediated by active transport, and regulated by PTH and calcitonin. Approximately 20% of filtered Ca^{2+} is reabsorbed in the TAL, primarily in the cortical segment. Ca^{2+} reabsorption in this segment occurs almost exclusively by passive, paracellular means, driven by electrochemical gradients (see Fig. 42.11). TAL Mg^{2+} uptake is also primarily paracellular and reclaims approximately 50% to 60% of filtered Mg^{2+}. Ca^{2+} reabsorption in the TAL is suppressed when serum Ca^{2+} is elevated, through activation of the basolateral calcium-sensing receptor (CaSR), which inhibits sodium chloride uptake in the TAL and reduces the electrical gradient driving paracellular Ca^{2+} uptake. In addition, the permeability of Ca^{2+} channels formed by claudins in the tight junctions is altered by stimulation of the basolateral CaSR.

The DCT and the connecting segment together reabsorb approximately 10% to 20% of filtered Ca^{2+}. Unlike the proximal tubule and TAL, calcium transport in the DCT is almost exclusively transcellular. Uptake occurs via apical Ca^{2+} channels (TRPV5); intracellular Ca^{2+} is bound by the cytoplasmic Ca^{2+}-binding protein, calbindin D28k, which is thought to facilitate Ca^{2+} diffusion to the basolateral region and may also help maintain Ca^{2+} absorption from the lumen by preventing increases in intracellular free Ca^{2+} levels, which would otherwise cause a decrease in apical TRPV5 expression. Ca^{2+} exits across the basolateral plasma membrane via active transport by Ca^{2+}ATPase (PMCA1b) and via Na^+/ Ca^{2+} exchange (NCX1) (see Fig. 42.12). Apical magnesium uptake is mediated by Mg^{2+} channels (TRPM6); basolateral Mg^{2+} exit is mediated by a Mg^{2+}/Na^+ exchanger. Only 1% to 2% of the filtered Ca^{2+} is reabsorbed in the collecting ducts, by mechanisms that have not been identified.

Regulation of Ca^{2+} transport primarily occurs in the DCT, the connecting segment, and the cortical TAL. PTH, 1α,25-$(OH)_2$–vitamin D_3, and calcitonin have important roles in controlling renal Ca^{2+} excretion. A decrease in plasma Ca^{2+} stimulates PTH

release, which affects bone, the intestines, and the kidneys to raise the plasma Ca^{2+} level. The major renal response occurs in the cortical TAL, DCT, and connecting segment by actions mediated through the PTH receptor, PTH1R. In the cortical TAL, PTH1R signaling alters the expression of claudins, which form the Ca^{2+}-channels in the tight junctions, thereby increasing paracellular Ca^{2+} permeability and Ca^{2+} reabsorption. In the DCT and connecting segment where Ca^{2+} reabsorption is transcellular, PTH1R activation increases the activity of the apical Ca^{2+} channel, TRPV5, and the expression of calbindin D28K, thus increasing apical Ca^{2+} uptake and cytoplasmic Ca^{2+} levels. PTH also increases the Cl^- conductance at the basolateral plasma membrane in the DCT, which hyperpolarizes the cells (the interior becomes more electrically negative) and thus increases the driving force for Ca^{2+} entry.

The hormone, *vitamin D,* is converted to its active form, 1α,25-$(OH)_2$–*vitamin* D_3 (*calcitriol*), in the proximal convoluted tubule; this process is stimulated by PTH and modulated by vitamin D3 through a feedback system that regulates the enzymes involved in its conversion. Receptors for vitamin D_3 are located predominantly in the DCT and connecting segment, where vitamin D_3 increases the cellular expression of the apical Ca^{2+} channel (TRPV5), the Ca^{2+}-binding protein, calbindin D28k, the basolateral plasma membrane Ca^{2+}-ATPase (PMCA1b), and the basolateral Na^+/ Ca^{2+} exchanger (NCX1), thus enhancing renal Ca^{2+} reabsorption.

Calcitonin reduces the serum Ca^{2+} concentration, largely by decreasing osteoclast-mediated bone resorption and thus increasing net Ca^{2+} deposition in bone. Calcitonin reduces renal Ca^{2+} excretion by enhancing Ca^{2+} reabsorption in the TAL and the DCT via mechanisms that are not yet characterized.

CLINICAL CORRELATIONS

Glucosuria

History

A client presents her 10-year-old female miniature schnauzer with the complaint of a dramatic increase in water consumption and urine volume over the previous 2 weeks.

Clinical Examination

No major abnormalities are found on physical examination. The dog appears alert and is moderately overweight. The urinalysis reveals 4 + glucose (normally negative) and urine specific gravity of 1.030. The plasma glucose level is tested immediately and is 275 mg/dL (normal, 80–120 mg/dL).

Comment

The dog has *diabetes mellitus,* which results from a relative or absolute deficiency of insulin secreted from the β cells of the pancreas, similar to type 1 diabetes mellitus in humans, or by a state of insulin resistance, similar to type 2 diabetes mellitus in humans. Regardless of the cause, the effective insulin deficiency results in elevated plasma glucose levels. Glucose is freely filtered by the glomerulus and normally is entirely reabsorbed by the proximal tubule, via Na^+-coupled apical uptake by SGLT1/2 glucose transporters, and basolateral facilitated diffusion via GLUT1/2 glucose carriers. As the plasma glucose level rises, the glucose concentration in the glomerular filtrate rises. When it exceeds the reabsorptive capacity of the proximal tubule (the renal threshold) of approximately 180 mg/dL, glucose appears in the urine (*glucosuria*). The glucose acts as an osmotic agent and increases the volume of urine excreted. The dog then drinks more water to replace the excessive fluid loss.

Treatment

Treatment of diabetes mellitus in veterinary patients usually involves the administration of insulin injections two to three times each day, with adjustments of the dose in accordance with frequent evaluations of the plasma or urine glucose values. When the insulin dosage is appropriate, the plasma glucose level normalizes, the glucosuria disappears, and the urine volume and water consumption decrease.

Hypoadrenocorticism

History

A concerned client presents her 1-year-old spayed female Samoyed with the complaint of severe weakness, inappetence, and vomiting since the previous day.

Clinical Examination

The dog is lethargic, weak, and extremely dehydrated. The heart rate is normal, but pulses are weak. No other abnormalities are detected on physical examination. Samples of blood and urine are collected immediately for a complete blood count, serum chemistry profile, and urinalysis, and then an intravenous catheter is placed to start volume replacement therapy with a balanced electrolyte solution. The urinalysis is normal, with a specific gravity of 1.025. Abdominal radiographs are normal, but the thoracic radiographs demonstrate a small cardiac silhouette and small thoracic vessels. The serum creatinine level is 2.5 mg/dL (normal, 0.6–1.2 mg/dL), serum K^+ level is 6.5 mEq/L (normal, 3.6–5.6 mEq/L), serum Na^+ is 129 mEq/L (normal, 141–155 mEq/L), serum Cl^- is 97 mEq/L (normal, 103–115 mEq/L), and serum HCO_3^- is 12 mEq/L (normal, 18–24 mEq/L).

Comment

The dog has *hypoadrenocorticism.* The metabolic disturbances result from a deficiency of the mineralocorticoid hormone aldosterone. In a normal animal, aldosterone stimulates connecting segment and collecting duct Na^+,K^+-ATPase activity and enhances apical Na^+ uptake by increasing apical ENaC abundance and open probability in principal cells. These effects on Na^+ reabsorption promote K^+ secretion by increasing active transport of K^+ into the cell from the interstitium and by creating a favorable gradient for apical K^+ exit. Aldosterone also enhances apical ROMK channel activity, further favoring K^+ secretion. When aldosterone is deficient, Na^+ conservation and K^+ secretion in these segments are impaired, and hyponatremia and hyperkalemia can occur. Cl^- uptake via pendrin is impaired by aldosterone deficiency and passive movement of Cl^- and water follows the path of Na^+, thus renal Cl^- and water excretion increases. *Hyperkalemia* (high serum K^+ level) has profound effects on excitable tissue, including nerve and muscle cells, and results in muscle weakness, decreased cardiac output, hypotension, and cardiac arrhythmias. The loss of NaCl and water results in volume depletion and decreased size of the heart and thoracic blood vessels, and exacerbates the hypotension and poor tissue perfusion.

Poor perfusion of the kidneys is probably the major cause of the elevated serum creatinine level (*azotemia*) because inadequate renal blood flow and reduced glomerular capillary hydrostatic pressure prevent adequate glomerular filtration. This is called prerenal azotemia. In most cases of *prerenal azotemia* the urine is maximally concentrated in an attempt to retain fluid and restore blood volume, but in hypoadrenocorticism, this response is often blunted, possibly by the *hyponatremia* (reduced serum Na^+ level) or by the absence of glucocorticoids, which may be necessary for maximal urine concentration (see Chapter 43). The decreased serum bicarbonate level indicates metabolic acidosis, which is the result of both the diminished renal ability to secrete H^+ and reabsorb HCO_3^- (see Chapter 44) and the increased production of acid from poorly perfused tissue.

Treatment

Prompt treatment is important for the animal's survival because the hyperkalemia and acidosis can cause fatal cardiac arrhythmias. Volume repletion with normal saline and correction of the *base deficit* (low serum HCO_3^- level) often stabilizes the animal. Replacement therapy with mineralocorticoid hormones (e.g., desoxycorticosterone acetate, desoxycorticosterone pivalate, fludrocortisone acetate) restores apical Na^+

Continued

CLINICAL CORRELATIONS—cont'd

channel and basolateral Na^+,K^+-ATPase activity and should be started as soon as possible. Frequently, glucocorticoid hormones are given early to treat shock even before the electrolyte status is known; these hormones are beneficial for two reasons. First, hypoadrenocorticism usually results in glucocorticoid deficiency, sometimes manifested by hypoglycemia, and replacement therapy is indicated. Second, the mineralocorticoid activity in many glucocorticoid preparations may be beneficial in correcting the hyperkalemia and hyponatremia.

A definitive diagnosis may be obtained by an adrenocorticotropic hormone (ACTH) challenge test, which maximally stimulates the release of cortisol from the adrenal gland and shows little or no response in animals with hypoadrenocorticism.

Chronic maintenance usually involves oral therapy with fludrocortisone acetate; the proper dosage is determined by periodic evaluations of the serum K^+ and Na^+ levels. Chronic replacement of glucocorticoids is also recommended.

Practice Questions

1. Which segment of the renal tubule is responsible for the reabsorption of the bulk of filtered solutes?
- **a.** Proximal tubule
- **b.** Thin limbs of Henle's loop
- **c.** Thick ascending limb of Henle's loop
- **d.** Distal convoluted tubule
- **e.** Collecting duct

2. The main driving force for the reabsorption of solutes from the tubule fluid is:
- **a.** Active transport of solutes across the apical plasma membrane.
- **b.** Secondary active transport of solutes across the apical plasma membrane.
- **c.** Active transport of Na^+ from the tubule epithelial cell across the basolateral plasma membrane by the electrogenic Na^+ channel.
- **d.** Active transport of Na^+ from the tubule epithelial cell across the basolateral plasma membrane by the Na^+,K^+-ATPase pump.
- **e.** Passive diffusion of solutes through the paracellular pathway.

3. Glucose is found in the urine of an animal when:
- **a.** Glucose transporters in the proximal tubule are inhibited by furosemide.
- **b.** Glucose secretion in the proximal tubule is stimulated by angiotensin II.
- **c.** Glomerular filtration barrier is defective causing increased glucose in the tubule fluid.
- **d.** Plasma glucose is elevated, increasing glucose concentration in the tubule fluid above the proximal tubule transport capacity.
- **e.** Elevated plasma glucose stimulates proximal tubule glucose secretion.

4. The ultimate rate of excretion of K^+ in the urine is determined by the: (Choose all correct answers)
- **a.** Concentration of K^+ in the glomerular filtrate.
- **b.** Proximal tubule, which reabsorbs or secretes K^+ to meet the physiological requirements of the animals.
- **c.** Thick ascending limb, where K^+ secretion is enhanced by high plasma K^+ concentrations.
- **d.** Distal convoluted tubule, which not only secretes K^+ during hyperkalemia, but also alters NaCl uptake and hence, distal NaCl delivery and plasma aldosterone, in response to changes in extracellular K^+ concentration, and thereby regulates K^+ secretion in the collecting duct.
- **e.** Collecting duct, where the rate of K^+ secretion is regulated, and K^+ reabsorption also occurs.

5. Which of the following are effects of aldosterone on Na^+ transport in the connecting segment and collecting duct?
- **a.** Enhances the permeability of Na^+ channels in the apical plasma membrane, thereby enhancing Na^+ reabsorption.
- **b.** Stimulates Na^+,K^+-ATPase activity in the basolateral plasma membrane, thereby enhancing Na^+ reabsorption.
- **c.** Reduces the Na^+ permeability of the apical plasma membrane, thereby inhibiting Na^+ reabsorption.
- **d.** Reduces Na^+,K^+-ATPase activity in the basolateral plasma membrane, thereby inhibiting Na^+ reabsorption.
- **e.** Increases the abundance of Na^+ channels in the apical plasma membrane, thereby enhancing Na^+ reabsorption.

Bibliography

Alexander RT, Dimke H. Effect of diuretics on renal tubular transport of calcium and magnesium. *Am J Physiol Renal Physiol*. 2017;312:F998–F1015.

Blaine J, Chonchol M, Levi M. Renal control of phosphate, calcium, and magnesium homeostasis. *Clin J Am Soc Nephrol*. 2015;10:1257–1272.

McDonough AA, Youn JH. Potassium homeostasis: the knowns, the unknowns, and the health benefits. *Physiology (Bethesda)*. 2017;32:100–111.

Muto S. Physiological roles of claudins in kidney tubule paracellular transport. *Am J Physiol Renal Physiol*. 2017;312:F9–F24.

Nielsen R, Christensen EI, Birn H. Megalin and cubilin in proximal tubule protein reabsorption: from experimental models to human disease. *Kidney Int*. 2016;89(1):58–67.

Skorecki K, Chertow GM, Marsden PA, et al, eds. *Brenner & Rector's the Kidney*. 10th ed. Philadelphia: Elsevier; 2016.

43

Water Balance

JILL W. VERLANDER

KEY POINTS

1. The kidney maintains water balance.
2. The proximal tubule reabsorbs more than 60% of filtered water.
3. The kidney can produce either concentrated or diluted urine.
4. A hypertonic medullary interstitium is needed to form concentrated urine.
5. Short-loop and long-loop nephrons both contribute to urine concentration.
6. Sodium chloride reabsorption by the medullary thick ascending limb generates medullary hypertonicity.
7. Urea reabsorption by the inner medullary collecting duct and urea recycling enhance medullary hypertonicity.
8. The countercurrent mechanism increases medullary interstitial osmolality with minimal energy expenditure.
9. Countercurrent exchange in the vasa recta removes water from the medullary interstitium without reducing medullary interstitial hypertonicity.
10. Active sodium chloride reabsorption in the thick ascending limb and the distal convoluted tubule dilutes the tubule fluid.
11. Antidiuretic hormone enhances urine concentration by increasing collecting duct water and urea permeability.
12. Cells in the inner medulla adapt to interstitial hyperosmolality by accumulation of organic osmolytes.

The Kidney Maintains Water Balance

One of the most important functions of the kidney is maintaining the water content of the body and the tonicity of the plasma. Terrestrial animals must constantly guard against desiccation, thus their kidneys evolved to reabsorb most of the water in the glomerular filtrate. Under normal conditions, a 10-kg beagle that produces 53.3 L of glomerular filtrate every day may reabsorb more than 99% of the water contained in the glomerular filtrate, excreting only 0.2 to 0.25 L of urine. A water-deprived dog with normal renal function can produce urine seven to eight times more concentrated than the osmolality of plasma, significantly higher than 2000 milliosmoles per kilogram of water (mOsm/kg H_2O). However, the kidney also can produce hypotonic urine in response to a water overload. After a water load, the same dog can excrete urine with an osmolality as low as 100 mOsm/kg H_2O, approximately one-third that of plasma. This chapter discusses how the kidney accomplishes these feats.

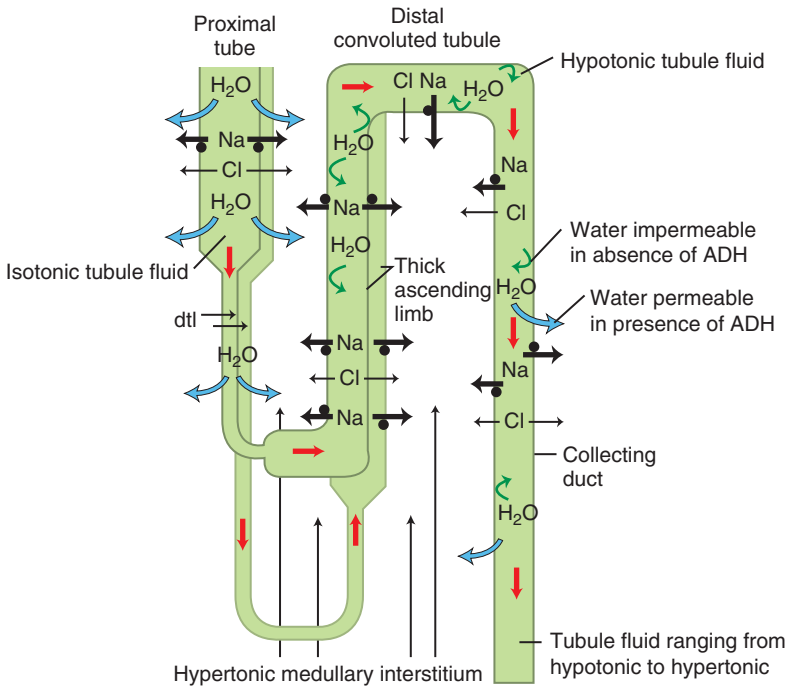
The Proximal Tubule Reabsorbs More Than 60% of Filtered Water

The proximal tubule reabsorbs the majority of the glomerular filtrate. It takes up solutes from the tubule fluid by both active and passive means. The sodium-potassium-adenosine triphosphatase (Na^+,K^+-ATPase) pump in the basolateral plasma membrane actively transports Na^+ and drives carrier-mediated secondary active transport and passive uptake of solutes. Removal of solute from the tubule fluid creates a slight gradient favoring the movement of water into the cells and the intercellular spaces. The complex apical brush border and basolateral plasma membrane infoldings create large surface areas that are highly permeable to water primarily because of the water channel, aquaporin-1 (AQP1), in both the apical and basolateral plasma membranes throughout the proximal tubule. Second, the tight junctions between proximal tubule cells contain water channels formed by claudin-2, which facilitate paracellular water movement. Thus the small chemical gradient results in rapid movement of water from the tubule fluid to the interstitial fluid. The high oncotic pressure and low hydrostatic pressure in the peritubular capillaries favor the movement of reabsorbed water and solute from the interstitial fluid to the blood.

The proximal tubules in the kidneys of our 10-kg beagle reabsorb 32 to 37 L of water per day. However, because proximal tubules reabsorb water nearly isotonically with salt, the osmolality of the tubule fluid remains similar from Bowman's space to the beginning of the thin descending limb of Henle's loop.

The Kidney Can Produce Either Concentrated or Diluted Urine

An elegant system has evolved in the mammalian kidney that allows excretion of either concentrated or diluted urine as needed. This system has three main components: (1) generation of a hypertonic medullary interstitium, which allows excretion of concentrated urine; (2) dilution of the tubule fluid by the thick ascending limb and the distal convoluted tubule, which allows excretion of dilute urine; and (3) variability in the water permeability of the collecting duct in response to antidiuretic hormone (ADH, vasopressin), which determines the final urine concentration. The beauty of this system is that all the factors necessary for urine concentration and dilution are operative at any given time, so the kidney can respond immediately to changes in ADH levels with corresponding changes in urine osmolality and water excretion. The basic elements of this system are illustrated in Fig. 43.1.



• **Fig. 43.1** Overview of renal mechanisms maintaining water balance. The proximal tubule reabsorbs approximately 60% of filtered water isotonically with solute. The thick ascending limb generates medullary hypertonicity by vigorous active reabsorption of NaCl without water and dilutes the tubule fluid. The distal convoluted tubule also reabsorbs NaCl without water and further dilutes the tubule fluid. Depending on the presence or absence of antidiuretic hormone (ADH, vasopressin), the collecting duct is either water permeable or impermeable. When ADH is present, the collecting duct is water permeable and the osmotic gradient produced by the hypertonic medullary interstitium causes water reabsorption and formation of concentrated urine. When ADH is absent, the collecting duct is water impermeable, and a dilute urine is excreted.

A Hypertonic Medullary Interstitium Is Needed to Form Concentrated Urine

Terrestrial animals usually produce urine concentrated well above the plasma osmolality. Excretion of concentrated wastes conserves water and thus reduces the volume of water that must be consumed to prevent dehydration. Two of the three factors just mentioned are responsible for the formation of concentrated urine: (1) generation of a hypertonic medullary interstitium and (2) enhanced water permeability in the collecting duct in the presence of ADH.

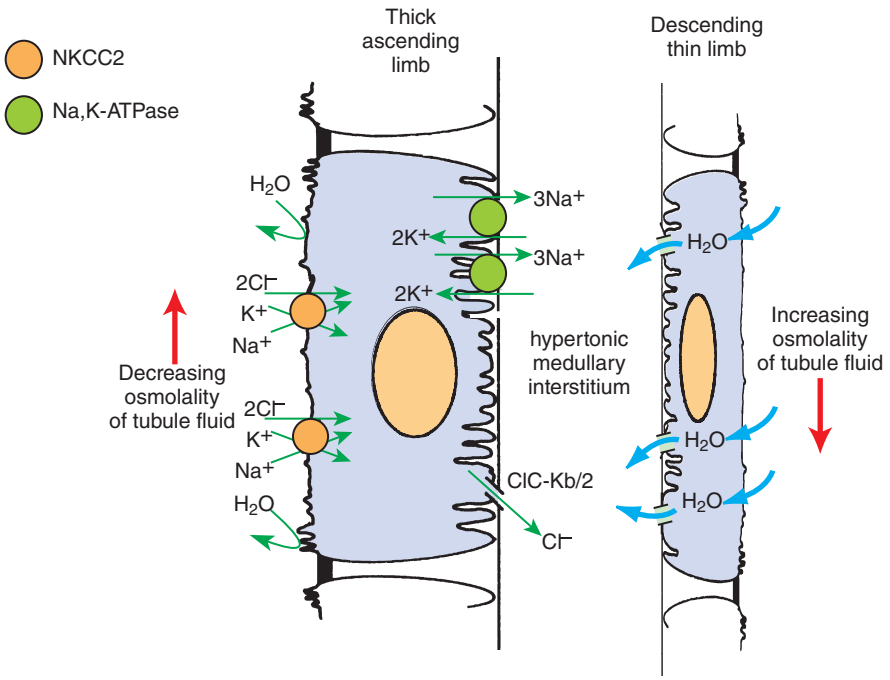
The hypertonicity of the medullary interstitium is produced and maintained primarily by (1) the reabsorption of osmotically active substances by tubules in the medulla and (2) the removal of water from the medullary interstitium by the vasa recta.

Short-Loop and Long-Loop Nephrons Both Contribute to Urine Concentration

The anatomical arrangement of the renal tubules in the medulla is a crucial element of the urine-concentrating mechanism. The nephrons of the mammalian kidney are classified as superficial and juxtamedullary nephrons based on the location of their respective glomeruli (see Fig. 41.1). *Superficial nephrons* have short loops of Henle that extend only into the inner stripe of the outer medulla. These *short-loop nephrons* have a descending thin limb that parallels the thick ascending limb, but they do not have an ascending thin limb; the thin descending limb merges with the thick ascending limb near the hairpin turn (Fig. 43.1). *Juxtamedullary, or long-loop, nephrons* have long loops of Henle that extend deep into the inner medulla. Midcortical nephrons, which have glomeruli located near the midregion of the cortex, may have long or short loops.

Both short-loop and long-loop nephrons contribute to urine-concentrating ability by increasing medullary tonicity via NaCl uptake in the thick ascending limb. However, the contribution of the two types of nephrons varies due to differences in the thin limbs; whereas the short-loop nephrons have only descending thin limbs and do not extend past the outer medulla, long-loop nephrons have several segments of descending and ascending thin limbs with specific urea and water transporter expression that adds to their role in maintaining the medullary hypertonicity and urine-concentrating ability.

In most mammals, short-loop nephrons outnumber long-loop nephrons, but the relative abundance of superficial and juxtamedullary nephrons varies among mammalian species. For example, humans generally have approximately 85% short-loop nephrons, mice approximately 75%, and pigs approximately 97%. Rabbits have a majority of long-loop nephrons, with only approximately 34% short-loop nephrons. Carnivores, including dogs, cats, and foxes, reportedly have only long-loop nephrons; however, the loops extend to a variable depth in the inner medulla. Some mammals that live primarily in fresh water, such as the mountain beaver and hippopotamus, do not have an inner medulla or long-loop nephrons. Despite the paucity of long-loop nephrons in many mammals, all of these animals are able to concentrate their urine, up to severalfold



• **Fig. 43.2** In the outer medulla, the thick ascending limb of both long- and short-loop nephrons actively transports sodium into the medullary interstitium via basolateral Na,K-ATPase, which drives apical NaCl uptake via the Na, K, 2 Cl cotransporter, NKCC2. Because the thick ascending limb is water impermeable, this increases the tonicity of the medullary interstitium. Descending thin limbs run parallel, with flow in the opposite direction, and are water permeable. Thus the osmolality of the luminal fluid in descending thin limbs progressively increases due to water abstraction to equilibrate with the ambient interstitial osmolality.

higher than plasma osmolality. The length of Henle's loop in long-loop nephrons varies as well. In general, animals with the greatest ability to concentrate the urine have longer renal papillae and longer loops of Henle than other animals. Beyond these simple features, the complex relationships among Henle's loop, vasa recta, collecting ducts, and other structures in the inner and outer medulla, as well as the specific segmental transport properties and permeabilities, all contribute to the efficiency of the urine-concentrating mechanism.

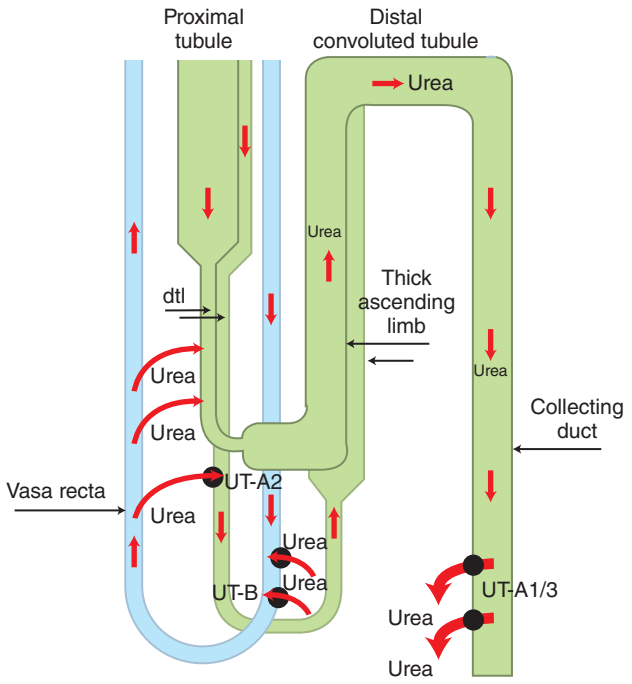
In birds, the reptilian-type nephrons have glomeruli near the surface in the renal cortex and have no loops of Henle. The mammalian-type nephrons have glomeruli that are deeper in the cortex and have either short or long loops of Henle that extend into the medullary cone. The mammalian-type nephrons have a countercurrent arrangement and are thought to be largely responsible for the ability of birds to excrete hypertonic urine. The kidneys of animals with no loops of Henle, such as reptiles, are unable to concentrate the urine above the osmolality of plasma.

Sodium Chloride Reabsorption by the Medullary Thick Ascending Limb Generates Medullary Hypertonicity

The thick ascending limb of Henle's loop actively reabsorbs sodium chloride (NaCl) but is impermeable to water. Therefore this segment raises the osmolality of the interstitial fluid, thus generating medullary interstitial hypertonicity and a lumen-to-interstitium osmotic gradient. This process occurs in both short-loop and long-loop nephrons. The hypertonic interstitium allows water to be abstracted from the descending thin limbs, which are water-permeable due to the presence of the water channel, AQP1, thus concentrating the tubule fluid in the descending thin limbs. Because of the parallel, yet countercurrent arrangement of the thick ascending limb and the descending thin limb, this results in progressively increasing osmolality in the tubule fluid of the descending thin limb. This process is illustrated in Fig. 43.2.

Urea Reabsorption by the Inner Medullary Collecting Duct and Urea Recycling Enhance Medullary Hypertonicity

The *inner medullary collecting duct* (IMCD) also actively reabsorbs NaCl, but its more important contribution to the medullary hypertonicity is the reabsorption of urea (see Fig. 43.3). Although the thick ascending limb, distal convoluted tubule, and cortical and outer medullary collecting ducts are impermeable to urea, the terminal IMCD is highly permeable to urea via specific urea transporters (UT-A1, UT-A3). Thus from the thick ascending limb distally, urea remains in the tubule fluid until it reaches the terminal IMCD deep in the medulla. Because urea reabsorption by the IMCD is enhanced by ADH, when conditions demand water conservation and ADH is released, urea reabsorption is enhanced and the osmotic gradient for water uptake increases. Because the thin limbs of Henle's loop are permeable to urea, the high interstitial urea concentration drives urea into the thin limb luminal fluid. Because the tubule segments that intervene between the ascending thin limb and the terminal IMCD are impermeable to urea, urea reabsorbed from the terminal IMCD and taken up by the thin

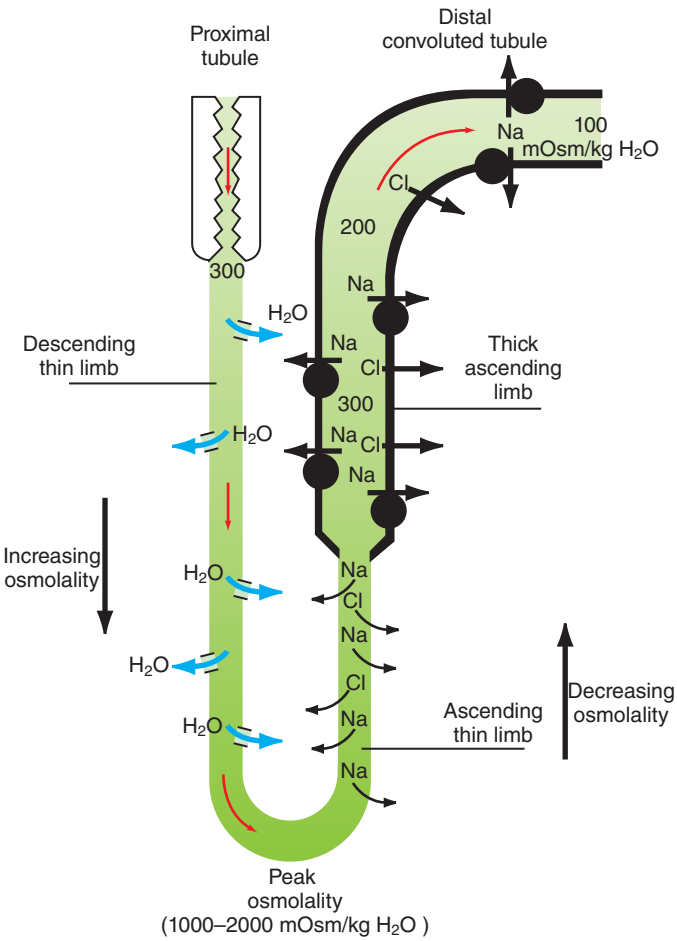


• **Fig. 43.3** Urea recycling in the kidney. Filtered urea is reabsorbed in the IMCD by carrier-mediated, facilitated transport (UT-A1, UT-A3) and diffuses down its concentration gradient into the vasa recta. Urea diffuses out of the fenestrated ascending vasa recta down its concentration gradient and returns to the tubule lumen by transport into the thin limbs of Henle's loop. Urea uptake in the descending thin limbs (*dtl*) and descending vasa recta is enhanced by the urea transporters, UT-A2 and UT-B, respectively. Because the thick ascending limb, distal convoluted tubule, and cortical and outer medullary collecting ducts are impermeable to urea, urea in the thin limb luminal fluid is recycled to the IMCD. Urea reabsorption in the IMCD is enhanced by ADH. The result is enhanced accumulation of urea in the medullary interstitium, which contributes to the medullary interstitial osmotic pressure and promotes water reabsorption. *ADH,* Antidiuretic hormone; *IMCD,* inner medullary collecting duct.

limbs is recycled back to the IMCD. In mammals, this system of *urea recycling* enhances the efficiency of the urine-concentrating mechanism. In birds, however, urea is nearly absent in the medullary interstitium; urates do not contribute appreciably to osmotic pressure because they have low water solubility. Thus medullary hypertonicity in birds appears to depend on single-solute (NaCl) recycling.

The Countercurrent Mechanism Increases Medullary Interstitial Osmolality With Minimal Energy Expenditure

The prevailing hypothesis for decades has been that a countercurrent mechanism in the thin limbs of Henle's loop in the inner medulla is responsible for the progressive amplification of the medullary hypertonicity initiated by the active reabsorption of salt by the thick ascending limb (Fig. 43.4). This may be accomplished with minimal energy expenditure because of two characteristics: (1) the anatomical arrangement of the thin limbs of Henle's loop and (2) the differential water and salt permeabilities of the descending and ascending thin limbs. Although recent data on the specific distribution of water and solute permeabilities and the complex anatomical associations of tubules and vessels in the medulla have raised some doubts about the countercurrent multiplier hypothesis, the fundamentals of the concept remain the basis of our understanding of the mechanisms maintaining medullary hypertonicity.



• **Fig. 43.4** Generation of medullary hypertonicity and a dilute tubule fluid in the distal nephron and preservation of medullary hypertonicity by the thin limbs of Henle's loop. The thick ascending limb of Henle's loop actively transports NaCl into the interstitium without water, thus diluting the tubule fluid and raising the medullary interstitial tonicity. The proximal tubule luminal fluid osmolality is approximately 300 mOsm/kg H_2O when it enters the progressively more concentrated medullary interstitium. Because portions of the descending thin limbs are impermeable to sodium (*Na*) but are permeable to water (*H_2O*), at least in part due to AQP1, water diffuses into the interstitium and the tubule fluid concentrates. After the hairpin turn deep in the inner medulla, the concentrated tubule fluid crosses regions of progressively lower interstitial osmolality as it flows through the ascending thin limb of Henle's loop. Because this segment is impermeable to water but is permeable to sodium, the gradient draws luminal sodium into the interstitium. The differential permeabilities of the descending and ascending thin limbs and the countercurrent arrangement preserve the medullary interstitial concentration gradient. The distal convoluted tubule continues NaCl reabsorption without water, so that the osmolality of the tubule fluid delivered to the collecting duct system is approximately 100 mOsm/kg H_2O, much less than that of plasma (295–300 mOsm/kg H_2O). All numerical values are in mOsm/kg H_2O.

The thin limbs of Henle's loop in juxtamedullary nephrons extend deep into the inner medulla. The descending and ascending thin limbs are joined by a sharp, hairpin turn. Thus the descending and ascending thin limbs run parallel to each other, with the tubule fluid flow in opposite directions.

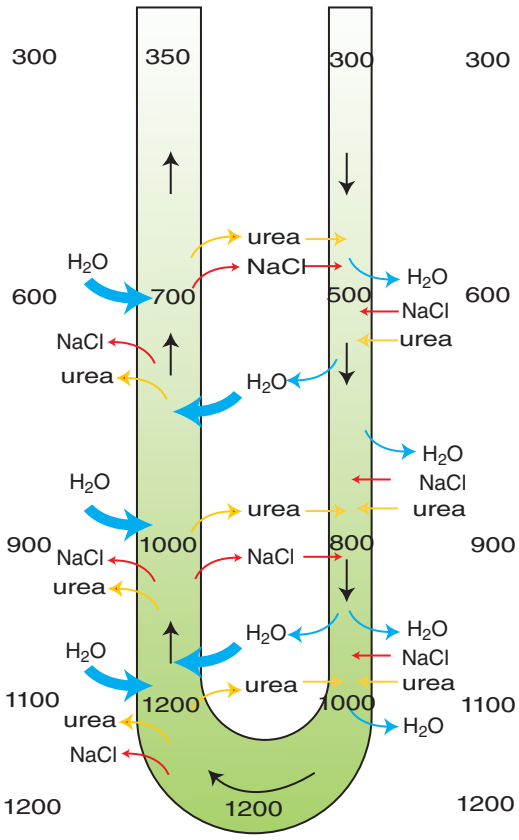
As mentioned before, the descending thin limb originates from the straight portion of the proximal tubule and runs parallel to the

thick ascending limb of Henle's loop. The tubule fluid entering the thin limb is essentially isosmotic to plasma, whereas the surrounding interstitial fluid in the outer medulla is hyperosmotic because of NaCl reabsorption by the water-impermeable thick ascending limb. Descending thin limbs contain AQP1 water channels and are highly permeable to water. However, the descending thin limb, at least in the outer medulla, is not permeable to salt. Thus the tubule fluid equilibrates with the hypertonic interstitial fluid by the movement of water into the interstitium and the tubule fluid osmolality rises.

The osmolality of the medullary interstitial fluid is progressively higher in the deeper regions of the medulla, and the tubule fluid osmolality also progressively rises until it reaches its maximal concentration at the hairpin turn. Formerly, this was believed to result from continued equilibration by diffusion of water without salt into the interstitium. However, more recent evidence shows that descending thin limbs in the inner medulla are salt permeable; with this knowledge, the precise explanation for how the tubule fluid osmolality continues to increase to the hairpin turn is in question.

As the thin limb ascends through regions of progressively lower interstitial osmolality, the concentrated luminal fluid flows through lower ambient osmolalities, and once again the tubule fluid equilibrates with the interstitial fluid. However, the ascending thin limb is impermeable to water and permeable to NaCl, so the equilibration occurs not by movement of water into the tubule fluid but by diffusion of NaCl from the tubule fluid into the interstitial fluid. Although the mechanisms responsible for the Na^+ permeability are not yet established, a specific Cl^- channel, ClC-Ka/1, mediates Cl^- permeability in ascending thin limbs. Due to NaCl diffusion out of the concentrated luminal fluid, the tubule fluid osmolality decreases and solute is added to the interstitium, contributing to the interstitial hypertonicity. This process continues until the ascending thin limb merges with the thick ascending limb in the outer medulla. At the transition to the thick ascending limb, the tubule fluid osmolality is only moderately hypertonic.

At this point, what has been accomplished? By passive means, the thin limbs have reabsorbed both water and salt. Water was reabsorbed from the descending thin limb, and salt was reabsorbed from the ascending thin limb. At the same time, the countercurrent flow and variable water and salt permeabilities in these two segments have helped to maintain the medullary hypertonicity.



• **Fig. 43.5** Countercurrent exchange in the vasa recta. The walls of the vasa recta are permeable to water, salt (*NaCl*), and urea, and the descending and ascending vasa recta are parallel and closely associated, with flows in opposite directions. Plasma entering the medulla in the descending vasa recta has an osmolality of approximately 300 mOsm/kg H_2O, and progressively increases by equilibration with the progressively higher interstitial osmolality of the inner medulla. Similarly, the plasma osmolality decreases as the vessel passes through regions of progressively lower interstitial osmolality. In both arms of the vasa recta, the gradient between the plasma and interstitial osmolality is reduced by movement of water and solute in opposite directions. This system prevents the dissipation of the medullary concentration gradient. In addition, there is net removal of water from the interstitium because of the relatively low hydrostatic pressure and relatively high oncotic pressure in the vasa recta. Increased color intensity represents increased solute content in the vasa recta lumen.

Countercurrent Exchange in the Vasa Recta Removes Water From the Medullary Interstitium Without Reducing Medullary Interstitial Hypertonicity

The diffusion of water from the descending thin limb into the interstitium would dilute the effect of salt and urea transport into the interstitium and cause swelling of the inner medulla if it were not for the ability of the vasa recta to remove the reabsorbed fluid. The vasa recta are permeable to water, salts, and urea. The relatively high plasma oncotic pressure in the vasa recta entering the medulla favors the movement of water into the capillary lumen, and luminal NaCl and urea concentrations equilibrate with the interstitial concentration. Thus as the vessels descend in the inner medulla, the plasma osmolality and urea concentration rise in the vasa recta near the hairpin turn, then fall as the vessels ascend out of the medulla (Fig. 43.5). What is the net effect of passive equilibration of the interstitial fluid with the plasma in the vasa recta? Two observations indicate that, by the time the ascending vasa recta leave the medulla, there has been net movement of fluid into the capillary: (1) the plasma oncotic pressure has fallen, and (2) the blood flow in the ascending vasa recta is approximately double that in the descending vasa recta. Thus the relatively high initial plasma oncotic pressure, the countercurrent arrangement of the vasa recta, and the passive equilibration of the plasma with the changing interstitial osmolalities in different regions of the medulla allow the removal of both water and solute from the medullary interstitium without dissipating the medullary hypertonicity.

Active Sodium Chloride Reabsorption in the Thick Ascending Limb and the Distal Convoluted Tubule Dilutes the Tubule Fluid

The thick ascending limb and the distal convoluted tubule actively reabsorb Na^+, which drives Cl^- reabsorption, via mechanisms

described in detail in Chapter 42. Because these segments are essentially impermeable to water, active solute reabsorption causes a progressive decline in the tubule fluid osmolality. Thus the thick ascending limb and distal convoluted tubule are often called the *diluting segments.* The result is that the tubule fluid delivered to the collecting duct is hypotonic, even in a dehydrated animal.

Antidiuretic Hormone Enhances Urine Concentration by Increasing Collecting Duct Water and Urea Permeability

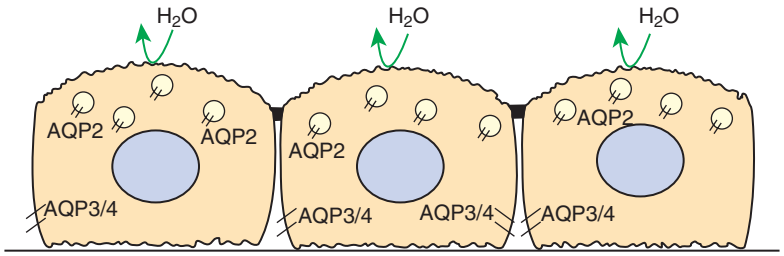
The generation of medullary hypertonicity and dilution of the tubule fluid in the distal nephron segments set the stage for the elimination of either concentrated or dilute urine, as warranted by the fluid volume status, plasma tonicity, and blood pressure of the animal. The water permeability of the collecting duct, which is regulated by ADH (arginine vasotocin in birds), determines the osmolality of the excreted urine.

During water overload, ADH is absent, and the collecting duct is relatively impermeable to water. The tubule fluid delivered by the distal convoluted tubule is hypotonic and remains so, because the water remains in the collecting duct lumen. Thus in the absence of ADH, dilute urine is formed, and excess water is excreted (Fig. 43.6).

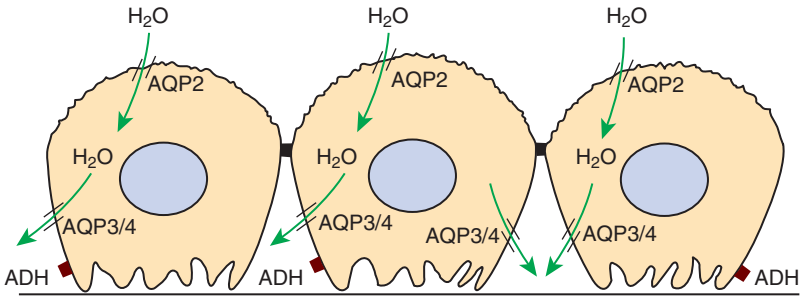
During dehydration, hypotension, or volume depletion, ADH is released from the pituitary. ADH release is triggered by a rise in plasma osmolality of as little as 3 to 5 mOsm/kg H_2O resulting from dehydration or salt overload and by decreased blood pressure as a result of systemic vasodilation, heart failure, or isosmotic volume depletion from vomiting, diarrhea, or hemorrhage. In these circumstances, the animal needs to reduce the plasma osmolality to normal or restore fluid volume or blood pressure.

When ADH is present, water flows from the dilute tubule fluid into collecting duct cells and into the hypertonic medullary interstitium down the concentration gradient, producing structural alterations that include cell swelling and dilation of the intercellular spaces (Fig. 43.7). As the now water-permeable collecting duct traverses the inner medulla through regions of progressively higher interstitial fluid osmolality, the tubule fluid equilibrates by diffusion of water into the interstitium, and highly concentrated urine is excreted.

ADH acutely regulates the water permeability of the collecting duct by regulating the location and abundance of the water-channel protein *aquaporin-2* (AQP2) in collecting duct cells. When ADH is absent, AQP2 is contained in cytoplasmic vesicles in principal cells and IMCD cells (see Fig. 43.6). ADH secretion activates the vasopressin V2 receptor present in principal cells and IMCD cells, which acutely stimulates the insertion of AQP2 into the apical plasma membrane of these cells, and water freely passes through these water channels (see Fig. 43.7). Chronic ADH stimulation leads to an overall increase in the amount of AQP2 in the collecting duct; conversely, chronically low ADH leads to decreased AQP2 expression. As already mentioned, ADH also enhances urea reabsorption by the IMCD, by increasing phosphorylation and apical expression of the urea transporters, UT-A1 and UT-A3, enabling increased urea accumulation in the medullary interstitium. The basolateral water channels, aquaporin-3 and aquaporin-4, are present in the basolateral plasma membrane and allow the movement of water from inside the cell to the hypertonic interstitial space. Forms of the clinical condition known as *nephrogenic diabetes insipidus,* which is characterized by renal unresponsiveness to ADH, are



• **Fig. 43.6** Collecting duct epithelium in the absence of antidiuretic hormone (ADH). When ADH is absent, the apical plasma membrane is impermeable to water, and dilute urine is excreted.



• **Fig. 43.7** Water permeability of apical plasma membrane of collecting duct epithelium in the presence of antidiuretic hormone (ADH). ADH binds to the type-2 vasopressin receptor (V2R; *red squares*) in principal cells, thereby stimulating the insertion of aquaporin-2 (*AQP2*) water channels into the apical plasma membrane and enhancing its water permeability. Water rushes into the cells, and across the basolateral plasma membrane via aquaporin-3 and -4 (*AQP3, AQP4*) into the lateral intercellular spaces. Concurrent morphological changes include translocation of membrane containing AQP2 from intracytoplasmic vesicles to the apical plasma membrane, cell swelling into the tubule lumen, and dilation of the lateral intercellular spaces.

caused by genetic defects in the V2 receptor or by congenital or acquired abnormalities or deficiency of AQP2 proteins, often accompanied by decreased expression of other aquaporins.

ADH effects on NaCl transport proximal to the collecting duct also contribute to conservation of fluid volume. ADH stimulates NaCl reabsorption in the thick ascending limb and the distal convoluted tubule. In the thick ascending limb, increased NaCl reabsorption increases the medullary interstitial hypertonicity and dilutes the tubule fluid. In the distal convoluted tubule, increased NaCl uptake returns salt to the systemic circulation, which leads to enhanced intravascular volume, and dilutes the tubule fluid further. These effects on NaCl reabsorption likely contribute to increased water reabsorption in the collecting duct by increasing the concentration gradient between the collecting duct luminal fluid and the medullary interstitium. ADH also increases the abundance and apical localization of the epithelial sodium channel, ENaC, in principal cells of the cortical collecting duct, thereby enhancing sodium uptake in this segment.

In birds, AQP2 homologs are present in collecting ducts of mammalian-type nephrons and are stimulated by arginine vasotocin to enhance water permeability; an avian AQP4 homolog has been identified in the basolateral plasma membrane of collecting duct cells. Salt and water reabsorption also occurs distal to the collecting ducts. Birds lack a urinary bladder; urine travels from the kidneys via the ureters to the cloaca, where both salt and water are reabsorbed. Furthermore, cloacal urine passes retrograde into the digestive tract, where additional salt and water are reabsorbed. The importance of this process with respect to both salt and water recovery varies among bird species.

Cells in the Inner Medulla Adapt to Interstitial Hyperosmolality by Accumulation of Organic Osmolytes

Cells in the inner medulla not only exist in a hypertonic environment but also regulate cell volume during changes in ambient osmolality. These cells accomplish this by accumulating organic *osmolytes* that maintain the intracellular osmotic pressure and prevent cell shrinkage without marked increases in the concentration of intracellular electrolytes. These substances include sorbitol, betaine, myoinositol, amino acids, and glycerophosphorylcholine (GPC). The intracellular concentrations of these osmolytes vary with the diuretic state of the animal, increasing during periods of urine concentration, when the medullary interstitial osmolality is maximized, and decreasing during diuresis, when medullary interstitial osmolality decreases. Changes in the intracellular content of organic osmolytes in response to changes in ambient osmolality occur by parallel changes in either the production (e.g., sorbitol, GPC) or transmembrane transport (e.g., betaine, myoinositol, amino acids) of the osmolytes, or by counterregulation of degradation of osmolytes (e.g., GPC).

CLINICAL CORRELATIONS

Diabetes Insipidus

History

A client presents her 6-month-old female Boston terrier with a complaint of excessive water consumption and urination.

Clinical Examination

The physical examination reveals no abnormalities. The dog is alert and active. A urinalysis is normal, and the urine specific gravity is 1.002 (osmolality, 152 mOsm/kg H_2O). The serum chemistry profile and a complete blood cell count (CBC) are normal.

You admit the dog to your clinic for a modified water deprivation test. The urine fails to become concentrated despite a 5% loss of body weight. You administer vasopressin (ADH), and the urine specific gravity is 1.029 (osmolality, 852 mOsm/kg H_2O) 1 hour later.

Comment

The dog has *central diabetes insipidus* (DI), which is a deficiency of ADH. The thick ascending limb of Henle's loop and the distal convoluted tubule dilutes the urine. Solute-free water absorption in the collecting duct depends on the action of ADH. In the absence of ADH, excessive water is excreted, and the dog drinks voraciously to prevent dehydration.

Other causes of excretion of dilute urine (urine osmolality much lower than serum osmolality) are psychogenic polydipsia, hyperadrenocorticism, glucocorticoid therapy, hypercalcemia, hypokalemia, and nephrogenic DI. Most of these can be identified from a thorough history, physical examination, CBC, and serum chemistry profile. When only psychogenic polydipsia, central DI, and nephrogenic DI remain in the differential diagnosis, the diagnosis usually can be made using the modified water deprivation test. Animals with psychogenic polydipsia can secrete ADH and have normal kidneys; therefore they concentrate their urine after water deprivation. Dogs with DI can concentrate their urine minimally or not at all after water deprivation. If the problem is insufficient ADH release (central DI), the urine concentration increases in response to exogenous ADH. If the kidney is unresponsive to ADH (nephrogenic DI), the urine concentration does not increase further in response to additional ADH.

Treatment

Treatment of central DI includes free access to water and daily administration of a vasopressin analog, such as desmopressin (DDAVP), by parenteral or intranasal route.

Chronic Renal Insufficiency

History

You recommend routine dentistry for a 15-year-old male miniature schnauzer that appears to be in good health. Before administering general anesthetic to this aged animal, you obtain a CBC, serum chemistry profile, and urinalysis to detect any subclinical organ dysfunction.

Clinical Examination

The CBC and serum chemistry profile are normal. The urinalysis is normal with a specific gravity of 1.010 (osmolality, 352 mOsm/kg H_2O). You ask the owner to submit a sample of the dog's urine from the first void of the day. The specific gravity of this sample is 1.012 (osmolality, 401 mOsm/kg H_2O).

Comment

Chronic renal insufficiency is common in geriatric patients and is probably responsible for the two urine specific gravity values in the "fixed range" of 1.008 to 1.012. These values correspond to osmolalities similar to or slightly higher than normal plasma osmolality. Additional evaluation would have to be performed to verify that this animal could neither dilute nor concentrate his urine significantly; however, in an animal of advanced age and in the absence of other clinical or laboratory abnormalities, further evaluation is probably not indicated.

In chronic renal insufficiency, the loss of functional nephrons is first manifested by the inability to significantly alter the urine concentration in response to a water load or water deprivation. The residual nephrons are initially able to sustain adequate filtration rates to prevent azotemia (elevated

Continued

CLINICAL CORRELATIONS—cont'd

serum creatinine and urea nitrogen levels), but the compensatory increase in flow rates in individual nephrons probably exceeds the capacity of the thick ascending limb and distal convoluted tubule to dilute the tubule fluid significantly or to generate a steep medullary concentration gradient. In addition, the loss of nephrons likely alters the medullary architecture such that the countercurrent mechanism and maintenance of the medullary hypertonicity is further compromised. As a result of these changes, the tubule fluid cannot be concentrated greatly above or diluted below the level of the plasma osmolality. If there is progressive nephron loss, the glomerular filtration rate will continue to decline, and renal failure will ensue.

Treatment

It is important to be aware that your patient has chronic renal insufficiency and is unable to respond efficiently to changes in fluid and salt intake. Water should be withheld only briefly, and caution should be exercised in supporting the animal with intravenous fluids during anesthesia while avoiding fluid overload.

A highly bioavailable, low-protein diet that is also low in sodium and phosphorus may slow the progression of chronic renal disease and delay the onset of renal failure, at least in some species.

Practice Questions

1. The bulk of filtered water is reabsorbed by which renal tubule segment?
- **a.** Proximal tubule
- **b.** Thin limbs of Henle's loop
- **c.** Thick ascending limb of Henle's loop
- **d.** Cortical collecting duct
- **e.** Inner medullary collecting duct

2. The kidney responds rapidly to changing water requirements. The ability to quickly alter the rate of water excretion by greatly concentrating or diluting the urine is the result of several factors. Which of the following does not contribute to this ability?
- **a.** Generation of hypertonic medullary interstitium
- **b.** Countercurrent flow and differential salt and water permeabilities in the thin limbs of Henle's loop
- **c.** Dilution of the tubule fluid by the thick ascending limb and the distal convoluted tubule
- **d.** Responsiveness of the collecting duct to antidiuretic hormone (ADH)
- **e.** ADH-regulated countercurrent flow and enhanced water permeability in the vasa recta

3. The hypertonic medullary interstitium is generated in large part by:
- **a.** Active transport of Na^+ by the straight portion of the proximal tubule.
- **b.** Active reabsorption of Na^+ by the water-impermeable, ascending thin limb of Henle's loop.
- **c.** Active reabsorption of Na^+ by the water-impermeable, thick ascending limb of Henle's loop.
- **d.** Increase in water channels in the apical plasma membrane of collecting duct cells under the influence of ADH.
- **e.** Enhanced urea permeability of the thick ascending limb of Henle's loop under the influence of ADH.

4. In dehydration, ADH is released, which reduces water excretion by:
- **a.** Enhancing water reabsorption in the proximal tubules by stimulating Na^+,K^+-ATPase.
- **b.** Enhancing water reabsorption in the thick ascending limb by stimulating the insertion of aquaporin-2 water channels into the apical plasma membrane.
- **c.** Enhancing water reabsorption in the collecting duct by stimulating Na^+,K^+-ATPase activity.
- **d.** Enhancing water permeability in the collecting duct by stimulating the insertion of aquaporin-2 water channels into the apical plasma membrane.
- **e.** Reducing the glomerular filtration rate by activation of tubuloglomerular feedback.

5. In clinical situations, the excretion of dilute urine may be caused by all the following *except:*
- **a.** Chronic renal disease
- **b.** Glucocorticoid administration
- **c.** ADH deficiency
- **d.** Hypercalcemia
- **e.** Acute renal hypoperfusion

Bibliography

Knepper MA, Kwon T, Nielsen S. Molecular physiology of water balance. *N Engl J Med*. 2015;372(14):1349–1358.

Laverty G, Skadhauge E. Adaptive strategies for post-renal handling of urine in birds. *Comp Biochem Physiol A Mol Integr Physiol*. 2008;149(3):246–254.

Nishimura H. Urine concentration and avian aquaporin water channels. *Pflugers Arch*. 2008;456(4):755–768.

Pallone TL, Edwards A, Mattson DL. Renal medullary microcirculation. *Compr Physiol*. 2012;2(1):97–140.

Pannabecker TL. Comparative physiology and architecture associated with the mammalian urine concentrating mechanism: role of inner medullary water and urea transport pathways in the rodent medulla. *Am J Physiol Regul Integr Comp Physiol*. 2013;304:R488–R503.

Rosenthal R, Gunzel D, Theune D, et al. Water channels and barriers formed by claudins. *Ann N Y Acad Sci*. 2017;1397:100–109.

Sands JM, Layton HE. Advances in understanding the urine-concentrating mechanism. *Annu Rev Physiol*. 2014;76:387–409.

Sands JM, Layton HE, Fenton RA, et al. Urine concentration and dilution. In: Skorecki K, Chertow GM, Marsden PA, eds. *Brenner & Rector's the Kidney*. 10th ed. Philadelphia: Elsevier, Inc; 2016.

Yang B, Sands JM, eds. Urea transporters. In: *Subcell Biochem*. Vol 73. Springer Nature; 2014.

44

Acid-Base Balance

JILL W. VERLANDER

KEY POINTS

1. Buffers, lungs, and kidneys together maintain acid-base balance.
2. The kidneys reabsorb essentially all filtered bicarbonate.
3. The kidneys excrete excess acid and generate "new" bicarbonate in the process of net acid excretion.
4. Renal net acid excretion is the sum of urinary titratable acid plus urinary ammonium, minus urinary bicarbonate.
5. Urinary titratable acid forms from protons secreted by tubule epithelial cells that bind to filtered buffers in the tubule fluid and contributes to new bicarbonate generation.
6. Renal ammonia metabolism generates the largest component of net acid excretion and new bicarbonate production.
7. The proximal tubule has a high capacity for H^+ secretion, bicarbonate reabsorption, and ammonia metabolism.
8. The thick ascending limb of Henle's loop absorbs filtered bicarbonate.
9. The collecting duct determines the final urine pH.
10. The collecting duct can secrete protons and ammonia, reabsorb bicarbonate, and generate acidic urine.
11. The collecting duct can secrete bicarbonate and generate alkaline urine.

Buffers, Lungs, and Kidneys Together Maintain Acid-Base Balance

Normal blood pH is approximately 7.4; normal cellular function requires a pH close to this value. Three systems maintain acid-base homeostasis: (1) intracellular and extracellular buffers, (2) the lungs, and (3) the kidneys. The first two make rapid corrections of blood pH, whereas the kidneys more slowly control acid-base homeostasis, generate new bicarbonate, and excrete excess hydrogen ion (H^+).

Maintaining acid-base balance usually requires preventing excess acid in the body, which constantly produces acid as a byproduct of metabolism. The amount of acid produced varies depending on changes in diet, exercise, other organ functions, and in birds the phases of the egg-laying cycle. Therefore the systems that maintain acid-base homeostasis must adapt to changes in the acid load. Less often, excess base must be eliminated.

Several intracellular and extracellular buffers titrate H^+ to maintain a physiological pH. These include hemoglobin and other proteins, carbonate in bone, phosphate, and bicarbonate (HCO_3^-). These buffers rapidly normalize the pH after acute changes in the acid load, unless the buffering capacity is exceeded. In addition, during chronic metabolic acidosis, bone provides a reservoir of buffer mobilized to help normalize systemic pH. Excess H^+ and low HCO_3^- in the extracellular fluid promote physicochemical as well as osteoclast-mediated dissolution of bone, releasing carbonate, which buffers H^+. In chronic acidosis, this can lead to abnormally low bone mineral density.

The respiratory system also responds rapidly to maintain normal blood pH by altering the rate of removal of carbon dioxide (CO_2) from the blood. The enzyme carbonic anhydrase (CA), present in red blood cells and many other cells, catalyzes the first step in the following reaction, whereas the second step occurs instantaneously:

$$CO_2 + H_2O \overset{CA}{\leftrightarrow} H_2CO_3 \leftrightarrow H^+ + HCO_3^-$$

Removal of CO_2 from the blood by respiration shifts this reaction to the left, and the concentration of H^+ is consequently reduced (pH is raised). Thus the lung is important for stabilizing blood pH, particularly in response to rapid changes in the acid load, but it does not eliminate acid from the body.

The kidney is the third line of defense of acid-base balance. Although buffering and respiration are able to stabilize blood pH, the kidneys are responsible for the actual excretion of excess H^+ and maintaining normal plasma HCO_3^- levels. The kidney reabsorbs filtered bicarbonate and excretes net acid efficiently by the combined effects of (1) carbonic anhydrases, which make protons and bicarbonate readily available for transport; (2) transporters that move H^+ from the epithelial cells into the tubule fluid and bicarbonate into the interstitial fluid; (3) buffers that "trap" protons and minimize increases in H^+ concentration in the tubule fluid; and (4) regulated renal ammonia metabolism and secretion, which accounts for the majority of the renal response to regulate acid excretion.

The Kidneys Reabsorb Essentially All Filtered Bicarbonate

As described in Chapter 42, the proximal tubule reabsorbs the majority of the filtered HCO_3^-. Plasma HCO_3^- is freely filtered, and essentially all of the filtered bicarbonate is reabsorbed by the renal tubule epithelial cells. The majority of filtered bicarbonate is reabsorbed by the end of the proximal tubule. Bicarbonate in the tubule fluid is not reabsorbed by transepithelial transport of HCO_3^- itself but involves a more complex mechanism. H^+ secreted across the apical plasma membrane of proximal tubule cells combines

with luminal HCO_3^- to form carbonic acid (H_2CO_3), which is converted to H_2O and CO_2; the latter reaction is catalyzed by carbonic anhydrase associated with the apical plasma membrane. The CO_2 crosses the plasma membrane into the cell, at least partly by diffusion, but also may be facilitated by the water channel, aquaporin-1, which has been shown to function as a CO_2 channel in proximal tubule cells in vitro. Intracellular carbonic anhydrase catalyzes CO_2 hydration to form H^+ and HCO_3^-, which are transported across the apical and basolateral plasma membranes respectively, via membrane transporters discussed in a later section. By this multistep process, illustrated in Fig. 42.7, filtered bicarbonate is reabsorbed. Although large amounts of luminal HCO_3^- are reabsorbed in the proximal tubule, in the early proximal tubule segments, roughly proportional amounts of H_2O are reabsorbed, and thus the luminal HCO_3^- concentration remains similar to that of the glomerular filtrate until the late portions of the proximal tubule.

The Kidneys Excrete Excess Acid and Generate "New" Bicarbonate in the Process of Net Acid Excretion

As previously mentioned, although the respiratory system and bone contribute to maintaining the blood pH acutely, only the kidney is able to excrete the excess acid generated as a byproduct of metabolism. Also, the kidney's normal reabsorption of essentially all filtered bicarbonate does not contribute to acid excretion but merely restores the filtered bicarbonate to the plasma where it originated. Renal net acid excretion (NAE) not only involves proton secretion by cells but also the intracellular formation of "new" bicarbonate and its transport into the interstitium where it can enter the blood. Thus renal NAE is essentially equivalent to the amount of "new" bicarbonate contributed to the systemic circulation, and these amounts vary depending on the physiological conditions.

Renal Net Acid Excretion Is the Sum of Urinary Titratable Acid Plus Urinary Ammonium, Minus Urinary Bicarbonate

Total renal NAE is determined by the rate of three main processes, which are titratable acid excretion, ammonium excretion, and bicarbonate excretion. It is represented by the equation:

$$NAE = V(U_{TA} + U_{NH_4^+} - U_{HCO_3^-})$$

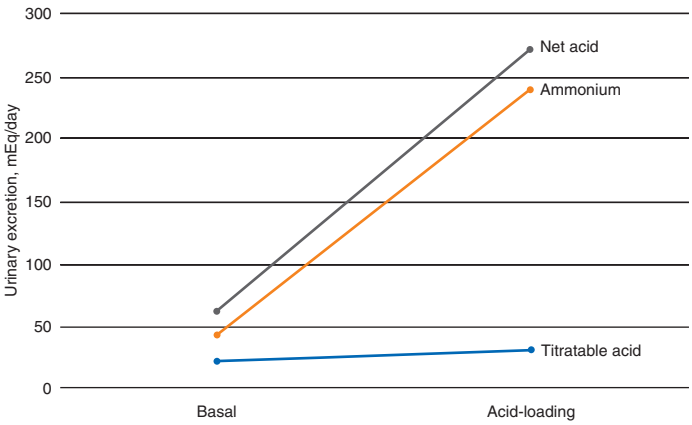
where V is the volume of urine excreted/unit time, U_{TA} is the urine concentration of titratable acids, U_{NH4}^+ is the urine concentration of ammonium, and U_{HCO3}^- is the urine concentration of bicarbonate. Under most circumstances in carnivores and omnivores, U_{HCO3}^- is essentially zero and NAE is effectively the sum of urinary excretion of titratable acids and ammonium. However, in some circumstances urinary bicarbonate excretion may be significant, in which case urinary bicarbonate excretion reduces NAE. In fact, in healthy silage-fed cattle, NAE is typically negative due to the net alkali load from the feed. Finally, although urinary pH can be as low as ~4.3 under extreme conditions, the contribution of free protons in the urine to NAE is trivial. This is because the lowest achievable urine pH is equivalent to an H^+ concentration in the range of 10 to 100 µmol/L, which is orders of magnitude less than the urinary concentration of titratable acids and ammonium.

Under most circumstances, the major components of NAE are titratable acid and ammonium excretion. *Titratable acids* are urinary solutes that are filtered in the base form and bind secreted protons in the tubule fluid. Titratable acids have a pKa higher than systemic pH and lower than typical urine pH, which enables filtration of the base form and titration in the tubule fluid to the acid form. *Ammonia* also serves as a buffer in the tubule fluid, but because the pKa of ammonia is high, ~9.15, it exists almost exclusively as ammonium in plasma and thus is distinct from titratable acids. The contribution of ammonia to NAE depends in large part on the intracellular generation of ammonia (NH_3) in the proximal tubule and subsequent protonation and excretion in the ammonium ion form (NH_4^+), rather than filtration of the base form and titration to the acid form in the tubule fluid. Compared with titratable acid excretion, ammonium excretion accounts for more acid excretion under basal conditions and increases more when acid excretion is upregulated; thus ammonium excretion is the most important component of renal NAE (Fig. 44.1).

Urinary Titratable Acid Forms From Protons Secreted by Tubule Epithelial Cells that Bind to Filtered Buffers in the Tubule Fluid and Contributes to New Bicarbonate Generation

The kidneys secrete H^+, primarily in the proximal tubule, the thick ascending limb of Henle's loop, and the collecting duct. These segments use different mechanisms to secrete protons and to control blood pH precisely. The proximal tubule secretes the majority of acid, whereas the collecting duct is the final controller of NAE and urine pH.

Most secreted H^+ is transported across the apical plasma membrane by the following three transporters: (1) a sodium ion (Na^+)/H^+ exchanger, (2) an H^+-adenosine triphosphatase (ATPase) pump, and (3) an H^+,K^+-ATPase pump. The Na^+/H^+ exchanger secretes acid by electrically neutral exchange of luminal Na^+ for intracellular H^+. The Na^+ gradient generated by basolateral Na^+,K^+-ATPase drives apical Na^+/H^+ exchange (secondary active transport).



• **Fig. 44.1** Comparison of the contribution of ammonium excretion and titratable acid excretion to net acid excretion (NAE) in humans at baseline and after acid load. Ammonium excretion accounts for more than half of NAE at baseline. With acid-loading, titratable acid excretion increases modestly, but the vast majority of the increase in net acid excretion is due to markedly increased ammonium excretion. Based on data from Lemann J, Bushinsky DA, Hamm LL: Bone buffering of acid and base in humans, *Am J Physiol Renal Physiol*. 2003;285:F811–F832.

Na^+/H^+ exchange is the main route of H^+ secretion in the proximal tubule and the thick ascending limb.

The electrogenic H^+-ATPase pump actively transports intracellular H^+ across the apical plasma membrane and contributes a net positive charge to the tubule fluid. The electrogenic proton pump is the primary mechanism of proton secretion in the collecting duct and also contributes to proton secretion in the proximal tubule.

The collecting duct H^+,K^+-ATPase pumps, which are similar to gastric and colonic proton pumps, actively secrete acid by electrically neutral exchange of intracellular H^+ for K^+ in the tubule fluid. Although the H^+-ATPase pump is responsible for most H^+ secretion by the collecting duct, H^+,K^+-ATPases may equal or exceed the acid secretion rate of the H^+-ATPase pump under some conditions.

Filtered phosphate is in the base form, and H^+ secreted into the tubule fluid titrates HPO_4^{2-} to form $H_2PO_4^-$. The monovalent phosphate ion ($H_2PO_4^-$) is lipid insoluble and is transported only at very low rates on the apical Na^+-inorganic phosphate (P_i) cotransporters present in the proximal tubule. Thus the secreted protons bound to $H_2PO_4^-$ are retained in the tubule fluid. This is the main component of titratable acidity excreted by mammalian kidneys. Other luminal buffers that contribute significantly to urinary titratable acidity are creatinine and citrate.

In birds, titration of luminal urate forms uric acid. Besides being lipid insoluble, uric acid also has a low aqueous solubility, and thus a significant portion of acid is removed as uric acid precipitates.

Ammonia is the other major buffer that contributes to acid excretion, and ammonium excretion is the major component of renal acid excretion under basal conditions and during metabolic acidosis. Unlike phosphate and bicarbonate, luminal ammonia/ammonium arises almost exclusively from proximal tubule production and secretion rather than from glomerular filtration. The critically important role of ammonia in acid excretion is discussed later.

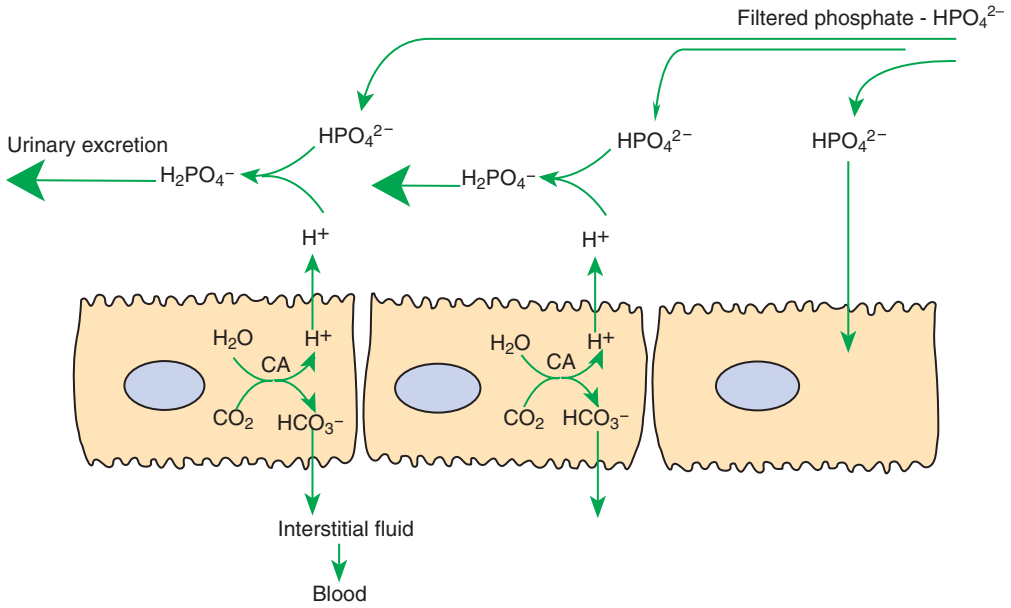
Although filtered bicarbonate buffers secreted protons, essentially all filtered bicarbonate is reabsorbed by the mechanism described previously. Thus although it is essential for maintaining acid-base homeostasis, it does not contribute to urinary titratable acid or to NAE.

Net apical acid secretion is coordinated with intracellular formation and basolateral absorption of "new" bicarbonate (vs. filtered bicarbonate), which replenishes plasma bicarbonate consumed by endogenous acid production. In the proximal tubule, bicarbonate formed in the cell is transported into the interstitial fluid primarily via the electrogenic sodium bicarbonate cotransporter, NBCe1-A. In the thick ascending limb, multiple basolateral bicarbonate transport mechanisms exist. Basolateral bicarbonate absorption is likely mediated in part by HCO_3^- substitution for Cl^- on the basolateral KCl cotransporter, KCC4, and by basolateral Cl^-/HCO_3^- exchange, possibly via anion exchanger 2 (AE2). In the collecting duct, the basolateral anion exchanger, kAE1, is a chloride/bicarbonate exchanger responsible for the majority of bicarbonate absorption by acid-secreting cells.

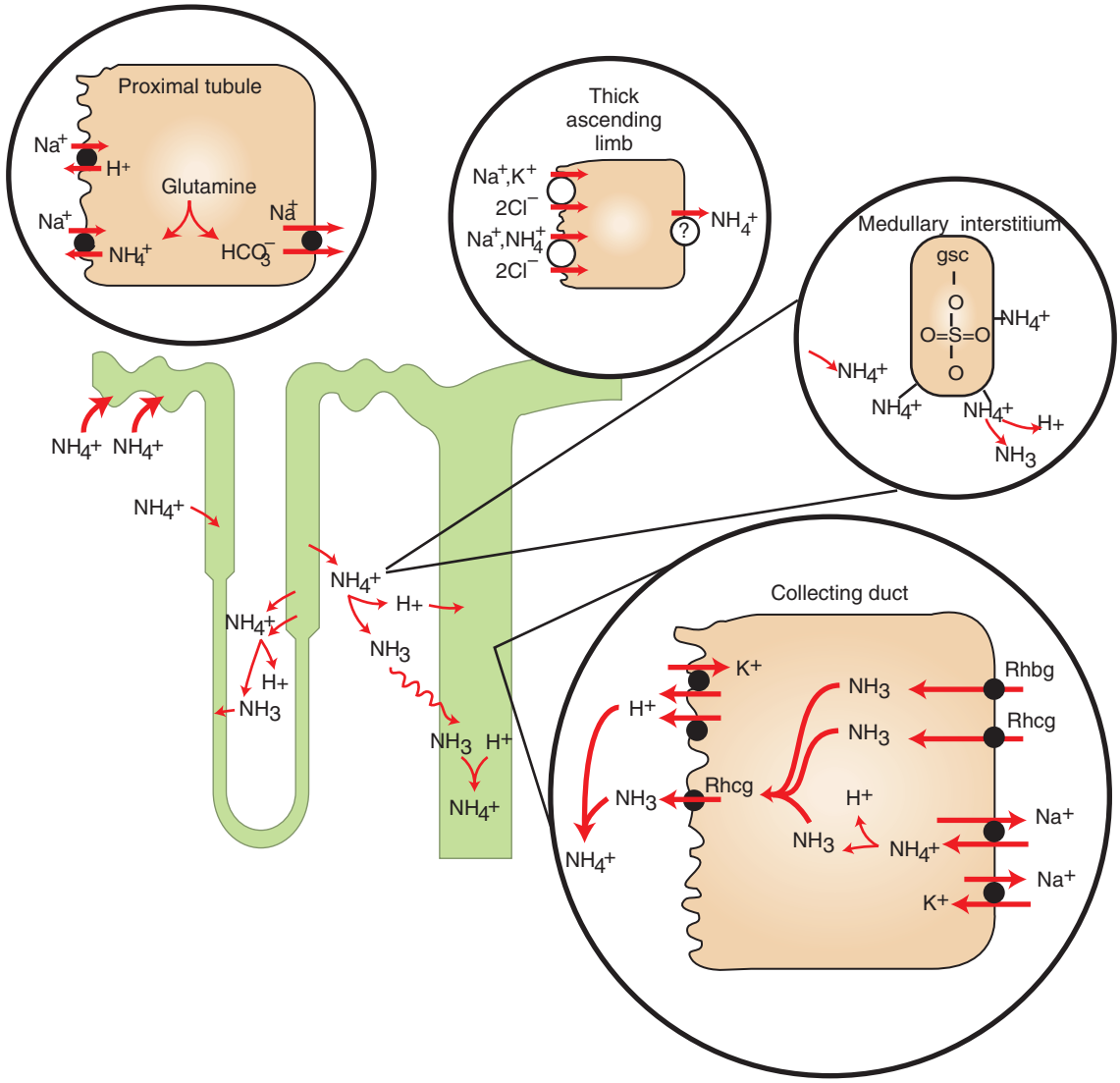
Fig. 44.2 schematically illustrates coordinated proton secretion, luminal buffering of secreted protons forming titratable acids, and absorption of "new" bicarbonate.

Renal Ammonia Metabolism Generates the Largest Component of Net Acid Excretion and New Bicarbonate Production

Renal ammonia metabolism is a major component in the maintenance of acid-base balance and is illustrated in Fig. 44.3. In proximal tubule cells, the amino acid *glutamine* is metabolized to produce NH_3 and H^+. This process is called *ammoniagenesis.* Net ammoniagenesis involves cytoplasmic and mitochondrial enzymes that metabolize glutamine as well as a process that assimilates



• **Fig. 44.2** Schematic illustration of proton secretion and new bicarbonate generation resulting from urinary titratable acid excretion. Filtered buffers, primarily HPO_4^{2-}, but also creatinine and citrate, are titrated to the acid form in the tubule fluid by secreted protons and excreted, and new bicarbonate formed in the cell is transported across the basolateral plasma membrane and absorbed into the blood.



• **Fig. 44.3** Schematic illustration of the roles of various nephron segments in ammonium excretion. In the proximal tubule, glutamine metabolism generates ammonium ion (NH_4^+) and bicarbonate (HCO_3^-). NH_4^+ is secreted into the lumen by substitution for H^+ on the Na^+/H^+ exchanger in the apical plasma membrane and HCO_3^- is transported into the interstitial fluid via the electrogenic sodium bicarbonate cotransporter, NBCe1-A. Ammonium ion recycles in the loop of Henle by reabsorption by the thick ascending limb, where NH_4^+ is reabsorbed by substitution for K^+ on the Na^+, K^+, $2Cl^-$ cotransporter in the apical plasma membrane, followed by some form of facilitated transport across the basolateral plasma membrane. The elevation of the interstitial NH_3 and NH_4^+ concentration results in movement into the descending thin limbs of Henle's loop and subsequent return to the thick ascending limb. This medullary recycling results in a high concentration of ammonia (NH_3) and NH_4^+ in the medullary interstitium, assisted by reversible binding to interstitial sulfatides (shown schematically in expanded view), and prevents its return to the cortex, where it would be reabsorbed into the blood. Ammonia or NH_4^+ is transported by specific ammonia transporters (the Rh glycoproteins, Rhbg and Rhcg) in the collecting duct and by substitution of NH_4^+ for K^+ on Na^+,K^+-ATPase in the inner medullary collecting duct, and is excreted in the urine.

ammonia to synthesize glutamine. The metabolism of glutamine involves mitochondrial phosphate-dependent glutaminase and cytoplasmic phosphoenolpyruvate carboxykinase, whereas the assimilation of ammonia to form glutamine is catalyzed by cytoplasmic glutamine synthetase. As mentioned previously, because the pKa of ammonia is ~9.15, at physiological pH essentially all ammonia is in the form NH_4^+. Intracellular NH_4^+ enters the tubule fluid through carrier-mediated transport by substitution for H^+ on the Na^+/H^+ exchanger. Glutamine metabolism also produces "new" bicarbonate anions, which are transported across the basolateral plasma membrane and replenish systemic bicarbonate. Thus proximal tubule ammoniagenesis enables acid secretion, new bicarbonate production and absorption, and distal delivery of ammonium. As shown in Fig. 44.1, renal ammoniagenesis and NH_4^+ excretion is enhanced by acidosis and is the main component of the renal response to an increase in the acid load.

In the thick ascending limb of Henle's loop, luminal NH_4^+ is reabsorbed by substitution for K^+ on the apical $Na^+,K^+,2Cl^-$ cotransporter (NKCC2). NH_4^+ reabsorption in this segment reduces the amount of ammonia species delivered in the luminal fluid to the late distal tubule and increases ammonium in the medullary interstitium.

High NH_4^+ concentrations are enhanced and maintained in the medullary interstitium by a countercurrent multiplication system in the loops of Henle, similar to that described in Chapter 43, and by reversible binding of NH_4^+ to sulfatides, anionic glycosphingolipids that are abundant in the medullary interstitium. This creates a steep concentration gradient for NH_3/ NH_4^+, which favors its movement from the medullary interstitium into medullary collecting duct cells. Until recently, the prevailing model for NH_4^+ excretion in the collecting duct was NH_3 *diffusion and trapping*. The belief was that NH_3 freely diffused across plasma membranes and into the luminal fluid where it bound protons to form NH_4^+, rendering it impermeant across plasma membranes and trapped in the luminal fluid. However, NH_3, although not a charged molecule, *is* a polar molecule much like water, and it is now known that the lipid bilayer is relatively impermeable to NH_3. In fact, specific ammonia transporters, the Rh glycoproteins, Rhcg and Rhbg, are present in plasma membranes of the majority of cell types in the collecting duct and facilitate transepithelial ammonia transport. These transporters are required for normal collecting duct ammonia transport and renal ammonia excretion, and their abundance and subcellular distribution are regulated in accordance with physiological conditions that increase renal ammonia excretion, such as acidosis. In addition, in the terminal inner medullary collecting duct, NH_4^+ is transported by substitution for K^+ on the basolateral Na^+,K^+-ATPase.

Essentially all secreted NH_3 binds to secreted H^+ to form NH_4^+ in the tubule luminal fluid. The formation of NH_4^+ from intraluminal NH_3 and H^+ lowers the concentration of both NH_3 and H^+ in the tubule fluid. This contributes to the maintenance of a favorable gradient for the transport of NH_3 into the tubule fluid, raises the tubule fluid pH, and reduces the electrochemical gradient for H^+ that is created by active proton secretion in the collecting duct.

Finally, the comparative aspects of ammonia excretion are intriguing. Ammoniagenesis and ammonia excretion are important mechanisms controlling acid-base homeostasis in mice, rats, dogs, chickens, and humans. In these species, ammonia excretion accounts for up to 60% of NAE in basal conditions and can increase to 90% of NAE in models of metabolic acidosis. However, these findings cannot be applied generally to other species. Rabbits have low basal urinary ammonia excretion rates and do not increase ammonia excretion during metabolic acidosis. In one model of metabolic acidosis, domestic cats acidified the urine but apparently did not increase renal ammoniagenesis and only increased urinary ammonia excretion to a level comparable to basal excretion in mice. In humans, rats, and mice, dietary potassium restriction increases renal ammonia excretion despite concurrent development of metabolic alkalosis, but potassium restriction causes metabolic acidosis in dogs and cats and reduces renal ammonia excretion in dogs.

The Proximal Tubule Has a High Capacity for H^+ Secretion, Bicarbonate Reabsorption, and Ammonia Metabolism

The proximal tubule not only reabsorbs the majority of the filtered bicarbonate, it also contributes to net acid secretion by secretion of protons excreted as titratable acids, generates and secretes ammonium ion, and generates new bicarbonate intracellularly. HCO_3^- is transported to the blood side of the tubule primarily by the electrogenic basolateral sodium bicarbonate cotransporter, NBCe1-A. Concurrently, H^+ is transported into the lumen, primarily by the Na^+/H^+ antiporter, but also by the H^+-ATPase pump, which may transport up to 35% of the total H^+ secreted by the proximal tubule.

In animals with normal kidney function, acid secretion, ammonia metabolism, and bicarbonate generation and absorption in the proximal tubule are increased during chronic metabolic acidosis. Several hormones regulate proximal tubule bicarbonate reabsorption and acid secretion: angiotensin II stimulates transport by the basolateral sodium bicarbonate cotransporter NBCe1-A, the apical Na^+/H^+ exchanger, and the apical vacuolar H^+-ATPase; glucorticoid receptor activation and endothelin enhance the Na^+/H^+ exchanger; and parathyroid hormone suppresses the Na^+/H^+ exchanger and basolateral sodium bicarbonate transport.

Although the proximal tubule has a large capacity for H^+ secretion and reabsorbs 80% or more of the filtered HCO_3^-, it cannot maintain a large pH gradient across the apical plasma membrane. Net H^+ secretion in this segment is particularly dependent on the intraluminal buffers discussed earlier, which combine with secreted H^+ and prevent the H^+ concentration in the tubule fluid from rising significantly. As a result, although the majority of renal acid secretion occurs in the proximal tubule, the pH of the tubule fluid when it leaves this segment is similar to that of the glomerular filtrate.

The Thick Ascending Limb of Henle's Loop Absorbs Filtered Bicarbonate

Bicarbonate reabsorption continues in the thick ascending limb of Henle's loop, which reabsorbs approximately 15% of filtered bicarbonate. Carbonic anhydrase generates protons and bicarbonate for transport as in the proximal tubule. In the thick ascending limb, protons are secreted primarily by apical Na^+/H^+ exchange, also similar to the proximal tubule. Several mechanisms of basolateral bicarbonate uptake have been described, including substitution of bicarbonate for chloride on the basolateral KCl cotransporter, KCC4, and via the basolateral anion exchanger, AE2.

The Collecting Duct Determines the Final Urine pH

The rate of acid secretion by the collecting duct determines the final urine pH and renal NAE. Despite robust acid secretion in the proximal tubule and the additional contribution by the thick ascending limb, because of luminal buffering the pH of the tubule fluid that reaches the connecting segment is still similar to that of the glomerular filtrate, pH approximately 7.4. However, the normal urine pH of carnivores ranges from 5.5 to 7.5, that of ruminants ranges from 6 to 9, and even greater extremes of pH occur in response to acidosis and alkalosis. The collecting duct is responsible for this ability to excrete urine with a pH extremely different from that of plasma.

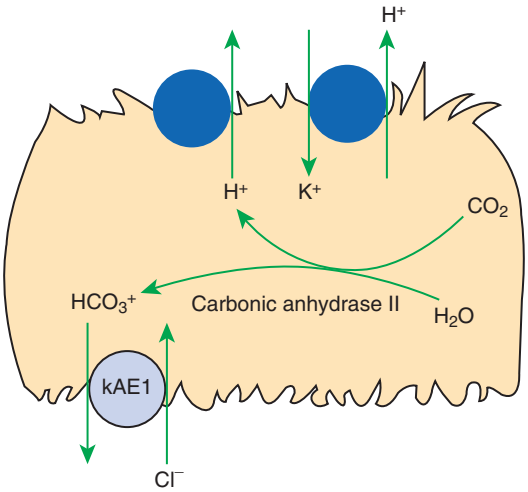
The Collecting Duct Can Secrete Protons and Ammonia, Reabsorb Bicarbonate, and Generate Acidic Urine

In contrast to the proximal tubule, which is a high-capacity, low-gradient system of H^+ secretion, the collecting duct has a lower capacity for H^+ secretion but can generate a steep H^+ concentration gradient.

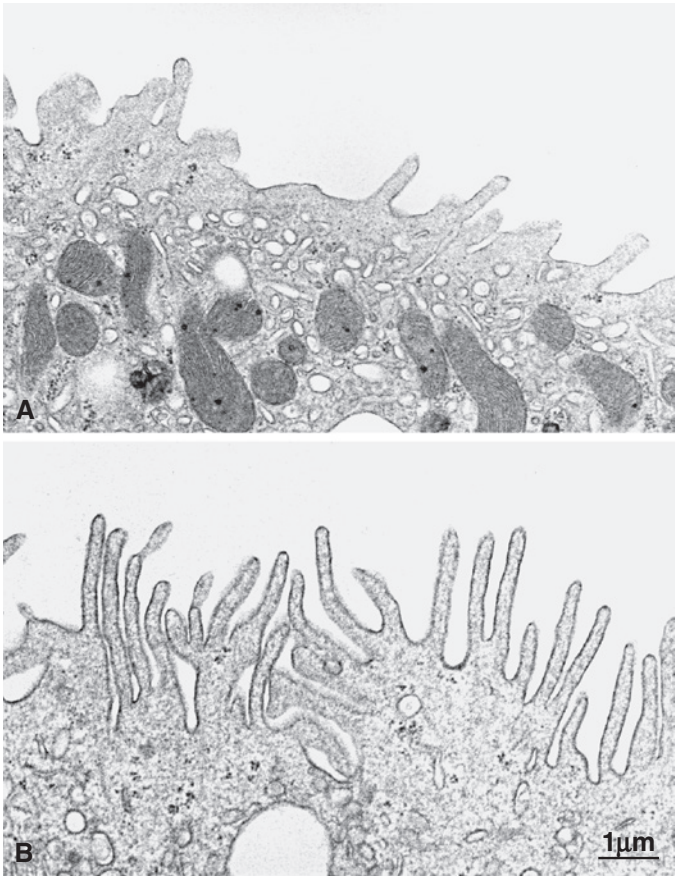
Acid secretion in most of the collecting duct system is primarily a function of specialized epithelial cells, the *intercalated cells,* illustrated in Fig. 42.13. Intercalated cells constitute approximately 40% of cells in the connecting segment, cortical collecting duct, and outer medullary collecting duct in most species examined, diminishing and ultimately disappearing in the inner medullary collecting duct. Intercalated cells contain abundant cytoplasmic carbonic anhydrase, which catalyzes the formation of intracellular H^+ and HCO_3^- from intracellular H_2O and CO_2. The collecting duct contains specific intercalated cell subtypes that are structurally and functionally distinct; approximately half of the intercalated cells in the renal cortex and virtually all in the renal medulla are an acid-secreting subtype (type A or α). In these cells, H^+ is secreted into the tubule fluid by the apical electrogenic proton pump, H^+-ATPase, or by electrically neutral H^+,K^+-ATPase pumps. HCO_3^- is transported to the blood side of the cell by a basolateral Cl^-/HCO_3^- exchanger (kidney anion exchanger 1, kAE1) similar to the Cl^-/HCO_3^- exchanger in red blood cell membranes (Fig. 44.4). Acid-secreting intercalated cells alter the rate of H^+ secretion by altering the number of proton pumps in the apical plasma membrane. The insertion or removal of proton pump–containing membrane vesicles causes structural changes that reflect the physiological response (Fig. 44.5). These structural changes during metabolic acidosis also coincide with translocation of the ammonia transporter, Rhcg, from intracellular compartments to the apical plasma membrane, which enhances ammonia excretion and thus NAE. In rabbits, the Cl^-/HCO_3^- exchanger, kAE1, also is translocated from intracellular compartments to the basolateral plasma membrane in acidosis. In this way, the acid-secreting intercalated cells respond to physiological conditions and alter acid secretion accordingly. Less is known about the roles and response mechanisms of the renal H^+,K^+-ATPases in acidosis, but metabolic acidosis and hypokalemia enhance renal H^+,K^+-ATPase activity.

The terminal segments of the inner medullary collecting duct, where there are few or no intercalated cells, also can secrete acid. A Na^+/H^+ exchanger, an electrogenic proton pump, a H^+,K^+-ATPase pump, and substitution of NH_4^+ for K^+ on the basolateral Na^+,K^+-ATPase all participate in acid secretion in this segment, but the relative importance of these mechanisms is currently unclear.

Acid secretion in the collecting duct is generally enhanced by acidosis and suppressed by alkalosis. The hormones angiotensin



• **Fig. 44.4** Schematic illustration of the mechanisms of H^+ secretion and HCO_3^- reabsorption in the acid-secreting intercalated cells of the collecting duct. Two means of active transport of H^+ across the apical plasma membrane are present: the electrogenic proton pump, H^+-ATPase, and the electrically neutral H^+,K^+-ATPase pump. The intracellular formation of H^+ and HCO_3^- from CO_2 and H_2O is catalyzed by cytoplasmic carbonic anhydrase. The basolateral plasma membrane contains a Cl^-/HCO_3^- exchanger (kAE1) that allows HCO_3^- reabsorption.



• **Fig. 44.5** Transmission electron micrographs of an acid-secreting (type A) intercalated cell from rat cortical collecting duct. (A) In a control animal, the apical plasma membrane contains few small membranous projections, and the apical cytoplasm is filled with numerous membrane vesicles. (B) In a rat with acute respiratory acidosis, the apical surface is covered with numerous long, membranous projections, and the number of apical cytoplasmic vesicular profiles is greatly reduced. This is the result of the insertion of membrane vesicles containing H^+ transporters and the ammonia transporter, Rhcg, into the apical plasma membrane in response to acidosis, thus enhancing the acid- and ammonia-secreting capacity of the cell.

II, aldosterone, and endothelin stimulate H^+-ATPase in acid-secreting intercalated cells in the collecting duct, but the roles of these hormones in systemic acid-base homeostasis are not fully understood.

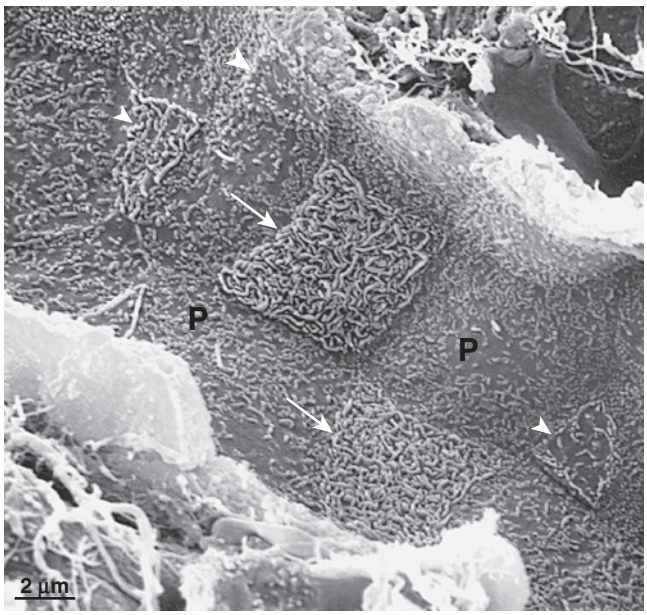
The Collecting Duct Can Secrete Bicarbonate and Generate Alkaline Urine

The proximal tubule reabsorbs HCO_3^- and secretes H^+, regardless of the plasma HCO_3^- concentration and the blood pH. In fact, as the plasma HCO_3^- concentration increases, the concentration of HCO_3^- in the glomerular filtrate increases, and the amount of HCO_3^- reabsorption by the proximal tubule epithelium also increases. However, portions of the collecting duct are capable of net HCO_3^- secretion in response to alkalosis. In addition to acid-secreting intercalated cells (type A or α) similar to those in the medulla, two additional intercalated cell subtypes (type B or β, and non-A, non-B or type C) that can secrete bicarbonate are present in the connecting segment and the cortical collecting duct (Fig. 44.6). These cells are rich in carbonic anhydrase and secrete HCO_3^- by an apical Cl^-/HCO_3^- exchanger, pendrin, which is distinct from the basolateral Cl^-/HCO_3^- exchanger in acid-secreting intercalated cells. Type B intercalated cells have a basolateral proton pump and functionally represent a mirror image of acid-secreting cells, with active H^+ reabsorption and exchange of Cl^- in the tubule fluid for intracellular HCO_3^- (Fig. 44.7). Type C (non-A, non-B) intercalated cells have the apical anion exchanger, pendrin, as well as apical H^+-ATPase. These intercalated cells secrete bicarbonate, but the apical proton pump suggests that, under some circumstances, they may not contribute to net bicarbonate excretion but may have a more important role in chloride reabsorption.

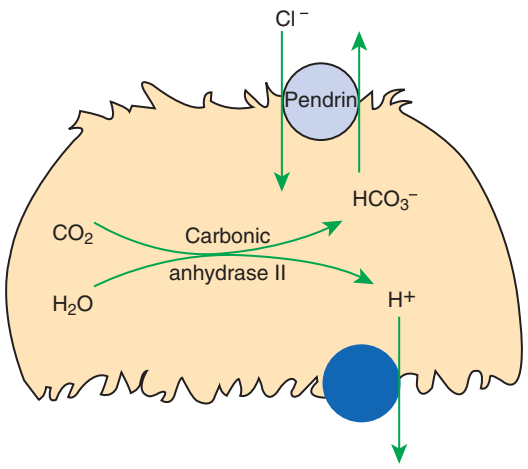
Alkalosis stimulates and acidosis suppresses bicarbonate secretion. Bicarbonate secretion is also stimulated by dietary restriction of NaCl or Cl^- alone, and by aldosterone analogs and angiotensin II. These hormones also enhance apical proton transport by acid-secreting intercalated cells, which seems counterproductive. However, the stimulation of both proton secretion and apical chloride/bicarbonate exchange by these hormones enhances Cl^- and Na^+ uptake, and is an additional mechanism supporting the NaCl retention stimulated by aldosterone and angiotensin II. The role of intercalated cells in NaCl uptake in the collecting duct is discussed in Chapter 42.

In Cl^- depletion, the amount of Cl^- delivered to the collecting duct may fall so low that bicarbonate secretion is impaired because inadequate luminal Cl^- is available for exchange with intracellular HCO_3^-. The blunted bicarbonate secretion contributes to the development and maintenance of metabolic alkalosis during Cl^- depletion, a condition commonly seen in patients after continuous gastric suction, gastrointestinal obstruction, diuretic treatment, and certain forms of diarrhea.

Little is known about the comparative physiology of renal control of acid-base balance in domestic animals, although the renal tubules are similar in all mammals examined to date, and intercalated cells have been observed at least in cat, dog, pig, and horse collecting ducts. However, it is likely that considerable anatomical and functional differences exist among species, particularly between carnivores, which usually excrete acid urine, and ruminants, which usually excrete neutral or alkaline urine.



• **Fig. 44.6** Scanning electron micrograph of rat cortical collecting duct viewed from the tubule lumen. Three cell types are evident. The principal cells (*P*) are large, with a single central cilium and few apical surface microprojections. The type A (acid-secreting) intercalated cells (*arrows*) have a large apical surface covered with extensive membrane folds (microplicae). The type B (bicarbonate-secreting) intercalated cells (*arrowheads*) have a small apical surface area covered with sparse microprojections.



• **Fig. 44.7** Schematic illustration of the mechanism of HCO_3^- secretion (H^+ reabsorption) in the type B intercalated cell of the cortical collecting duct. These cells contain H^+-ATPase in the basolateral plasma membrane and are rich in cytoplasmic carbonic anhydrase. The Cl^-/HCO_3^- exchanger, pendrin, is present in the apical plasma membrane and mediates bicarbonate secretion.

CLINICAL CORRELATIONS

Respiratory Acidosis With Renal Compensation

History

A 6-year-old male German shepherd is brought to you with complaints of weakness, exercise intolerance, and poor appetite that have progressively worsened over the past 6 weeks.

Clinical Examination

The dog is recumbent and anxious. The heart rate is rapid, but pulses are strong and regular. Respiration is labored, and crackles are heard over all lung fields. Thoracic radiographs reveal a diffuse, severe pulmonary interstitial and alveolar infiltrate with enlargement of the hilar lymph nodes. You obtain samples of blood and urine for a complete blood cell count (CBC), serum chemistry panel, urinalysis, and arterial blood gas (ABG) measurement. The urine pH is 5.0, and the ABG results are as follows: pH, 7.37 (normal, 7.45); P_{O_2} (oxygen tension), 58 mm Hg (normal, 80–100 mm Hg); P_{CO_2} (carbon dioxide tension), 70 mm Hg (normal, 31–35 mm Hg); and HCO_3^-, 37 mEq/L (normal, 18–24 mEq/L).

Comment

The dog has *chronic respiratory acidosis* caused by severe pulmonary infiltrates. The lung is unable to ventilate adequately, and the blood level of CO_2 rises. The elevated CO_2 favors the production of carbonic acid, which releases H^+ and lowers the blood pH. Although the increased CO_2 level contributes to an increase in blood HCO_3^- levels, the marked increase in the blood HCO_3^- in this case results from enhanced renal retention of HCO_3^- and secretion of H^+. Respiratory acidosis activates the acid-secreting intercalated cells in the collecting duct, where HCO_3^- is reabsorbed and H^+ is secreted. The blood HCO_3^- concentration rises and helps return the blood pH toward normal. Despite the increased blood HCO_3^-, expression of the apical Cl^-/HCO_3^- exchanger, pendrin, is suppressed by respiratory acidosis and thus bicarbonate secretion is likely impaired. A steep H^+ gradient is established in the collecting ducts, and acid urine is excreted. (Although metabolic acidosis enhances ammoniagenesis in the proximal tubule, thereby enhancing the generation of new bicarbonate and acid excretion in the form of ammonium ion, this does not occur in respiratory acidosis, at least in rats.)

Treatment

Diagnose and correct the pulmonary disease, if possible. Bicarbonate therapy is not indicated, because the blood bicarbonate level is already high, and the blood pH is partially corrected. Oxygen therapy may improve the Pa_{O_2} and help support the animal until specific treatment is instituted.

Metabolic Alkalosis With Paradoxical Aciduria

History

You examine a 3-year-old Holstein-Friesian cow that has had a poor appetite for 2 to 3 days. The cow recently calved and freshened normally, but milk production has dropped in the last 2 days, and the feces are loose.

Clinical Examination

Physical examination reveals dehydration and an elevated heart rate. Percussion of the abdomen detects an area of high-pitched resonance on the right side. A distended abomasum is palpable on rectal examination. You diagnose a right displaced abomasum and suspect abomasal torsion. Attempts to correct the displacement by rolling the cow fail. The cow is transported to your clinic for surgery, and samples are obtained for a CBC, serum chemistry profile, and urinalysis. The serum K^+ level is 2.7 mEq/L (normal, 4.0–5.1 mEq/L), serum Cl^- level is 77 mEq/L (normal, 85–103 mEq/L), and total CO_2 concentration (approximately equivalent to serum HCO_3^- concentration) is 35 mEq/L (normal, 24–27 mEq/L). The urine pH is 6.0.

Comment

The cow has *hypokalemic, hypochloremic, metabolic alkalosis* secondary to abomasal displacement. The alkalosis was initiated by continued secretion of HCl by the abomasum, sequestration of the secreted HCl in the abomasal lumen, and blunted HCO_3^- secretion by the intestine after the gastrointestinal obstruction. The hypokalemia is likely a result of intracellular movement of K^+ secondary to alkalosis and may not reflect a decrease in total body K^+ levels, although in chronic chloride-depletion metabolic alkalosis, significant renal potassium losses may occur and cause K^+ depletion.

The kidney normally responds to alkalosis by excreting alkaline urine. However, in this case the volume contraction and hypochloremia prevent the formation of alkaline urine, and the result is *paradoxical aciduria*. The proximal tubule reabsorbs filtered HCO_3^-, regardless of the plasma pH or serum HCO_3^- concentration. The volume depletion enhances Na^+ reabsorption in the collecting duct and Cl^- and H_2O reabsorption are enhanced secondary to the increased Na^+ uptake.

Renal secretion of HCO_3^- occurs by apical exchange of Cl^- in the tubule fluid for intracellular HCO_3^- in type B and type C (non-A, non-B) intercalated cells in the collecting duct and connecting segment. Because NaCl is avidly reabsorbed to combat volume depletion, little Cl^- remains for exchange with HCO_3^-, and net HCO_3^- secretion does not occur. Acid secretion in the collecting duct increases in response to aldosterone and may be enhanced in this volume-depleted animal. In rodents, hypokalemia activates the acid-secreting intercalated cells in the collecting duct, increases apical proton pump activity, increases proximal tubule ammoniagenesis, and increases ammonia excretion via principal cell and intercalated cell ammonia transporters; these events may also occur in cattle and may contribute to increased net acid excretion and maintenance of the metabolic alkalosis in this case.

Treatment

Treatment involves vigorous volume replacement with intravenous normal saline, with KCl added, and surgical correction of the abomasal displacement.

Practice Questions

1. In carnivores, the usual role of the kidney in maintaining acid-base homeostasis is to:
- **a.** Secrete excess bicarbonate.
- **b.** Secrete excess ammonia.
- **c.** Secrete excess acid.
- **d.** Secrete excess carbon dioxide.
- **e.** Secrete excess phosphate buffer.

2. The bulk of acid secretion is accomplished by which renal tubule segment?
- **a.** Proximal tubule
- **b.** Thin limbs of Henle's loop
- **c.** Thick ascending limb of Henle's loop
- **d.** Distal convoluted tubule
- **e.** Collecting duct

3. Which of the following factors does *not* contribute to efficient acid excretion by the renal tubules?
- **a.** Primary active transport of bicarbonate
- **b.** Ammoniagenesis in the proximal tubule
- **c.** Intraluminal buffering by phosphate
- **d.** Intracellular and membrane-associated carbonic anhydrase
- **e.** Transmembrane proton transport by the Na^+/H^+ exchanger, H^+-ATPase pump, and H^+,K^+-ATPase pump

4. Which of the following statements regarding mechanisms of acid-base regulation by the collecting duct is *false?*
   - **a.** The cortical collecting duct responds to acidosis by increasing the net rate of acid secretion.
   - **b.** The cortical collecting duct responds to alkalosis with net bicarbonate secretion.
   - **c.** Proton and bicarbonate transport in the collecting duct are only slightly altered in response to systemic acid-base disturbances.
   - **d.** The collecting duct determines the ultimate pH of the urine.
   - **e.** The intercalated cells are largely responsible for acid secretion by the collecting duct.

5. What is the role of renal ammonia metabolism in the renal response to acidosis, at least in dogs and rodents?
   - **a.** Acidosis increases ammoniagenesis in the proximal tubule, which increases net acid excretion and the generation of new bicarbonate ions.
   - **b.** Acidosis increases collecting duct ammonia secretion, which enhances acid secretion.
   - **c.** Acidosis stimulates ammoniagenesis in the proximal tubule and inhibits collecting duct ammonia secretion, which increases ammonia buffering of the plasma.
   - **d.** Renal ammonia metabolism does not contribute to renal acid-base regulation.
   - **e.** Both a and b.

Bibliography

Hamm LL, Nakhoul N, Hering-Smith KS. Acid-base homeostasis. *Clin J Am Soc Nephrol*. 2015;10(12):2232–2242.

Wall SM, Lazo-Fernandez Y. The role of pendrin in renal physiology. *Annu Rev Physiol*. 2015;77:363–378.

Weiner ID, Verlander JW. Ammonia transporters and their role in acid-base balance. *Physiol Rev*. 2017;97(2):465–494.

Weiner ID, Verlander JW. Renal acidification mechanisms. In: Skorecki K, Chertow GM, Marsden PA, et al, eds. *Brenner & Rector's the Kidney*. ed 10. Philadelphia: Elsevier; 2016.

45

Overview of Respiratory Function: Ventilation of the Lungs

SUSAN L. EWART

KEY POINTS

Respiratory Function

1. The respiratory system transports oxygen and carbon dioxide between the environment and the blood.

Ventilation

1. Lung volumes vary throughout the respiratory cycle.
2. Alveolar ventilation is the fresh air per minute that reaches perfused alveoli while dead space ventilation remains in the conducting airways.
3. Ventilation requires muscular energy.
4. Lung elasticity results from tissue and surface tension forces and determines airway compliance.
5. Surfactant reduces elastic recoil of the lungs and thus increases compliance.
6. The lung is functionally coupled to the chest wall.
7. Airflow is opposed by frictional resistance.
8. Smooth muscle contraction affects the diameters of the trachea, bronchi, and bronchioles.
9. The work of breathing is increased by exercise, thermoregulation, and disease.
10. Dynamic compression can narrow the airways and limit airflow.
11. The distribution of ventilation depends on the local mechanical properties of the lung.
12. In some species, air travels between adjacent regions of lung through collateral pathways.

Respiratory Function

The Respiratory System Transports Oxygen and Carbon Dioxide Between the Environment and the Blood

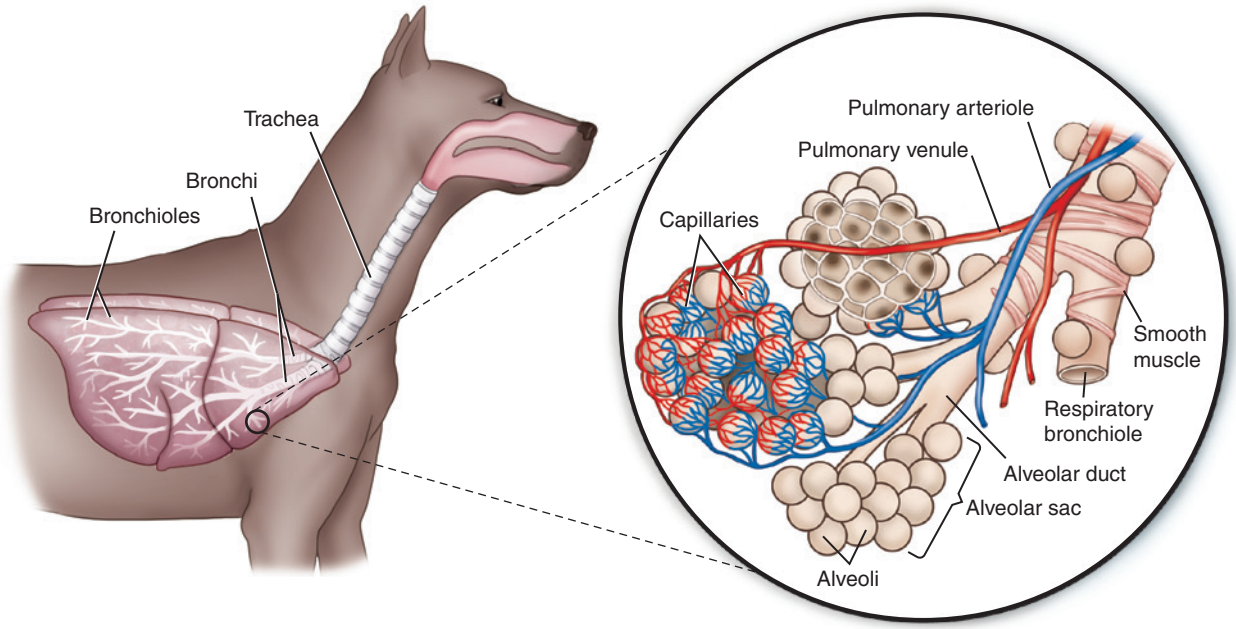
Structurally, the respiratory system can be divided into upper and lower airways, with the former being composed of the *nares, nasal passages, pharynx, larynx,* and (by some authors) *trachea;* these structures conduct air to and from the lungs. After entering the thorax, the trachea divides into paired mainstem *bronchi,* which give rise to several generations of bronchi followed by multiple generations of smaller airways, termed *bronchioles* (see Fig. 45.1). This *tracheobronchial tree* is lined by a secretory, ciliated epithelium. The larger airways—trachea and bronchi—are supported by cartilage and contain bronchial glands and goblet cells, the secretions of which contribute to the mucous lining of the airways. The bronchioles lack cartilage, glands, and goblet cells; the liquid lining the bronchiolar epithelial surface originates in club cells (previously referred to as Clara cells). The *conducting airways,* from the nares to the bronchioles, serve to conduct air from the environment to the gas exchange regions of the lungs, as well as *condition* it by filtering out particles, warming it to body temperature, and humidifying it to saturation with water vapor.

At the periphery of the lungs, the terminal bronchioles branch into *respiratory bronchioles,* then *alveolar ducts, alveolar sacs,* and finally *alveoli.* These peripheral structures are the site of gas exchange as oxygen (O_2) moves from the alveolar air to the pulmonary capillaries and carbon dioxide (CO_2) moves in the reverse direction.

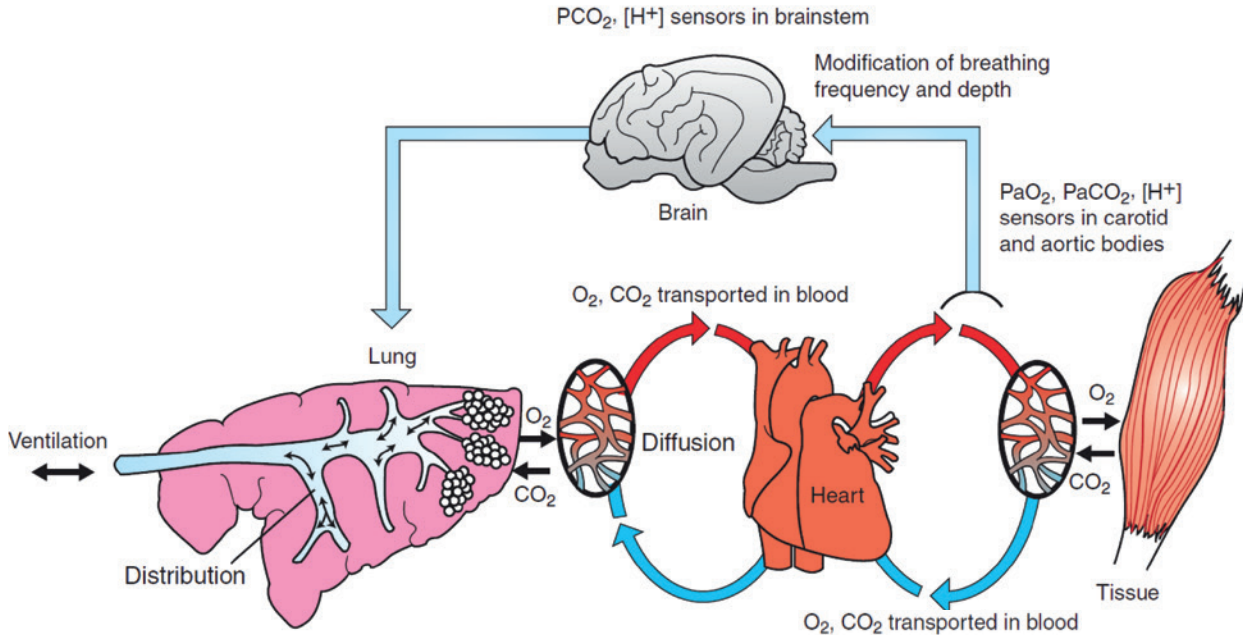
Fig. 45.2 shows the processes involved in gas exchange, including *ventilation, distribution* of gas within the lung, *diffusion* between the alveoli and the pulmonary capillary blood, *transport* of O_2 in the blood from the lungs to the tissue capillaries and of CO_2 in the reverse direction, and *diffusion* of gases between blood and tissues.

The respiratory system provides O_2 to support tissue metabolism and removes the byproduct of metabolism, CO_2. *Oxygen consumption* and *carbon dioxide production* vary with the metabolic rate, which is dependent on the animal's level of activity and physical condition. *Basal metabolism,* the metabolism rate of the resting animal, is a function of *metabolic body weight* (body weight in $kg^{0.75}$). The consequence of this relationship is that smaller animals consume more O_2 per kg of body weight than larger animals. For example, a 20 g mouse consumes six times more O_2 per unit body mass than does a 70 kg pig. This difference is largely due to the metabolic requirements necessary to maintain constant body temperature. Because smaller animals have a greater surface area to body weight ratio, they have a greater surface for heat loss and less heat storage capacity so they need higher basal metabolism to generate more heat.

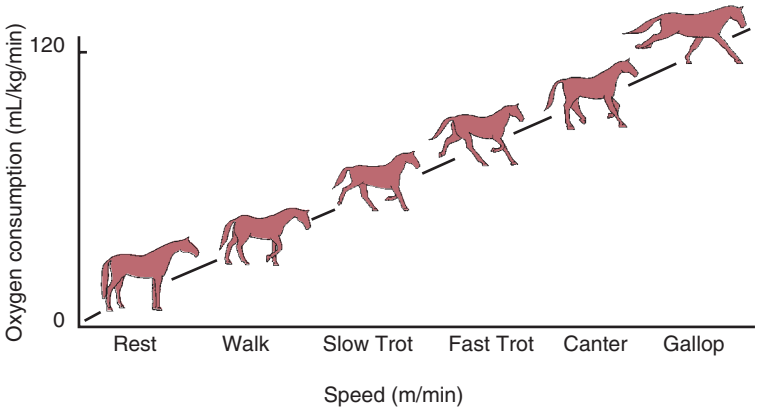
When animals exercise, their muscles need more O_2, which leads to an increase in O_2 consumption. *Maximal O_2 consumption* ($\dot{V}O_{2max}$) is directly related to the total mass of mitochondria within



• **Fig. 45.1** Anatomy of the respiratory system. The upper airways are extra-thoracic and include the nares, nasal passages, pharynx, and larynx. Connecting the upper and lower airways is the trachea, which has both extra-thoracic and intra-thoracic parts, according to the region of the body that it traverses. The lungs are intra-thoracic and consist of a tube (the trachea) that branches repeatedly (over 20 to 40 times) until it reaches the terminal airways where gas exchange with capillary blood takes places in alveoli. The branching structures within the lungs make up the lower airways and include left and right mainstem bronchi, which branch into lobar bronchi, followed by segmental bronchi, then multiple generations of smaller bronchi and finally into multiple generations of bronchioles. The branching continues as bronchioles lead to gas exchange regions of respiratory bronchioles, alveolar ducts, alveolar sacs, and alveoli.



• **Fig. 45.2** Diagrammatic representation of the processes involved in gas exchange.



• **Fig. 45.3** Effect of exercise on O_2 consumption in the horse. Oxygen consumption increases in a linear manner as the horse increases speed. (Modified from Hörnicke H, Meixner R, Pollman U: *Equine exercise physiology,* Cambridge, UK, 1983, Granta Editions.)

the skeletal muscles. Athletic species, such as horses and dogs, have greater mitochondrial density and therefore greater $\dot{V}O_{2max}$ than less athletic species of similar body size, such as cattle and goats. *Gas exchange* requirements vary with metabolism and increase greatly during strenuous exercise; the $\dot{V}O_{2max}$ in racing horses can approach 120 mL/kg/min, which is up to thirtyfold higher than baseline (Fig. 45.3).

The respiratory system is also involved in communication by sound and pheromones, thermoregulation, metabolism of endogenous and exogenous substances, and protection of the animal against inhaled dusts, toxic gases, and infectious agents (see Chapter 50). Additionally, the increase in abdominal pressure that facilitates urination, defecation, and parturition requires active participation of the respiratory muscles.

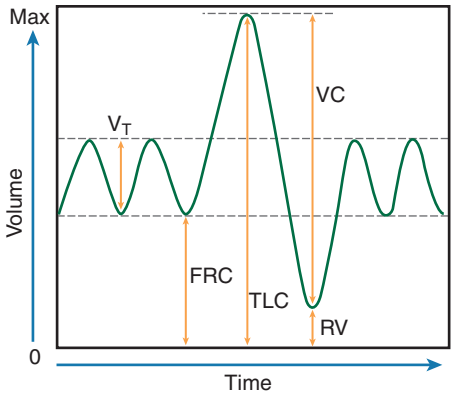
Ventilation

Lung Volumes Vary Throughout the Respiratory Cycle

The ease with which the normal lung expands and contracts with each breath is due to its extremely compliant nature. This allows the dramatic changes in lung volume that occur during the respiratory cycle. The volumes of air within the lung range from *total lung capacity* (TLC), which is the amount of air the lungs can hold following a maximum inspiration, to *residual volume* (RV), which is the volume of air that remains in the lungs after maximal, forced expiration (Fig. 45.4). *Vital capacity* (VC) is the maximum volume of air that can be moved, that is, the volume difference between TLC and RV. *Tidal volume* (V_T) is the volume of air moved into or out of the lungs during quiet resting breathing. At the end of a quiet resting exhalation, some air (~45 mL/kg) remains in the lung; this air volume is known as *functional residual capacity* (FRC).

Alveolar Ventilation is the Fresh Air Per Minute That Reaches Perfused Alveoli While Dead Space Ventilation Remains in the Conducting Airways

The O_2 needs of metabolism require that an animal take a certain volume of air into its lungs, specifically into its *alveoli,* each minute.



• **Fig. 45.4** Lung volumes and capacities throughout the respiratory cycle. *FRC,* Functional residual capacity; *RV,* residual volume; *TLC,* total lung capacity; *VC,* vital capacity; V_T, tidal volume. See text for description of terms.

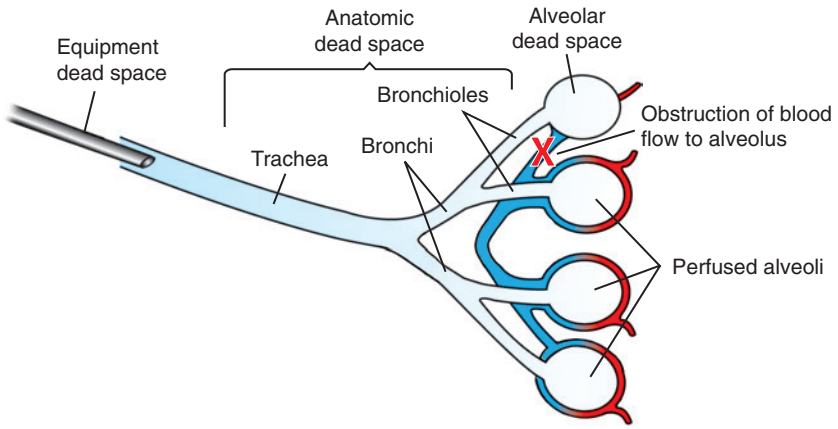
The total volume of air breathed per minute, known as *minute ventilation* ($\dot{V}_E$), is determined by the volume of each breath (V_T), and the *breathing frequency* (f) in breaths per minute, as follows:

$$\dot{V}_E = V_T \times f$$

An increase in V_E, which occurs in response to an increase in metabolic rate requiring more O_2, can be brought about through an increase in V_T, f, or both.

Before reaching the gas exchange regions of the lungs, air flows through the conducting airways, which are the anatomic structures from the nares to the bronchioles. Because gas exchange does not occur in these locations, they are termed *anatomic dead space* (Fig. 45.5). *Dead space,* which refers to inhaled air that does not participate in gas exchange, can also occur within the alveoli. This *alveolar dead space* is the result of ventilated alveoli that are not perfused with blood, so that gas exchange cannot occur (see Chapter 47). *Physiologic dead space* is the sum of the anatomic and alveolar dead space; it is the physiologic dead space that determines the portion of each breath that is unavailable for gas exchange. The V_T is the sum of the volume of air that enters the perfused alveoli (V_A) and the volume that remains in the physiologic dead space (V_D), as follows:

$$V_T = V_A + V_D$$



• **Fig. 45.5** Respiratory dead space includes the ventilated parts of the respiratory system where gas exchange does not occur. Four schematic alveoli are shown attached to the conducting airways and perfused by differing amounts of blood. The volume of the conducting airways constitutes the anatomic dead space; the part of the endotracheal tube extending beyond the respiratory system constitutes equipment dead space; and the volume of air ventilating unperfused alveoli is alveolar dead space. *Top alveolus,* an unperfused alveolus is dead space because there is no blood flow to allow gas exchange; *Bottom three alveoli,* ideally perfused alveoli contribute no dead space because all of the air participates in gas exchange.

To determine the *ventilation,* we multiply each side of this equation by f, as follows:

$$V_T \times f = (V_A \times f) + (V_D \times f)$$

The result is:

$$\dot{V}_E = \dot{V}_A + \dot{V}_D$$

Therefore $\dot{V}_E$ is the sum of *alveolar ventilation* ($\dot{V}_A$), which is essential for gas exchange, and *dead space ventilation* (V_D), which is "wasted" ventilation.

Alveolar ventilation is the most important ventilatory parameter because it is the amount of fresh air available for gas exchange. Only air that enters alveoli that are perfused with blood constitutes alveolar ventilation. Alveolar ventilation is tightly regulated to match the O_2 uptake and CO_2 elimination necessitated by metabolism. Thus when an animal exercises, its alveolar ventilation increases to take in more O_2 and eliminate more CO_2.

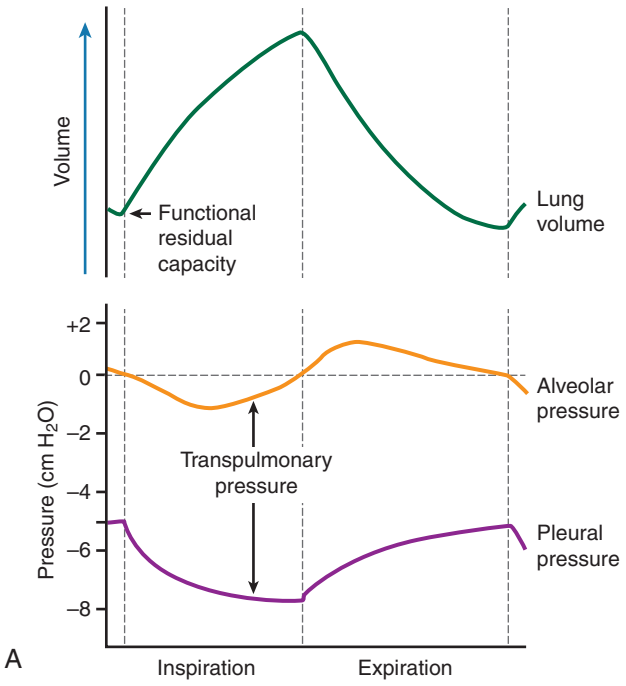
The *fraction* of each breath ventilating the physiologic dead space is known as the *dead space/tidal volume ratio* (V_D/V_T). The V_D/V_T varies considerably among species. In smaller species, such as dogs, it approximates 33%, whereas in some larger species, such as cattle and horses, it is approximately 50% to 75%. Because the anatomic dead space volume is relatively constant, changes in V_T, f, or both can alter the relative amounts of air that ventilate the alveoli and the dead space. These changes in V_T and f occur in animals during exercise and thermoregulation.

The anatomic dead space is important in thermoregulation (see Chapter 53). Air entering the respiratory system is usually cooler than body temperature and not saturated with water vapor. As air passes through conducting airways into the lung, it is warmed by transfer of heat from capillaries of the bronchial circulation, which encircle the conducting airways, and humidified to saturation by evaporation of water from the mucosal surfaces. When the animal exhales, heat is lost because the warmed and humidified air leaves the body. When some species, such as dogs, are heat-stressed, they pant. The small V_T and high f characteristics of panting cause more air to ventilate the dead space, resulting in increased water evaporation and heat loss. Cattle, pigs, and mules subjected to heat stress also increase respiratory rate and dead space ventilation in an attempt to lose heat. By the same principle, animals undergoing cold stress alter their pattern of ventilation to reduce dead space ventilation by increasing V_T and decreasing f in an attempt to retain heat. Furthermore, cold-stressed animals have a higher metabolic rate, which is necessary to maintain body temperature in cold conditions. This leads to an increase in both O_2 consumption and CO_2 production, making it necessary for cold-stressed animals to optimize efficiency and decrease heat loss by minimizing dead space ventilation, as well as increase their overall alveolar ventilation, and thus $\dot{V}_E$.

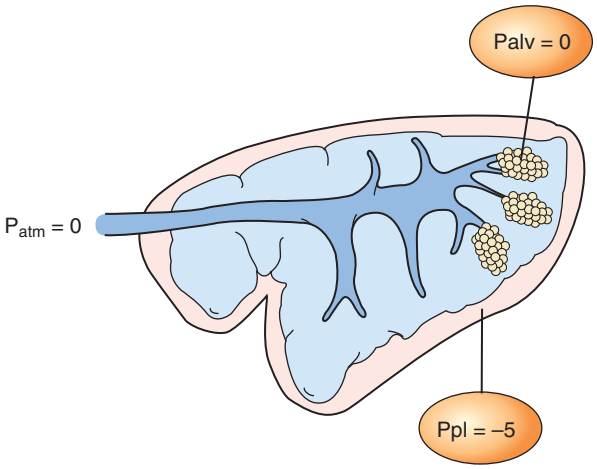
The veterinarian needs to ensure that equipment used for anesthesia or respiratory therapy does not increase the dead space. Excessively long endotracheal tubes or overly large face masks create a large amount of *equipment dead space,* which is the additional volume of air in a breathing circuit where both inspired and expired air go through a common path. The consequence of equipment dead space is that the animal must increase its V_T to obtain adequate alveolar ventilation.

Ventilation Requires Muscular Energy

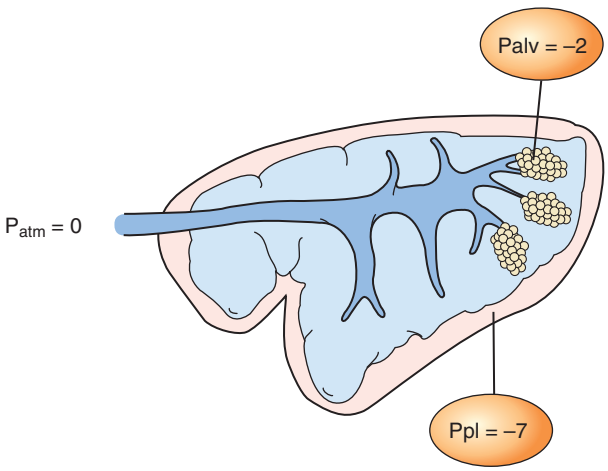
Inhalation occurs when the respiratory muscles contract to expand the thorax, stretch the lung, and create a subatmospheric pressure within the alveoli, which causes air to enter the respiratory system as it moves down the pressure gradient (Fig. 45.6). During *exhalation,* the elastic energy stored in the stretched lungs and thorax causes them to decrease in volume, leading to an increase in alveolar pressure (above atmospheric pressure) that drives air out of the respiratory system. Therefore in most resting mammals, exhalation is *passive* in that it does not require muscular effort. Horses are an exception because they have an active phase to exhalation, even at rest. In all species during exercise or in the presence of some respiratory diseases, exhalation has an active phase during which it is assisted by muscle contraction.



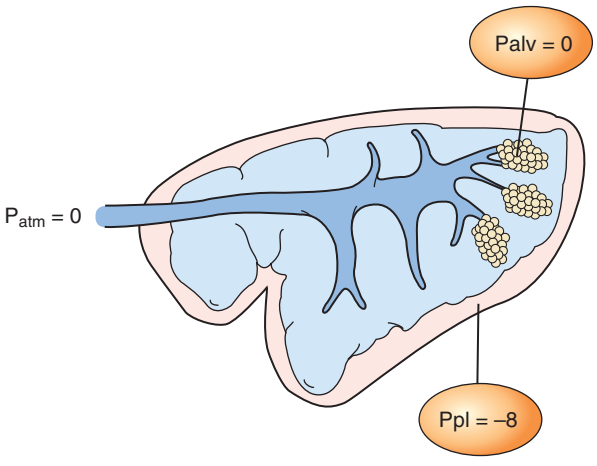
• **Fig. 45.6** **A,** Changes in volume and pressures during normal resting (tidal) breathing. **B–E,** Examples of pressure changes that might be expected in the respiratory system during different phases of quiet breathing. *Palv,* Alveolar pressure; *Patm,* atmospheric pressure; *Ppl,* pleural pressure; V_T, tidal volume. Numbers represent the pressure difference (cm H_2O) from atmospheric pressure. The gradient Palv – Ppl is the transpulmonary (or transmural) pressure gradient necessary to keep the lung inflated. Patm – Palv is the driving pressure gradient that drives airflow through the airways. Patm – Ppl is the pressure gradient that inflates the lung. **B,** Before the start of inhalation, when the respiratory system is resting at functional residual capacity (FRC), there is no airflow into the lungs because Patm – Palv = 0 cm H_2O and the negative pleural pressure is keeping the lung partially inflated (Patm – Ppl = 5 cm H_2O). **C,** During inhalation, Patm – Ppl = 7 cm H_2O to enlarge the lung (Palv – Ppl = 5 cm H_2O) and make air flow through the airways (Patm – Palv = 2 cm H_2O). **D,** At the end of a tidal inhalation, flow ceases because Patm – Palv = 0 cm H_2O, but the lung contains more air (Palv – Ppl = 8 cm H_2O). **E,** During exhalation, airflow reverses direction: Patm – Palv = –2 cm H_2O, the lung volume is still greater than FRC (Palv – Ppl = 8 cm H_2O), and the total pressure gradient remains positive (Patm – Ppl = 6 cm H_2O). It is important to remember that these pressure gradients change continually throughout a breath and with changes in tidal volume, respiratory frequency, lung compliance, and airway resistance. (Part A adapted from Hall JE: *Guyton and Hall textbook of medical physiology,* ed 12, Philadelphia, 2011, Saunders)



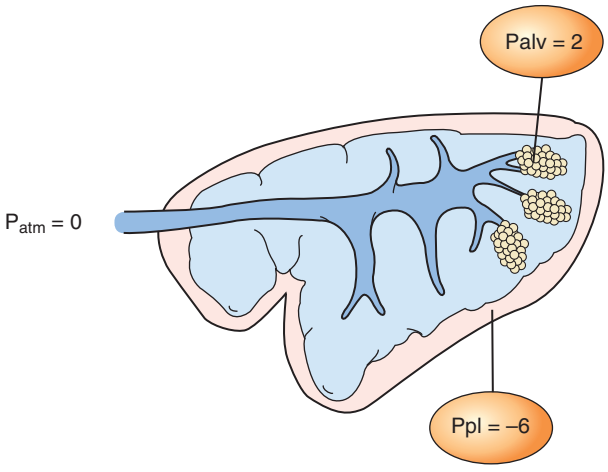
B. FRC



C. Inhalation FRC + 0.5V_T



D. FRC + V_T



E. Exhalation FRC + 0.5V_T

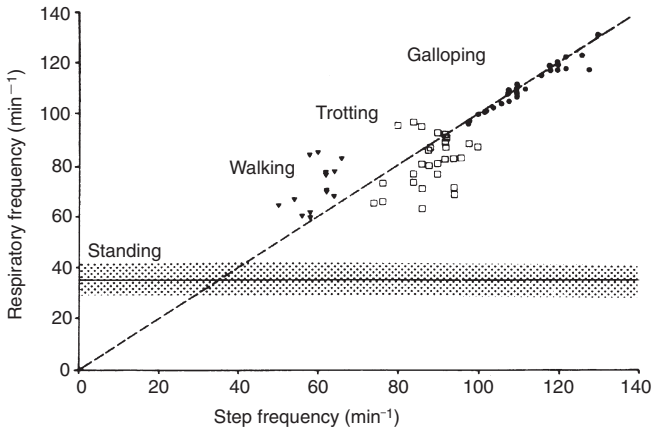
The most important inspiratory muscle is the *diaphragm,* which is a domed musculotendinous sheet separating the abdomen from the thorax and innervated by the *phrenic nerve.* During contraction of the diaphragm, the dome is pulled caudally and thereby enlarges the thoracic cavity; its tendinous center pushes against the abdominal contents, elevating intra-abdominal pressure, which displaces the abdominal wall and caudal ribs outward, thus also tending to enlarge the thorax. It is the enlargement of the thorax that creates the negative (subatmospheric) pressure within the airways necessary to make air flow down a pressure gradient into the lungs during inhalation.

The *external intercostal* muscles also are active during inhalation. The fibers of these muscles are directed caudoventrally, from the caudal border of one rib to the cranial border of the next, so that muscle contraction moves the ribs rostrally and outward. Because the cranial ribs support the forelimbs in quadrupeds, they participate less in ventilation than the more caudal ribs. Other inspiratory muscles, including those connecting the sternum to the head, contract during strenuous breathing and move the sternum rostrally to assist in thoracic enlargement.

The subatmospheric pressure generated within the respiratory tract during inhalation tends to draw the nonrigid structures of the external nares, pharynx, and larynx toward midline. Contraction of *abductor muscles* (which move structures away from the midline) attached to these structures is essential for preventing them from obstructing the airways. For example, *recurrent laryngeal neuropathy* (also known as *laryngeal hemiplegia*) in horses is a condition in which axonopathy of the left recurrent laryngeal nerve results in atrophy of the unilateral laryngeal abductor muscle (cricoarytenoideus dorsalis). The left side of the larynx fails to abduct on inhalation, and affected animals become symptomatic when exercising as the left arytenoid cartilage and vocal fold impede airflow during inspiration. Along with limiting inspiratory airflow, the fluttering of the left vocal fold creates an abnormal breathing sound referred to as *roaring.* Similar pathology occurs in older large breed dogs with the condition *geriatric onset laryngeal paralysis polyneuropathy* (GOLPP); these dogs have bilateral laryngeal paralysis that limits airflow during inspiration and creates the sound of inspiratory *stridor* when the dogs are subject to light exercise or heat stress; in severe cases, symptoms occur at rest.

While expiration is normally *passive* (requires no muscle contraction) at rest, it is assisted by the *abdominal* and *internal intercostal muscles* during exercise and in some diseases. Contraction of the abdominal muscles (rectus abdominus, internal abdominal oblique, external abdominal oblique, and transversus abdominus) increases abdominal pressure, which forces the relaxed diaphragm forward and reduces the size of the thorax. The fibers of the internal intercostal muscles are directed cranioventrally, from the cranial border of one rib to the caudal border of the next cranial rib, so that their contraction decreases the size of the thorax by moving the ribs caudally and ventrally. As the thorax becomes smaller, the intrathoracic pressures increases, which increases the pressure in the alveoli and forces air out of the lungs.

During exercise, respiratory muscle activity increases to generate an increased $\dot{V}_E$. In cursorial (running) mammals, ventilation is synchronized with gait at the canter and gallop but not at the walk or trot (Fig. 45.7). Inhalation occurs as the forelimbs are extended and the hind limbs are accelerating the animal forward. Exhalation occurs when the forelimbs are in contact with the ground. In the galloping horse and perhaps in other galloping quadrupeds, much of the increase in the size of the thorax during inhalation is a consequence of elongation of the trunk as the spine extends, rather than an increase in the thoracic diameter.



• **Fig. 45.7** Relationship between gait and respiration in the horse. At the walk and trot, step and respiratory frequency are not correlated. At the gallop (and canter), respiratory and step frequency bear a 1 : 1 relationship. (Modified from Hörnicke H, Meixner R, Pollman U: *Equine exercise physiology,* Cambridge, UK, 1983, Granta Editions.)

Lung Elasticity Results From Tissue and Surface Tension Forces and Determines Airway Compliance

The nonrigid airways of the lungs are held open or distended because, under normal conditions, the pressure in the alveoli (Palv) is greater than the pressure surrounding the lung in the pleural space (Ppl) (Fig. 45.6). This relationship is described as the *transmural pressure* (or in this case *transpulmonary pressure*), which is the difference between the pressure inside a structure (Palv) minus the pressure outside the structure (Ppl). During maximal forced exhalation the Ppl can become positive, approaching the value for Palv, which effectively limits expiratory flow as the airways are no long held open by a transmural pressure difference.

Compliance (C) is the measure of the distensibility of an elastic structure and for the lung is calculated as the change in lung volume (ΔV) that occurs for a given change in transpulmonary pressure (ΔP), as follows:

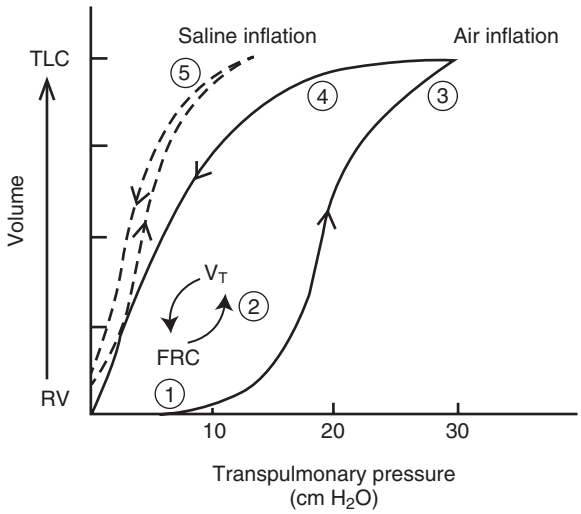
$$C = \Delta V / \Delta P$$

The normal lung is very compliant and thus readily increases volume with only modest changes in transpulmonary distending pressures, especially around resting tidal volumes (Figs. 45.6 *A, C, D,* and 45.8).

Materials that rebound (or recoil) to their original shape after having been stretched are demonstrating *elasticity.* In the lungs, the property of *elastance* (E) relates to the work required to inflate and deflate the lungs and is the reciprocal of compliance, in that it is the pressure required to elicit a volume change, as follows:

$$E = \Delta P / \Delta V$$

Lung compliance and elastance are fairly linear through the range of normal resting breathing volumes, but are less linear at the extremes of lung volume; this relationship between lung volumes and distending pressures is demonstrated by the *pressure-volume curve* (Fig. 45.8). The pressure-volume curve of the normal air-filled lung is characterized by a fairly steep slope in the middle of the curve at lung volumes near FRC, such that relatively small changes



• **Fig. 45.8** Pressure-volume curve of the normal lung (solid line) and under experimental conditions of inflation with saline (dotted line). Notice that (1) a greater change in pressure is required to inflate the normal lung from residual volume as compared to the change in inflation pressure required for normal tidal breathing (2; FRC ≒ V_T); (3) the lung reaches its elastic limits *(total lung capacity)* at transpulmonary pressure of approximately 30 cm H_2O; (4) the lung's elastic properties differ during inspiration and expiration such that the lung volume at any given pressure during inspiration is less than the lung volume at the same pressure during exhalation (a phenomenon known as *pressure-volume hysteresis*); (5) when saline is used instead of air to inflate the lung, less pressure is required for inflation, and the pressure-volume hysteresis is abolished. There is no air-liquid interface when the lung is inflated with saline, so the elastic recoil due to surface tension is abolished and the lung becomes easier to inflate. *FRC,* Functional residual capacity; *RV,* residual volume; *TLC,* total lung capacity; V_T, tidal volume.

in transpulmonary pressure result in sizeable changes in lung volume. At both low (RV) and high (TLC) lung volumes, the slope of the curve is much less steep, indicating that less change in lung volume occurs for a given pressure (or conversely, more pressure is required to elicit a given volume change). This is because lung compliance is reduced at low and high lung volumes.

At FRC, the slightly subatmospheric pressure in the pleural cavity keeps the lung inflated. If the thorax is opened and the negative pressure pulling the lungs open is lost, the lungs collapse, although some air remains trapped within the alveoli behind closed bronchioles. The collapse of the lung when the thorax is opened, as well as its recoil to a smaller volume during exhalation, are a result of the lung's inherent elasticity, which is generated by elastin fibers and by surface tension forces.

Elastin fibers form a woven mesh that extends along the length of the airways and within the interstitium of the alveolar septa. The distensible nature of this network is illustrated by the fact that a 10 kg dog's lungs containing approximately 100 mL of air at RV can expand to 450 mL at FRC and to 1100 mL at TLC. As the lung approaches TLC, its further inflation is limited by the collagen network of the pleural surface and also by the rib cage.

The *surface tension forces* contributing to the lung's elastic recoil arise from the air-liquid interface within the terminal airspaces (alveoli, alveolar sacs, alveolar ducts, and respiratory bronchioles). This surface tension arises from *cohesive forces* generated by hydrogen bonding between molecules in the *hypophase,* which is a thin layer of aqueous fluid that lines the alveoli. Cohesive forces at the surface of aqueous solutions (where liquid meets air) continually contract to assume the smallest surface area possible. Because the alveoli are lined with a thin layer of aqueous fluid, and the lungs are so compliant, the surface tension-driven propensity to minimize surface area at the air-liquid interface within alveoli produces a major portion of the lungs' elastic recoil. The importance of surface tension is demonstrated by the experiment comparing the pressures required to inflate excised lungs when they are filled with saline versus when they are filled with air (Fig. 45.8). Filling the lungs with saline abolishes surface tension because there is no longer an air-liquid interface; in this experimental condition, the pressure needed to inflate the lungs is greatly reduced because the elastic recoil of the lungs is decreased.

Surfactant Reduces Elastic Recoil of the Lungs and Thus Increases Compliance

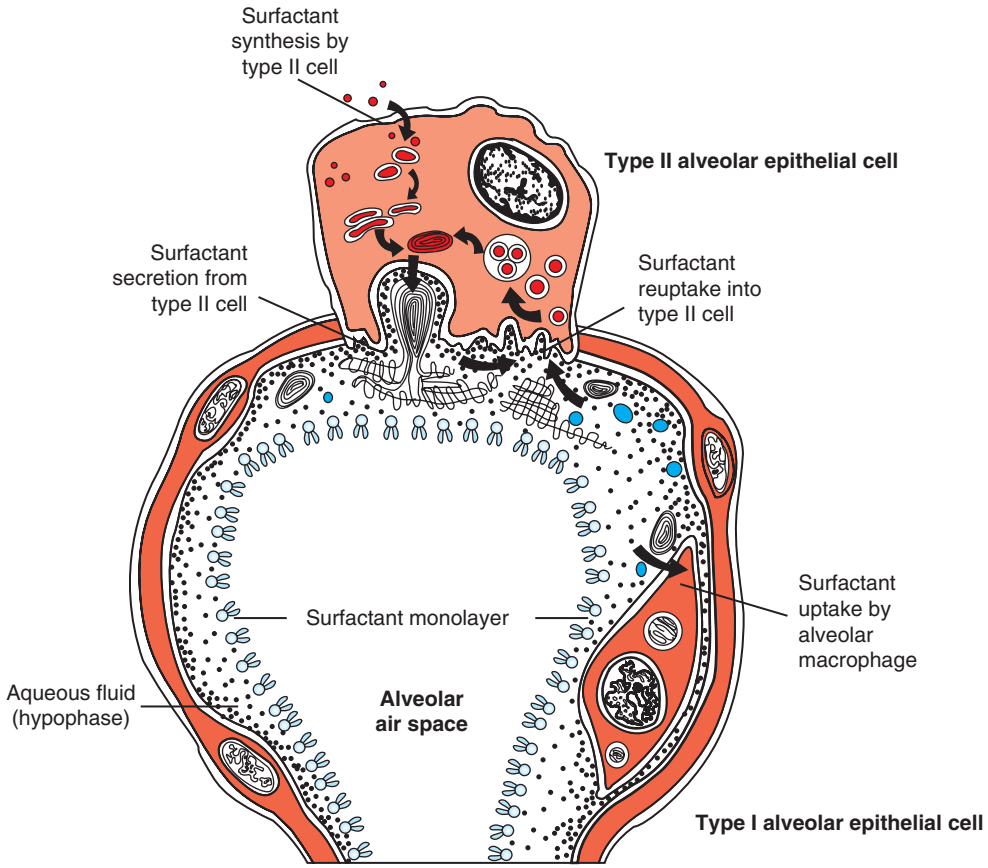
Lung compliance is the slope of the lung pressure-volume curve (Fig. 45.8). The pressure-volume curve is not linear, indicating that compliance varies with the state of lung inflation. Compliance is usually measured over the range of V_T and, when adjusted for differences in lung size, does not vary greatly among adult mammals. Consequently, most mammals generate similar changes in pleural pressure during breathing. Anesthesiologists must consider lung compliance when artificially ventilating an animal. A *compliant lung* inflates readily. A *lung with low compliance,* as occurs in some diseases, is more difficult to inflate. *Surfactant* is a critical factor underlying the high compliance of the normal lung.

If the air spaces were simply lined with water, the surface tension from the water's cohesive forces would be so great that the alveoli would collapse at the inflation pressures generated during breathing. Alveolar patency and stability is a consequence of the presence of *pulmonary surfactant* (Fig. 45.9). Surfactant is produced in *type II alveolar epithelial cells* and contains a mixture of lipids and proteins, with the most plentiful component being the lipid *dipalmitoyl phosphatidylcholine.* Surfactant has hydrophilic and hydrophobic portions that cause it to seek the surface of the alveolar lining fluid, where it displaces water molecules at the air-liquid interface and thus reduces the cohesive forces of the fluid lining the airways. In this manner surfactant reduces the surface tension of the alveolar lining fluid and increases lung compliance. As lung volume decreases and the alveolar surface area shrinks, surfactant molecules become more concentrated on the alveolar surface, further reducing surface tension and promoting alveolar stability.

Premature newborns may have surfactant deficiency as surfactant is released into the alveolar spaces and tracheal fluid late in gestation. Its appearance correlates with the rise in fetal plasma cortisol levels. Animals born prematurely have difficulty inflating their lungs because of inadequate surfactant. Synthetic surfactants can be used to treat premature newborns who lack adequate surfactant.

Postnatally and throughout life, release of surfactant from type II alveolar epithelial cells is aided by deep inspirations, such as sighing, which also redistribute surfactant over the alveolar surface from stores in the hypophase. Anesthetized animals and those with chest pain may not sigh and consequently some of the alveoli collapse, termed *atelectasis.* Provision of deep breaths with a ventilator or bag valve mask (Ambu bag) helps to maintain surfactant release and distribution, prevent lung collapse, and maintain normal lung compliance.

Surfactant also has a role in defense against pathogens as it aids the beating of cilia, thus supporting the mucociliary clearance functions. Surfactant proteins also support the innate and adaptive immune systems in a number of ways, including opsonizing



• **Fig. 45.9** Diagram of an alveolus showing the synthesis, release, distribution and uptake of surfactant. Note the stores of surfactant in the hypophase (*swirls* and *concentric circles* below the surfactant monolayer).

pathogens, facilitating their uptake by alveolar macrophages, and recruitment of other inflammatory cells by activating the complement pathway (see Chapters 54 and 55).

The Lung is Functionally Coupled to the Chest Wall

The lung is covered by the *visceral pleura* and the thorax is lined by the *parietal pleura.* These two pleural surfaces, while separated by a thin layer of pleural fluid, are maintained in close apposition by *adhesive forces* of the pleural fluid. Thus the lungs are functionally coupled to the chest wall. There is little frictional resistance to the movements of breathing because of the lubricating properties of the pleural fluid. Furthermore, while they are less than millimeters apart, the two pleural surfaces don't quite touch because of repulsive forces between adjacent surfaces exerted by like charges on phospholipids adhered to these mesothelial surfaces.

The pleural space is closed to the environment and is under a slight negative pressure that is generated by the opposing elastic forces of the chest wall (recoiling outward) and the lung (recoiling inward) in the face of the pleural fluids' adhesive forces. Additionally, a small part of the pleural space negative pressure is the result of pleural fluid absorption via lymphatics. As a consequence of the functional coupling of the lungs to the thorax provided by the pleural fluid, the respiratory system behaves as a single unit. Therefore, when the thorax expands during inhalation the lungs must expand as well.

Fig. 45.6 shows that, at FRC, the *pressure in the pleural cavity* (Ppl) that surrounds the lungs is approximately 5 cm H_2O below atmospheric pressure ($^-$5 cm H_2O). During inhalation, as the inspiratory muscles contract, the thorax enlarges and Ppl becomes more negative. This decrease in Ppl, along with the adhesive forces coupling the visceral and parietal plural surfaces, causes the elastic lung to stretch and enlarge its volume, which decreases pressure within the alveoli (Palv). The decrease in Palv below atmospheric pressure (Patm) causes air to flow into the lung through the tracheobronchial tree.

By the same principles, during expiration as the inspiratory muscles relax, the thorax decreases in size as do the lungs within. This moves the pressure in the pleural space toward 0 and raises the Palv above Patm so that air flows out of the lung down its pressure gradient. Furthermore, when an animal actively exhales below FRC, the stiff thorax increasingly resists deformation, so that RV—the volume of air in the lung at the end of a maximal exhalation—is determined by the limits to which the rib cage can be compressed. The FRC—the volume of the lungs at the end of passive expiration—is also the balance point between the elastic tendency of the lungs to recoil inwards with the tendency of the thoracic wall to recoil outward.

The thorax is generally less compliant in large animals than small animals; the stiff chest wall of the horse and cow is in contrast to the very compliant chest wall of small rodents. Neonates need to have a compliant chest to pass through the birth canal. *Atelectasis,* which is the collapse of individual alveoli or entire segments of lung, is more likely to occur in species that have a compliant

thorax. This is one reason that atelectasis is more common in newborn animals than in adults.

Airflow is Opposed by Frictional Resistance

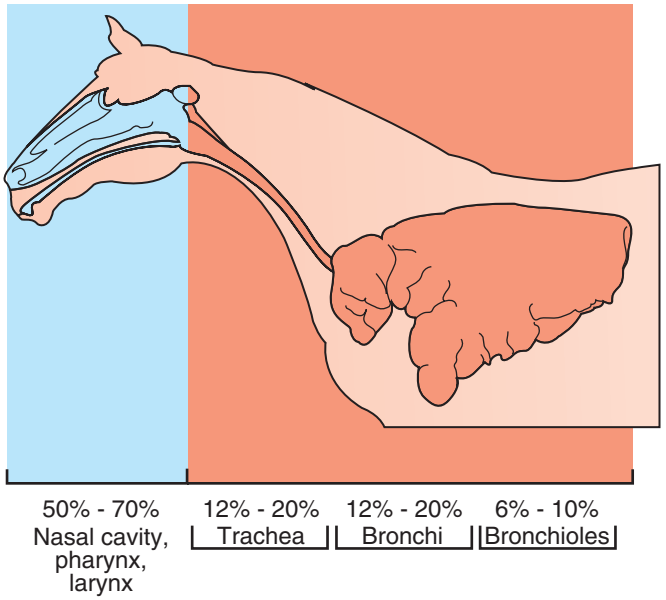
The force that impedes airflow along the respiratory passages is termed *airway resistance* (R). The magnitude of R determines the pressure drop (ΔP) along the airways (termed *driving pressure*) that is required to achieve the needed airflow through the respiratory tract, and can be calculated as follows:

$$R = \Delta P/\text{flow}$$

where ΔP = Patm − Palv.

Resistance to airflow is determined primarily by the *cross-sectional area* and the related measure, *radius,* of the airways, such that as the radius decreases the resistance increases by the fourth power (resistance is proportional to $1/\text{radius}^4$). However, the branching of the trachea into bronchi and bronchioles dictates that air flows through a large number of *parallel paths;* for this reason, it is the *total cross-sectional area* of these many parallel branches, rather than the caliber of the individual air passages that determines overall R in different segments of the respiratory system. Thus the total cross-sectional area of the airway increases greatly toward the periphery of the lung as the single trachea repeatedly branches to end in several million alveoli. While the caliber of each individual peripheral airway is small, because there is such a large total cross-sectional area across the parallel air passages generated by repeated branching, the resistance to airflow is lowest in the peripheral airways.

The structures of the *upper airways* are a major source of R. In the resting animal, the nasal cavity, pharynx, and larynx provide approximately 60% of the total R (Fig. 45.10). While individually these airways have the largest diameter of any single air passage, they precede any branching so air does not flow through them in parallel; this makes their total cross-sectional area relatively small. The upper airways must accommodate the full volume of each breath and this large flow through a tube of modest radius results in their high resistance relative to peripheral airways.



• **Fig. 45.10** Distribution of resistance in the conducting airways of a horse. The upper airways contribute the majority of resistance. Within the tracheobronchial tree, the bronchioles contribute the least fraction of resistance.

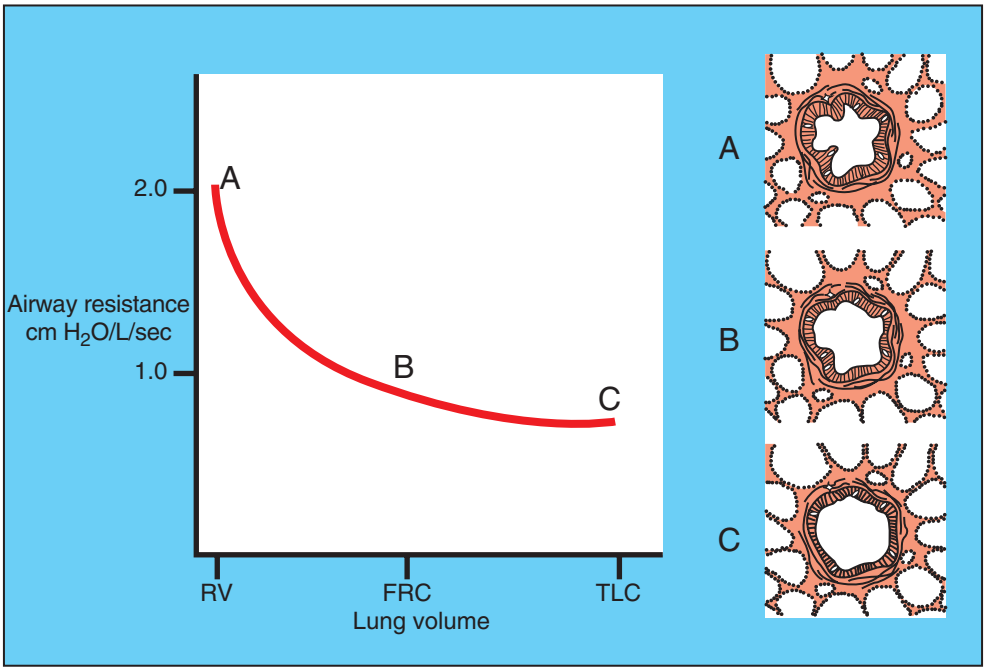
Nasal resistance can be decreased (e.g., during exercise) by dilation of the external nares and by vasoconstriction of the extensive vascular tissue in the nose. Vasoconstriction reduces the volume of blood in the vascular sinuses within the nasal mucosa and, as a consequence, the mucosal thickness decreases and the space available for air within the nose increases. When airflow rates increase during exercise or when the nasal cavity is obstructed, some species, such as dogs, breathe through the mouth to bypass the high-resistance nasal cavity. Other species, such as horses, are *obligate nose breathers* and are solely dependent on a decrease in nasal resistance to keep the work of breathing at a reasonable level. Horses accomplish this by flaring their nostrils and by constricting blood vessels to shrink the nasal mucosa.

The branching system of the tracheobronchial tree delivers air to the alveoli. The number of branches depends on the animal's size. Humans have 24 branches, mice range from more than 12 to 20, and horses have 40 or more. The trachea contains incomplete cartilaginous rings, whereas the bronchi are supported by cartilage plates. The bronchioles lack cartilage and are encircled by smooth muscle, thus they are particularly subject to narrowing when airway smooth muscle constricts.

The lungs of most species have a total of six lobes, each supplied by a *lobar bronchus,* which gives rise to daughter segmental bronchi. Even in species that lack lobation, such as the horse, the same pattern of six lobar bronchi persists. At each division of a parent bronchus, the diameters of the daughter airways are not equal; one is much narrower than the parent, whereas the diameter of the other is similar to that of the parent. This system of branching continues through at least the first six generations of bronchi. At the level of the bronchioles, the airways bifurcate giving rise to a pair of daughter bronchioles each of similar diameter to the parent structure. Thus along with the multigeneration branching pattern, the size of the daughter branches also contributes to the total cross-sectional area of the tracheobronchial tree expanding toward the periphery; the cross-sectional area increases only slightly between the trachea and the first four generations of bronchi, but it doubles at each division of the peripheral airways. Air flows down these many parallel branches, which makes the effective cross-sectional area of the peripheral airways very large, and thus resistance to airflow through this region is very small. As a result of the branching pattern of the tracheobronchial tree, airways larger than 2 to 5 mm in diameter contribute up to 90% of the resistance to breathing in the tracheobronchial tree; bronchioles and smaller airways contribute as little as 10% (Fig. 45.10).

Velocity is a measure of distance traveled per unit time and, in the respiratory system, is a function of the volume of air moving through the system and the cross-sectional area through which it moves. Because the total cross-sectional area increases dramatically toward the periphery of the lung, and because velocity is inversely proportional to the total cross-sectional area, the *velocity of airflow* diminishes progressively from the trachea toward the bronchioles. The high-velocity *turbulent airflow* in the trachea and bronchi produces the *lung sounds* heard through a stethoscope in a normal animal. *Laminar airflow* (low-velocity flow) in the bronchioles produces no sound.

The respiratory bronchioles, alveolar ducts, alveolar sacs, and alveoli are gas exchange structures present at the ends of the tracheobronchial tree. The surface area (cross-sectional radius) of these structures is very large and thus under normal conditions



• **Fig. 45.11** Effect of change in lung volume on airway resistance. The airway is represented in the diagrams on the right side of the figure by the *large circle,* to which are attached alveoli, the septa of which link the airway wall to the pleural surface. As lung volume increases, the alveolar septa become stretched, apply tension to the airway walls, and thus dilate the airway and reduce resistance. *FRC,* Functional residual capacity; *RV,* residual volume; *TLC,* total lung capacity.

they present minimal resistance to airflow. Furthermore, as the lung inflates, these peripheral airways dilate passively and R decreases even further. This occurs because the alveolar septa are attached to the airways, and as the alveoli inflate, tension increases in their septa, pulling on the attached airways and causing them to dilate (Fig. 45.11).

Smooth Muscle Contraction Affects the Diameters of the Trachea, Bronchi, and Bronchioles

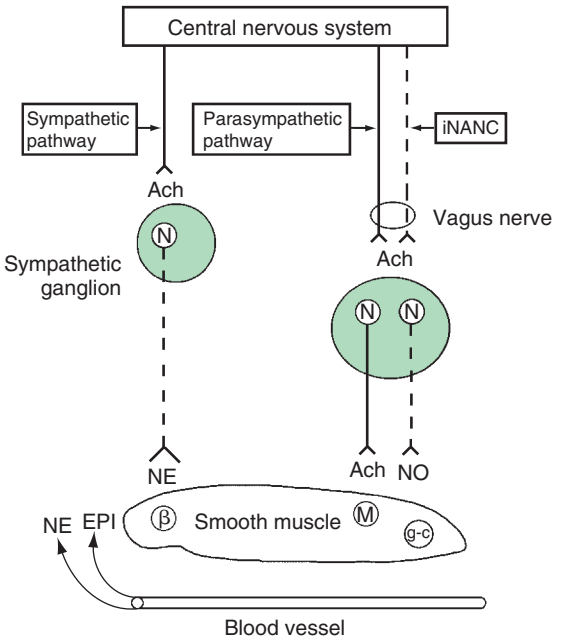
While the previously described structural characteristics determine the caliber of the larger airways, the caliber of the peripheral airways can be altered by both passive and active forces. Contraction of smooth muscle is a major factor determining intrapulmonary airway caliber. There is *smooth muscle* in the walls of the airways from the trachea to the alveolar ducts. In the trachea, it forms the *trachealis muscle,* which connects the ends of the C-shaped tracheal cartilages. In the bronchi and the bronchioles, smooth muscle encircles the airways. Smooth muscle actively regulates airway diameter in response to neural and other stimuli. The *parasympathetic nervous system* supplies airway smooth muscle through the vagus nerve, with parasympathetic ganglia located in the walls of the airways (Fig. 45.12). Activation of this system causes the release of *acetylcholine* from postganglionic fibers. The acetylcholine binds to *muscarinic receptors* on airway smooth muscle, leading to contraction. This narrows the trachea, bronchi, and bronchioles, a phenomenon known as *bronchoconstriction.* Parasympathetically mediated bronchoconstriction is one of the lung's protective mechanisms. Inhalation of irritating materials, such as dusts, can activate tracheobronchial sensory receptors connected to vagal afferent nerves. This, in turn, leads to activation of the parasympathetic system, resulting in bronchoconstriction.

Airway smooth muscle also contracts in response to many *inflammatory mediators,* particularly *histamine* and some *leukotrienes,* which are released from mast cells during an allergic reaction. Some inflammatory mediators act directly on the smooth muscle; others act in a reflexive manner involving *nonmyelinated afferent nerves* and the parasympathetic system. They are most likely responsible for the airway obstruction that occurs in diseases such as recurrent airway obstruction (a.k.a., heaves) in horses and asthma in cats.

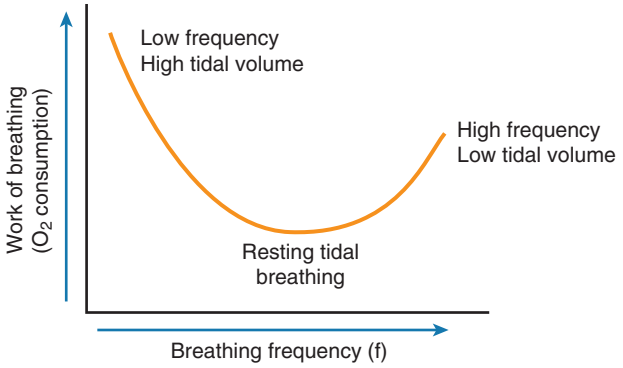
Relaxation of smooth muscle, and therefore dilation of the airways, occurs during activation of *β_2-adrenergic receptors* by circulating *epinephrine* released from the adrenal medulla. *Norepinephrine* released from the *sympathetic nervous system* also causes airway dilation through β_2-adrenergic receptors but to a lesser extent. Another bronchodilator system, the *inhibitory nonadrenergic noncholinergic (iNANC) nervous system,* exists in some species; the iNANC efferent fibers are in the vagus nerve and neurotransmission involves *nitric oxide.*

The Work of Breathing is Increased by Exercise, Thermoregulation, and Disease

The metabolic demands of breathing normally require only a small energy cost. Resting animals take relatively few breaths per minute and have low airflow rates (Fig. 45.13). The primary work of the respiratory muscles at rest is to overcome the inward elastic recoil of the lungs. When heat-stressed animals increase their breathing frequency to dissipate heat, the increased dead space ventilation (wasted ventilation) decreases breathing efficiency and thereby increases the work of maintaining adequate alveolar ventilation. During exercise, both respiratory rate and V_T increase in response to the need for increased alveolar ventilation. Respiratory muscles must therefore work harder to provide the necessary change in Ppl to generate the



• **Fig. 45.12** Diagrammatic representation of efferent autonomic innervation of the tracheobronchial tree. Muscarinic receptors (*M*) are activated by acetylcholine (*Ach*) released from postganglionic parasympathetic nerve terminals. Smooth muscle β_2-adrenergic receptors (β) are activated by circulating catecholamines such as epinephrine (*EPI*) or, in a few species, by release of norepinephrine (*NE*) from sympathetic nerves. The inhibitory nonadrenergic noncholinergic (*iNANC*) nervous system that travels in the vagus nerve releases nitric oxide (*NO*), which activates guanylyl cyclase (*g-c*) in the smooth muscle. *N,* Nicotinic receptors. (From Nadel JA, Barnes PJ, Holtzman MJ: Autonomic factors in hyperreactivity of airway smooth muscle. In Fishman AP, Macklem PT, Mead J, et al, editors: *Handbook of physiology,* section 3, vol 3, part 2, Bethesda, MD, 1985, American Physiology Society.)



• **Fig. 45.13** The work of breathing is minimized by moderating both tidal volume and breathing frequency to achieve the required minute ventilation. When high frequency and low tidal volumes are used to achieve minute ventilation, dead space ventilation increases and thus breathing efficiency decreases. Additionally, the reduced compliance at low lung volumes increases the work of breathing. Conversely, more energy is required when low frequency and large tidal volumes are used to achieve the same minute ventilation because lung compliance is reduced as volumes approach maximal (total lung capacity) limits.

greater airflow. Lung diseases often decrease compliance or increase R or both. As a consequence, animals with lung disease must do more work with their respiratory muscles to ventilate their lungs. In some cases, the work of breathing is substantial and becomes too high to maintain normal alveolar ventilation; in these cases, alveolar ventilation is insufficient to eliminate the CO_2 produced by metabolism. Insufficient alveolar ventilation that results in increased CO_2 retention is termed *hypoventilation* (see Chapter 47). When the energy cost of breathing is increased by disease or thermal stress, this results in less energy available for exercise or weight gain, which can result in poor performance or failure to thrive.

Dynamic Compression Can Narrow the Airways and Limit Airflow

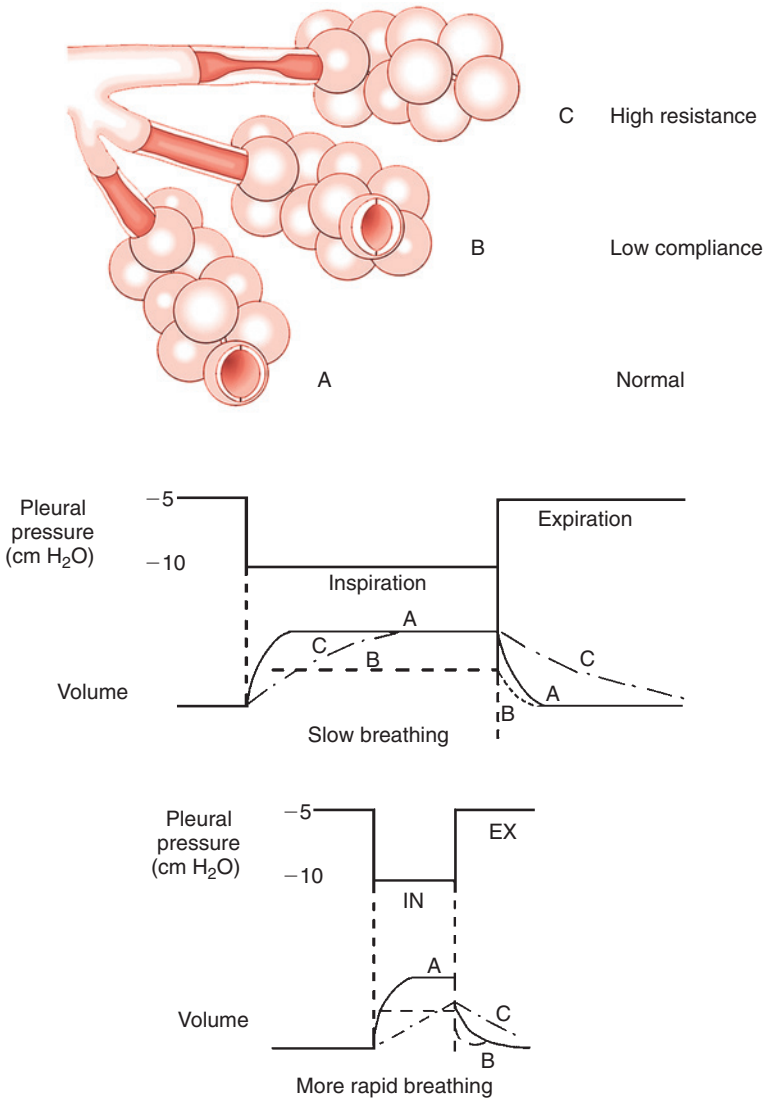
The walls of the airways are not rigid and therefore the airways can be compressed or expanded by the pressure gradient across their walls. Understanding when such *dynamic compression* is most likely to occur in different parts of the airways can provide diagnostic clues to the location of an airway obstruction. In the nasal cavity, pharynx, and larynx, dynamic compression of the airway occurs during inhalation. These extrathoracic airways are surrounded by atmospheric pressure, whereas the pressure within the airways is subatmospheric during inhalation; the resulting *negative transmural pressure* therefore tends to make the *extrathoracic airways collapse* during inhalation. Because of its bony support, the nasal cavity is not prone to compression, but the less well-supported nares, pharynx, and larynx are prone to compression or collapse. Normally, contraction of the abductor muscles of the nares, pharynx, and larynx during inhalation maintains patency and prevents collapse of these regions.

Recurrent laryngeal neuropathy in horses provides an excellent example of dynamic collapse of the extrathoracic airway during inhalation. In this disease, the intrinsic muscles on the left side of the larynx lose their nerve supply and undergo atrophy. As previously mentioned, when the abductor muscles of the larynx fail to contract, the left vocal fold is sucked into the lumen of the airway during inhalation, producing an inspiratory noise and obstructing airflow, which leads to poor performance when affected horses undergo strenuous exercise. Dynamic collapse of the extrathoracic airway does not occur during exhalation because the pressure within the airway is greater than atmospheric pressure and the resulting positive transmural pressure keeps the airway open.

In the intrathoracic airways, dynamic collapse may occur during forced exhalation as intrapleural pressure approaches pressures within the intrathoracic airway lumen. *Cough* is a *forced exhalation* during which dynamic collapse narrows the airways. The high air velocity through the narrowed portion of the airway facilitates removal of foreign material. Toy breeds of dogs have a high incidence of *collapsing trachea.* In this disease, the intrathoracic trachea is weakened and thus dynamically collapses during the forceful ventilation of exercise. Affected dogs make a "honking" expiratory noise as air is forced past the collapsed intrathoracic portion of the trachea.

The Distribution of Ventilation Depends on the Local Mechanical Properties of the Lung

Optimal gas exchange between inhaled air and capillary blood requires bringing together air and blood at the alveolus, that is, the *matching of ventilation* and *blood flow* (also termed *perfusion, Q*). Obviously, gas exchange cannot occur if an alveolus receives blood but no ventilation, or vice versa. Ideally, each region of lung would receive approximately equal amounts of ventilation, however, distribution of ventilation is always uneven to some degree and becomes more so in disease. Uneven distribution of ventilation can be caused by local decreases in lung compliance (e.g., in



• **Fig. 45.14** Effects of mechanical properties of the lung on alveolar filling. Alveolus *A* is normal, alveolus *B* has low compliance due to its thickened walls, and the airway supplying alveolus *C* has high resistance as a result of a partial obstruction. Step changes in pleural pressure are applied to these three schematic alveoli, and the changes in volume are shown during slow breathing and during more rapid breathing. During slow breathing alveoli *A* and *C* fill to equal degrees because their compliance is the same but alveolus *C* fills more slowly because its airway is partially obstructed; alveolus *B* with low lung compliance fills less. During rapid breathing, alveolus *C* does not have time to fill, it receives less ventilation than *A* and as a consequence ventilation becomes more unevenly distributed.

pneumonia) or local airway obstructions (e.g., by mucus, bronchoconstriction, foreign body obstruction) (Fig. 45.14).

The uneven distribution of ventilation becomes magnified in recumbent large animals, especially in the supine and laterally recumbent positions. This is because the lowermost regions of the lung are compressed to such an extent that they receive little or no ventilation. This can cause severe derangements of gas exchange, especially in anesthetized large animals, such as horses.

In Some Species, Air Travels Between Adjacent Regions of Lung Through Collateral Pathways

The lungs of mammalian species differ in the degree to which they are subdivided by connective tissue into *secondary lobules.* In the lungs of pigs and cattle, there is complete separation of lobules, and in dogs and cats, there is no separation. In horses and sheep, there is partial separation. The connective tissue septa prevent *collateral ventilation* (i.e., movement of air between adjacent lobules) in cattle and pigs. Collateral ventilation is extensive in dogs and intermediate in horses. Collateral ventilation provides air to alveoli when their main parent bronchus is obstructed. The differences in collateral ventilation mean that gas exchange abnormalities that follow airway obstruction are more serious in pigs and cattle than in dogs.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Lung Fibrosis in a Dog

History

A 3-year-old English Setter in respiratory distress is presented at a teaching hospital. The owner first noticed reluctance of the dog to exercise 3 weeks ago. Since then, the animal has had progressive difficulty in breathing. It appears hungry but cannot eat because it "gets out of breath."

Clinical Examination

Inspection reveals a thin dog breathing through its mouth. The respiratory rate is elevated, but the dog seems to be moving little air despite strong inspiratory efforts during which the intercostal spaces sink. Exhalation presents no difficulty; the ribs collapse rapidly and there is no accentuated abdominal effort. Examination reveals slightly blue-colored mucous membranes. Lung sounds are not remarkable. All other systems are normal.

Radiographs of the thorax show diffuse miliary density (whiteness) over the parts of the lung that are normally air-filled. The bronchi are normal. An elevated change in pleural pressure during breathing, normal R, and decreased lung compliance are the key findings on lung function testing. V_T is greatly reduced.

Comment

The history and clinical signs indicate a respiratory problem. The elevated change in pleural pressure during breathing confirms the increased effort necessary to breathe. This could be caused by (1) increased air movement resulting from an increased metabolic rate, (2) airway obstruction, or (3) a decrease in lung compliance (stiffening of the lung). The increased density in the normally air-filled elastic part of the lungs, coupled with the normal air passages, suggests a decrease in lung compliance rather than airway obstruction. This was confirmed when measurements of lung function revealed a normal R and decreased lung compliance.

The retraction of intercostal spaces indicates that the stiff lung is resisting expansion. Exhalation is not a problem because the lung has an increased tendency to collapse, and the airways are normal.

Treatment

This dog has pulmonary fibrosis, which is a diffuse disease of the gas exchange area of the lung. It is characterized by decreased compliance, which increases the work of breathing. The blue tinge to the mucous membranes indicates increased desaturated hemoglobin as a result of impaired O_2 exchange in the diseased lung. A biopsy reveals diffuse fibrosis around mineral particles in the walls of the alveoli. The prognosis for the dog is poor.

Chronic Airway Disease in a Horse

History

A 10-year-old Arabian mare is presented with a 2-year history of coughing and a progressive loss of exercise tolerance. Recently, the horse's problem has become so severe that it has difficulty breathing while resting in its stall. The cough is frequent and is usually worse when the horse is kept inside. The horse has a normal appetite; however, it is losing weight even though the teeth are normal and it is on a good diet and parasite control program.

Clinical Examination

Inspection reveals a thin horse with flared nostrils and an anxious expression. The respiratory rate is elevated and respiratory movements are accentuated. During inhalation the intercostal spaces are pulled in between the ribs. The initial part of exhalation is characterized by a rapid relaxation of the rib cage. This is followed by a prolonged contraction of the abdominal muscles, which is terminated immediately before the next inhalation. During the prolonged contraction of the abdominal muscles, wheezes can be heard when you place your ear close to the nostrils.

The horse has an elevated pulse rate. The mucous membranes of the gums have a bluish tinge. Auscultation of the thorax reveals increased breath sounds over all the lung fields and musical wheezes audible at end exhalation. Excessive mucus pooled in the airway can be seen through an endoscope advanced into the trachea.

Because the horse is being examined at a teaching hospital, there are facilities for measurement of lung function. The change in pleural pressure (ΔPpl) during each breath is 25 cm H_2O (normal, 5 to 10 cm H_2O), and R is 3 cm H_2O/L/sec (normal, 1 cm H_2O/L/sec). Administration of atropine intravenously decreases ΔPpl to 7 cm H_2O and R to 1.5 cm H_2O/L/sec. The horse looks less distressed and wheezes are reduced after atropine therapy.

Comment

The respiratory distress, cough, and lack of exercise tolerance indicate a respiratory problem. The increased effort of breathing documented by the elevated ΔPpl could be caused by airway obstruction, a decrease in lung compliance, or increased breathing resulting from increased metabolic rate. The mucus in the airway and elevated R confirm airway obstruction. Musical wheezes at the end of exhalation result from increased air turbulence or vibration of mucus within the airway walls. Airway obstruction is caused in part by bronchospasm resulting from parasympathetic activity because it is reversed by atropine, a parasympathetic antagonist. Atropine does not return resistance to normal, so there is also considerable obstruction by mucus and swelling of the airway wall.

The flared nostrils are an effort to reduce the R by dilating the upper airway. The blue-tinged mucous membranes indicate increased desaturated hemoglobin because of impaired O_2 uptake in the diseased lungs.

Retraction of the intercostal muscles during inhalation indicates a major decrease in pleural pressure as the respiratory muscles work to inflate the lung and pull air through the obstructed airways. The prolonged contraction of abdominal muscles, or heaving, represents an effort by the horse to force air out through obstructed airways. Weight loss is probably a result of the increased work of breathing. Coughing is an effort by the horse to expel the excessive mucus.

Treatment

This horse has *recurrent airway obstruction* (RAO), (previously also known as *heaves*), a problem exacerbated by stabling in a dusty barn and eating poorly cured, moldy hay. RAO is an inflammatory response to inhaling particles, antigens, and endotoxin in the hay and barn dust. The best treatment for the horse is to keep it outdoors at pasture and supplement its diet with pelleted feed rather than hay. In many cases, including this horse, additional treatments are needed when the horse endures a crisis. Treatment is aimed at dilating the airways (bronchodilator such as clenbuterol) and reducing inflammation (inhaled or systemic corticosteroids). Oxygen therapy is rarely necessary. With good management that involves reduction of exposure to hay dust, some horses do not need constant treatment. In advanced stages and when treating some performance horses, however, ongoing treatment with bronchodilators and inhaled steroids may be necessary.

Practice Questions

1. Which of the following is true?
- **a.** O_2 consumption per kilogram body weight is greater in a 50-g mammal than in a 50-kg mammal.
- **b.** Maximal O_2 consumption in mammals is directly related to the volume of mitochondria in the skeletal muscles.
- **c.** O_2 consumption increases when metabolic rate increases.
- **d.** O_2 consumption can increase up to thirtyfold during intense exercise.
- **e.** All of the above are true.

2. Functional residual capacity is:
- **a.** The volume of air remaining in the lung at the end of a maximal forced exhalation.
- **b.** The mechanical equilibrium of the respiratory system.
- **c.** Less than residual volume.
- **d.** Greater than total lung capacity.
- **e.** Determined by metabolic rate.

3. Which of the following lists includes only structures that compose the anatomic dead space?
- **a.** Respiratory bronchioles, alveoli, trachea, nasal cavity
- **b.** Pharynx, bronchi, alveolar ducts, larynx
- **c.** Capillaries, respiratory bronchioles, trachea, bronchi
- **d.** Pharynx, nasal cavity, trachea, bronchi
- **e.** Capillaries, respiratory bronchioles, alveolar ducts, alveoli

4. A horse has a tidal volume (V_T) of 5 L, respiratory rate of 12 breaths/min, and V_D/V_T ratio of 0.5. Calculate minute ventilation ($\dot{V}_E$) and alveolar ventilation ($\dot{V}_A$).
- **a.** $\dot{V}_E$ = 60 L/min; $\dot{V}_A$ = 2.5 L/min
- **b.** $\dot{V}_E$ = 30 L/min; $\dot{V}_A$ = 30 L/min
- **c.** $\dot{V}_E$ = 60 L/min; $\dot{V}_A$ = 30 L/min
- **d.** $\dot{V}_E$ = 2.5 L/min; $\dot{V}_A$ = 1.25 L/min
- **e.** $\dot{V}_E$ = 5.0 L/min; $\dot{V}_A$ = 2.5 L/min

5. Which of the following occur during inhalation?
- **a.** Diaphragm contracts, pleural pressure increases, alveolar pressure decreases.
- **b.** Diaphragm relaxes, external intercostal muscles contract, pleural pressure increases.
- **c.** Diaphragm relaxes, pleural pressure decreases, internal intercostal muscles relax.
- **d.** External and internal intercostal muscles contract, pleural and alveolar pressures increase.
- **e.** Diaphragm and external intercostal muscles contract, pleural and alveolar pressures decrease.

6. Lung compliance:
- **a.** Has the units of pressure per volume (cm H_2O/L).
- **b.** Is greater at functional residual capacity (FRC) than at total lung capacity (TLC).
- **c.** Is less when the lung is inflated with saline than when the lung is inflated with air.
- **d.** Is greater in small mammals than in large mammals, even when adjusted for differences in lung size.
- **e.** Is the only determinant of the change in pleural pressure during breathing.

7. Pulmonary surfactant:
- **a.** Can be deficient in premature newborns.
- **b.** Is produced in type II alveolar epithelial cells.
- **c.** Is in part composed of dipalmitoyl phosphatidylcholine.
- **d.** Decreases surface tension of the fluid lining the alveoli.
- **e.** All the above.

8. Which of the following *increases* the frictional resistance to breathing?
- **a.** Intravenous administration of a β_2-adrenergic agonist
- **b.** Contraction of the abductor muscles of the larynx
- **c.** A decrease in lung volume from FRC to residual volume
- **d.** Relaxation of the trachealis muscle
- **e.** Inhibition of the release of histamine from mast cells

9. The distribution of ventilation within the lung is influenced by:
- **a.** Regional variations in lung inflation.
- **b.** Regional variations in airway resistance.
- **c.** Regional variations in lung compliance.
- **d.** Collateral ventilation.
- **e.** All the above.

Bibliography

Boron WF. Mechanics of respiration. In: Boron WF, Boulpaep EL, eds. *Medical Physiology: A Cellular and Molecular Approach*. 2nd ed. Philadelphia: Saunders; 2012.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2016.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lekeux P, Art T. The respiratory system: anatomy, physiology and adaptations to exercise and training. In: Hodgson DR, Rose RJ, eds. *The Athletic Horse: Principles and Practice of Equine Sports Medicine*. 2nd ed. Philadelphia: Saunders; 2013.

Notter RH. Lung surfactants: basic science and clinical applications. In: Lenfant C, eds. *Lung Biology in Health and Disease*. Vol. 149. New York: Marcel Dekker; 2000.

Orgeig S, Hiemstra PS, Veldhuizen EJ, et al. Recent advances in alveolar biology: evolution and function of alveolar proteins. *Respir Physiol Neurobiol*. 2010;173S:S43–S54.

Robinson NE. Some functional consequences of species differences in lung anatomy. *Adv Vet Sci Comp Med*. 1982;26:1–33.

Weibel ER, Bacigalupe LD, Schmitt B, et al. Allometric scaling of maximal metabolic rate in mammals: muscle aerobic capacity as determinant factor. *Respir Physiol Neurobiol*. 2004;140(2):115–132.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

West JB, Luks AM. *West's Pulmonary Pathophysiology: The Essentials*. 9th ed. Philadelphia: Wolters Kluwer; 2017.

46

Blood Flow Through the Lungs

SUSAN L. EWART

KEY POINTS

Pulmonary Circulation

1. Pulmonary arteries have thinner walls and less muscle than systemic arteries.
2. Pulmonary blood vessels can be classified as alveolar and extra-alveolar vessels.
3. Passive changes in vascular resistance result from changes in vascular transmural pressure.
4. Pulmonary blood vessels offer a low resistance to flow.
5. The distribution of pulmonary blood flow within the lung differs between quadrupeds and bipeds.
6. Neural and humoral factors cause contraction of pulmonary artery smooth muscle.
7. Alveolar hypoxia is a potent constrictor of small pulmonary arteries.
8. The pulmonary circulation accommodates a large increase in blood flow during exercise.
9. Pulmonary microvasculature filters small particles from the blood.

Bronchial Circulation

1. The bronchial circulation provides a blood supply to airways and large vessels and aids in conditioning inhaled air.

The lung receives blood flow from two circulatory systems: the *pulmonary circulation* and the *bronchial circulation,* which is a branch of the systemic circulation. There are several important differences between the pulmonary and systemic circulations. First, the *pulmonary circulation* receives the total output of the right ventricle, perfuses the alveolar capillaries, and participates in gas exchange. The *bronchial circulation* provides a nutritional blood supply to airways and other structures within the lung. Second, the *pulmonary arteries carry deoxygenated blood* from the right ventricle to the pulmonary capillaries, where it becomes oxygenated and this *oxygenated blood is returned to the left side of the heart in pulmonary veins.* In contrast, in the systemic (including bronchial) circulation arteries carry oxygenated blood and veins carry deoxygenated blood (Fig. 46.1). Next, the entire cardiac output passes through the pulmonary circulation, and thus a single capillary bed, whereas the systemic circulation has a number of different capillary beds, one for each end organ. Additionally, the pressures within the pulmonary circulation are much lower than the systemic blood pressures. Finally, the vascular response to tissue hypoxia differs between the pulmonary (vasoconstriction) and systemic (vasodilation) circulations. The mechanisms for and consequences of these differences are described in detail later.

Pulmonary Circulation

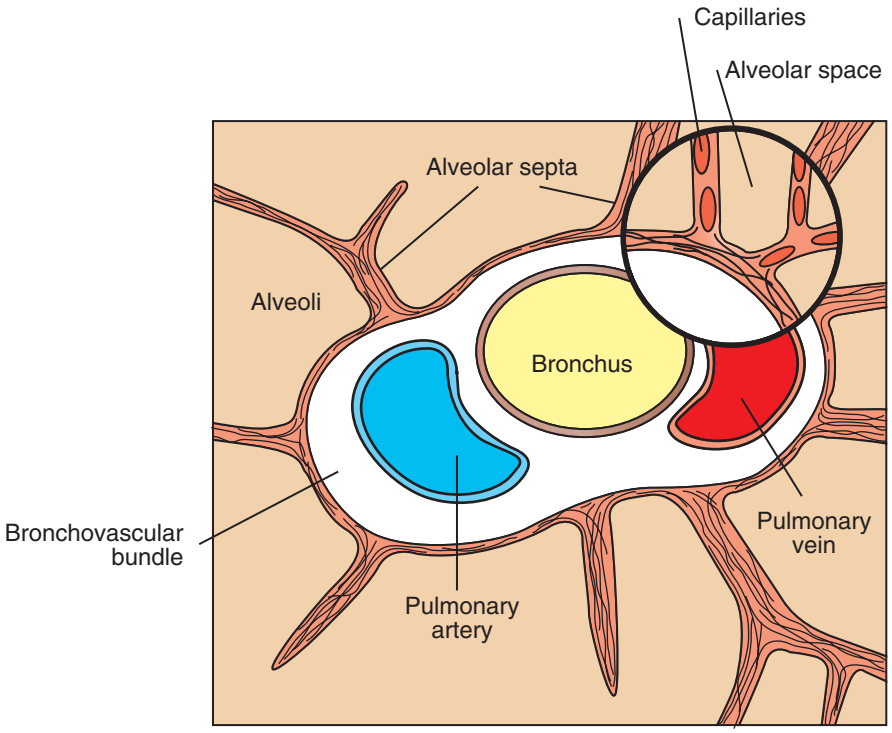
The total cardiac output of the right side of the heart passes through the lungs by means of the single vascular bed of the pulmonary circulation. This vascular bed holds approximately 10% to 20% of the total blood volume. Thus while the *rate* of blood flow through the pulmonary circulation approximates that of the systemic circulation, the blood *volume* in the pulmonary circulation is substantially less than that in the systemic circulation. When cardiac output increases, as occurs during exercise, the pulmonary circulation is able to accommodate this increase in blood flow without a large increase in the work of the right ventricle. In addition, control mechanisms exist to regulate the distribution of blood within the lung so that blood preferentially perfuses the well-oxygenated regions of the lung. The ability to regulate blood flow depends on the presence of smooth muscle in the walls of small pulmonary arteries. The quantity of smooth muscle varies among species.

Pulmonary Arteries Have Thinner Walls and Less Muscle Than Systemic Arteries

Pulmonary arteries carry deoxygenated blood from the right ventricle to the pulmonary capillaries, which surround the alveoli. The arteries and arterioles of the pulmonary system have thinner walls and less smooth muscle than systemic arteries of comparable diameter. This results in their lower basal (resting) tone and thus lowers blood pressures of the pulmonary circulation. The main pulmonary arteries that accompany the bronchi are elastic, and the smaller arterioles adjacent to the bronchioles and the alveolar ducts are muscular but are still thinner than systemic arteries and arterioles. Pigs and cattle have a thicker muscle layer in the smaller pulmonary arteries; horses have less muscle; and sheep and dogs have only a thin muscle layer. These species differences have clinical relevance as the amount of smooth muscle in the walls of small pulmonary arteries determines the *reactivity* of the vasculature to alveolar hypoxia and other neural and humoral stimuli.

The small pulmonary arterioles lead into pulmonary capillaries, which form an extensive branching network of vessels within the *alveolar septa,* essentially blanketing the alveolar surface and thus maximizing the surface area for gas exchange (see Chapter 47). However, not all pulmonary capillaries are perfused with blood in the resting animal. As a result, vessels that are unperfused in an animal at rest can be recruited when pulmonary blood flow increases (e.g., during exercise).

Pulmonary veins have thin walls and conduct *oxygenated blood* from pulmonary capillaries to the left atrium. Blood within pulmonary veins serves as a reservoir of blood for the left heart; this



• **Fig. 46.1** Diagrammatic representation of the extra-alveolar vessels (pulmonary artery and vein) in the bronchovascular bundle and (*inset*) an enlargement of the alveolar vessels (capillaries) in the alveolar septum. The alveolar septa are attached to the bronchovascular bundle so that they exert a radial traction on the bundle. Notice that the pulmonary artery carries deoxygenated (*blue*) blood and the pulmonary vein carries oxygenated (*red*) blood.

reservoir is available for sudden increases in cardiac output (e.g., at the start of a sudden burst of exercise).

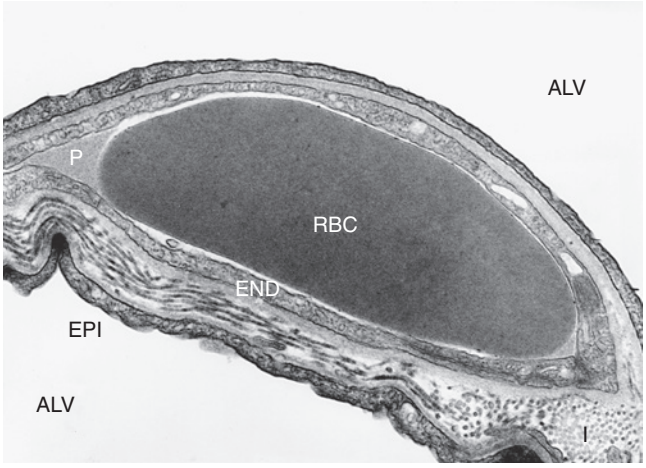
Pulmonary Blood Vessels Can Be Classified as Alveolar and Extra-Alveolar Vessels

Thin-walled *alveolar capillaries* perfuse the alveolar septa (Fig. 46.2). They are exposed almost directly to the cyclic pressure changes that occur in the alveoli during breathing. At low lung volumes, alveolar vessels are exposed to decreased pressures and have lower resistance, whereas at higher lung volumes, alveolar vascular resistance increases due to increasing alveolar pressures.

Extra-alveolar vessels include the pulmonary arteries and veins, which occur together with bronchi in a loose connective tissue sheath called the *bronchovascular bundle.* This bundle is bounded by a limiting membrane to which alveolar septa are attached (Fig. 46.1). The behavior of extra-alveolar vessels is determined by pressure changes within the connective tissue space of the bronchovascular bundle, which approximate pleural pressure. During inspiration, there is a drop in intrapleural pressure, which expands the lungs and distends the bronchovascular bundle and the extra-alveolar vessels contained therein. The bronchovascular bundle is also the initial site of accumulation of edema fluid—termed *pulmonary interstitial edema*—when animals develop pulmonary edema (see Chapter 50).

Passive Changes in Vascular Resistance Result From Changes in Vascular Transmural Pressure

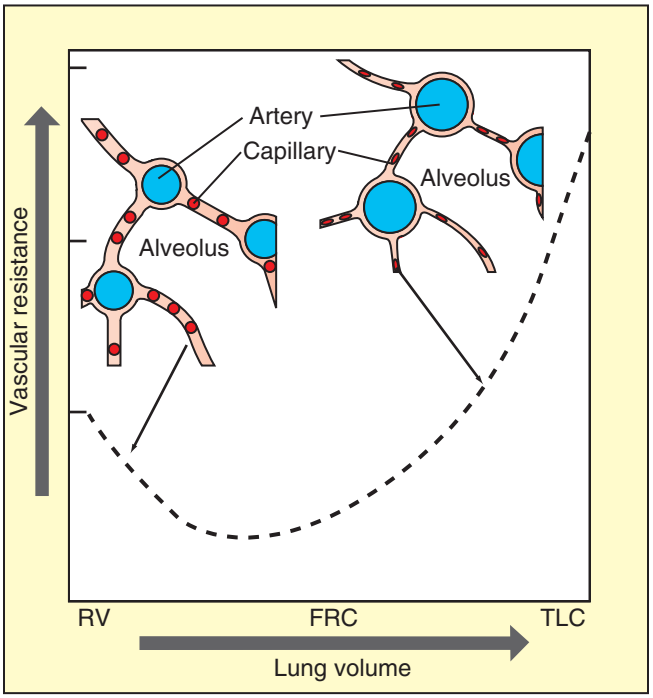
The diameter of blood vessels is a function of the pressure difference between the inside (Pin) and the outside (Pout) of the vessel, called the *transmural pressure* (P_{TM}).



• **Fig. 46.2** Transmission electron micrograph of capillary in the alveolar septum of a horse lung. A red blood cell (*RBC*) is shown bathed by plasma (*P*) in a capillary surrounded by endothelium (*END*). Alveoli (*ALV*) are on both sides of the septum and separated from the capillary by the epithelium (*EPI*) and a layer of interstitium (*I*). The interstitium is much thicker on one side of the capillary than on the other. Fluid exchange between the capillary and the interstitium occurs primarily on the thicker side. (Courtesy WS Tyler, Department of Anatomy, University of California–Davis.)

$$P_{TM} = Pin - Pout$$

Pressure within the vessels increases when the blood volume therein increases, as occurs during exercise. This leads to an increase in *transmural pressure,* which causes the vessels to dilate. Transmural pressure can also increase if the pressure surrounding the vessel



• **Fig. 46.3** Changes in pulmonary vascular resistance that occur across lung volumes. The diagram shows alveolar capillaries (*red*) and extra-alveolar vessels, in this case, arteries (*blue*). At residual volume (*RV*), the arteries are narrowed, but the capillaries are distended. At total lung capacity (*TLC*), the arteries are distended, but the capillaries are flattened because of the tension in the alveolar septa. Vascular resistance, which is the sum of the resistance provided by extra-alveolar vessels and capillaries, is minimal just below functional residual capacity (*FRC*).

decreases. This occurs in extra-alveolar pulmonary arteries and veins as the lung inflates. These vessels are contained in the bronchovascular bundle, which is enlarged by the traction of the surrounding alveolar septa during lung inflation. Consequently, pressure in the perivascular connective tissue of the bronchovascular bundle decreases. This leads to an increase in vascular transmural pressure, and the extra-alveolar arteries and veins therefore dilate. These dilated pulmonary vessels can accommodate more blood.

The overall effects of lung volume on pulmonary vascular resistance (PVR) reflect opposing effects on alveolar and extra-alveolar vessels (Fig. 46.3). At low volumes (e.g., residual volume, see Chapter 45), PVR is increased because extra-alveolar vessels are narrowed. As the lungs inflate toward functional residual capacity (FRC, see Chapter 45), resistance decreases, primarily because of dilation of extra-alveolar arteries and veins. Further inflation above FRC increases PVR, primarily because alveolar capillaries are flattened by the high tension in the stretched alveolar septa. As lung volume approaches total lung capacity (TLC, see Chapter 45), capillaries become progressively more elliptical and therefore offer more resistance to flow.

Pulmonary Blood Vessels Offer a Low Resistance to Flow

Even though the pulmonary circulation receives the total output of the right ventricle, pulmonary arterial pressures are much less than systemic pressures. In mammals, pulmonary arterial systolic, diastolic, and mean pressures average approximately 25, 10, and 15 mm Hg, respectively. These pressures are somewhat higher in larger animals than in smaller ones. *Pulmonary vascular pressures* can be measured by advancing a catheter through the jugular vein into the right atrium, then right ventricle and on into the pulmonary artery. If the catheter is advanced until it becomes wedged in a pulmonary artery, the occluded vessel becomes an extension of the catheter, allowing estimation of *pulmonary venous pressure,* also known as *pulmonary wedge pressure.* Pulmonary wedge pressure (average, 5 mm Hg) is only slightly greater than left atrial pressure (average, 3 to 4 mm Hg). The small difference between the mean pulmonary artery (15 mm Hg) and left atrial (4 mm Hg) pressures indicates that the pulmonary circulation offers little resistance to blood flow. PVR is calculated as follows:

$$PVR = (Ppa - Pla)/\dot{Q}$$

where Ppa is mean pulmonary arterial pressure, Pla is left atrial pressure, and $\dot{Q}$ is cardiac output.

PVR is low in the normal resting animal and it *decreases even further* when pulmonary blood flow and/or pulmonary vascular pressures increase, as occurs during exercise. This is because an increase in pressure recruits previously unperfused vessels. Additionally, pulmonary vascular smooth muscle relaxes during exercise so that small arteries and veins dilate. Together these changes increase the capacity of the pulmonary vasculature.

In the systemic circulation, arteries and arterioles provide the primary source of vascular resistance; in the pulmonary circulation, the resistance is more evenly distributed across the arteries or arterioles and capillaries, with smaller contributions from veins (Fig. 46.4). Unlike the arterioles in the systemic circulation, the small arteries in the pulmonary circulation neither provide large resistance nor dampen the arterial pulsations; consequently, pulmonary capillary blood flow is *pulsatile.*

The Distribution of Pulmonary Blood Flow Within the Lung Differs Between Quadrupeds and Bipeds

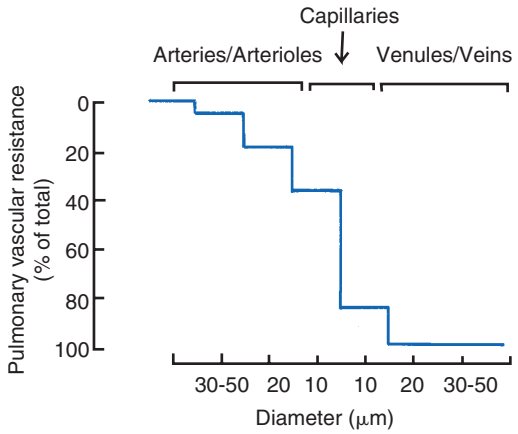
In bipedal animals—most notably humans—the distribution of blood across the pulmonary circulation is gravity-dependent, with blood flow per unit lung volume highest at the bottom of the lungs and lowest at the top of the lungs.

The distribution of pulmonary blood flow in quadrupeds is not determined by gravity but, rather, is preferentially distributed to the *caudodorsal region of the lungs* (Fig. 46.5). This caudodorsal distribution of blood in the lungs is accentuated by exercise and may even persist when posture changes during anesthesia. The branching pattern of pulmonary arteries and arterioles and the relative resistances of each vessel are the major determinants of blood flow distribution.

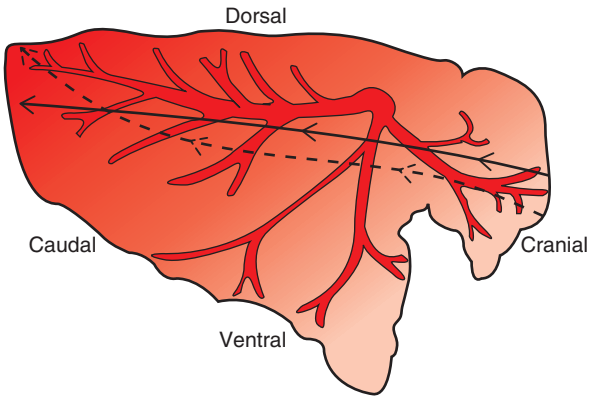
Neural and Humoral Factors Cause Contraction of Pulmonary Artery Smooth Muscle

A variety of neural and humoral factors can contract or relax pulmonary vascular smooth muscle and thereby alter the resistance to blood flow. The magnitude of the response of vessels to these stimuli is largely determined by the amount of smooth muscle in the small pulmonary arteries, which varies with species, and is greatest in cattle (Fig. 46.6).

An array of humoral factors moderate pulmonary vascular tone both in the resting state and in response to exercise or disease (Fig. 46.7). The mediators that cause *vasodilation* include nitric oxide (NO), bradykinin, adenosine, atrial natriuretic protein (ANP),



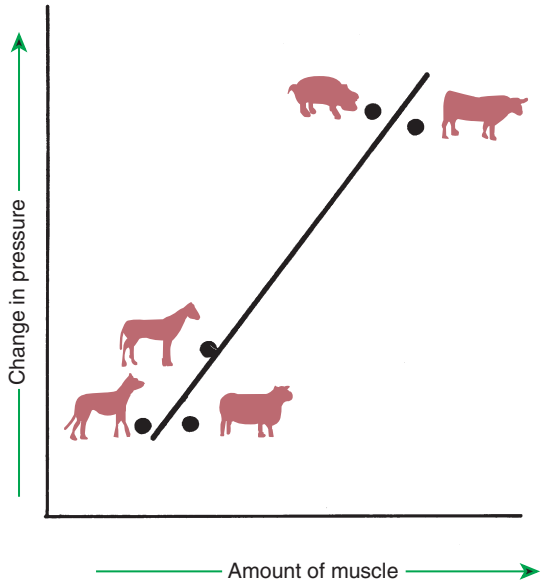
• **Fig. 46.4** Distribution of vascular resistance in the pulmonary circulation as determined by micropuncture studies. Unlike the resistance in the systemic circulation, a major portion of the resistance to blood flow in the pulmonary circulation is in the capillary bed. (From Bhattacharya J, Staub NC. Direct measurement of microvascular pressures in the isolated perfused dog lung, *Science* 1980;210(4467):327–328. Copyright 1980 by the American Association for the Advancement of Science.)



• **Fig. 46.5** Graphic representation of the distribution of pulmonary blood flow in the equine lung. Relative blood flow is indicated by the intensity of the *red shading*. Blood flows preferentially to the caudodorsal regions of the lung at rest (*solid line*). During exercise the caudodorsal distribution of pulmonary blood flow is exacerbated (*dashed line*) as more blood is directed away from the cranioventral and toward the caudodorsal lung regions. (Compiled from data in Hlastala MP, Bernard SL, Erikson HH, et al. Pulmonary blood flow distribution in standing horses is not dominated by gravity. *J Appl Physiol*. 1996;81(3):1051–1061.)

prostaglandin I_2 (PGI_2, a.k.a. prostacyclin), and acetylcholine, among others. Pulmonary vascular *vasoconstrictors* include hypoxia (discussed later), enthothelin-1, serotonin (a.k.a. 5-hydroxytryptamine, 5-HT), histamine, norepinephrine, the leukotrienes C_4, D_4, and E_4, and the prostaglandins E_2, $F_{2\alpha}$, and thromboxane.

The pulmonary arteries receive both *sympathetic* and *parasympathetic* innervation; the density of this *autonomic innervation* varies among species. Although the pulmonary circulation has both alpha-adrenergic receptors (which cause vasoconstriction) and beta-adrenergic receptors (which cause vasodilation), the net effect of sympathetic activation is vasoconstriction. Acetylcholine released from parasympathetic nerves activates muscarinic receptors causing vasodilation. Similarly, pulmonary inhibitory nonadrenergic noncholinergic (iNANC) nerves release various mediators, such as nitric oxide, vasoactive intestinal peptide, and substance P to effect vasodilation.

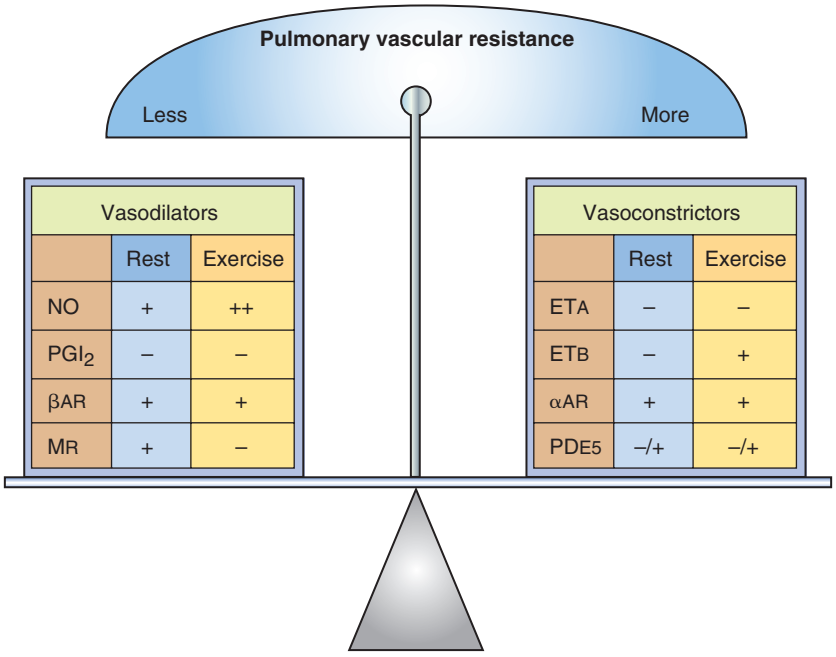


• **Fig. 46.6** Relationship between the amount of muscle in the media of small pulmonary arteries and the change in pulmonary arterial pressure when animals are exposed to a hypoxic environment. Animals with thicker muscle layers in their small pulmonary arteries, such as cattle and pigs, have a greater vascular response to hypoxia than do animals with a smaller amount of muscle in these blood vessels, such as dogs and sheep. Horses have an intermediate response.

Fig. 46.7 compares several important vasoactive agents and receptors involved in regulation of PVR. Responses to receptor activation may vary between species and with the initial degree of vascular smooth muscle tone. Responses may also change between rest and exercise. Some mediators, such as bradykinin or stimulation of inhibitory nonadrenergic noncholinergic nerves, relax smooth muscle and cause vasodilation by releasing nitric oxide from the endothelium. Release of nitric oxide also occurs in response to the increased shear stress across the endothelium when blood flow increases. The increased nitric oxide release is partly responsible for the dilation of the pulmonary circulation during exercise. Catecholamines, bradykinin, and prostaglandins are metabolized by the vascular endothelium, so their effects may be modified by endothelial damage.

Alveolar Hypoxia is a Potent Constrictor of Small Pulmonary Arteries

The distribution of air to alveoli can be reduced by the local presence of fluid, exudates, or other airway obstructions. The air in such poorly ventilated alveoli has a low partial pressure of oxygen (O_2), and therefore it is of limited benefit to the animal to keep sending blood to such alveoli. To correct this problem, *alveolar hypoxia* induces vasoconstriction of pulmonary arteries serving the hypoxic region. This *hypoxic pulmonary vasoconstriction* reduces blood flow to poorly ventilated alveoli and redistributes pulmonary blood flow toward better-ventilated regions of the lung. Under conditions of *atelectasis,* when there is no ventilation to the collapsed region of lung, local blood flow is greatly reduced by a combination of vasoconstriction in response to the local hypoxia, as well as vessel closure as the lung collapses. In contrast to this hypoxia-induced vasoconstriction of the pulmonary arteries, systemic arteries respond to local tissue hypoxia by vasodilating and thereby increasing blood flow and O_2 delivery to affected tissues.

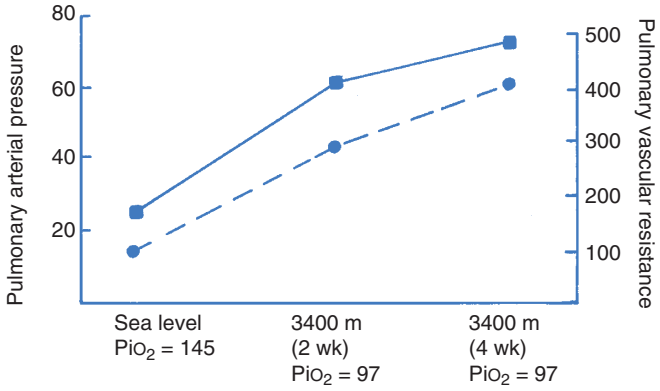


• **Fig. 46.7** Important vasoactive agents and receptors regulating pulmonary vascular resistance at rest and during exercise. Vasodilators tip the scale toward less pulmonary vascular resistance and vasoconstrictors tip it toward more resistance. The + and – signs, or their combination, represent the weight of the effect that the agent or receptor produces on its side of the scale. –, No effect; *NO,* nitric oxide; *PGI_2,* prostacyclin; *βAR,* beta-adrenergic receptor; *MR,* muscarinic receptor; *ETA,* endothelin A receptor; *ETB,* endothelin B receptor; *αAR,* alpha-adrenergic receptor; *PDE5,* phosphodiesterase-5. (Drawn with information from Merkus D, de Beer VJ, Houweling B, et al. Control of pulmonary vascular tone during exercise in health and pulmonary hypertension. *Pharmacol Therap*. 2008;119(3):242–263.)

Although the pulmonary vasoconstrictor response to hypoxia is present in all species, the magnitude of the response varies greatly. Among domestic mammals, the response is most vigorous in cattle and pigs, less vigorous in horses, and trivial in sheep and dogs. The increase in pulmonary vascular pressure in response to alveolar hypoxia and other stimuli is greater in cattle than in most other species because of the greater amount of smooth muscle in their pulmonary arteries (Fig. 46.6).

Hypoxic vasoconstriction is beneficial when there is localized alveolar hypoxia, but when pulmonary hypoxia is generalized, as can occur when animals live at high altitude or have diffuse lung disease, the vasoconstriction can have serious negative consequences. In some cattle grazing at high altitude, the hypoxia of altitude causes vigorous, generalized hypoxic pulmonary vasoconstriction (Fig. 46.8). This leads to an increase in pulmonary arterial pressure, which increases the work of the right ventricle and leads to *right-sided heart failure.* This clinical syndrome is known as *bovine high-altitude disease* or *brisket disease* because edematous fluid accumulates in the dependent (ventral) portions of the chest, termed brisket. When animals have generalized hypoxic pulmonary vasoconstriction as a result of lung disease, the resultant right-sided heart failure is known as *cor pulmonale.* In animals that experience an ongoing hypoxic constrictor response to either altitude or lung disease, the chronic hypoxia can result in sustained pulmonary hypertension, which is caused by an increase in the quantity of smooth muscle in the media of the small pulmonary arteries.

The magnitude of the pulmonary vasoconstrictive response to hypoxia is genetically influenced, and measurements of pulmonary artery pressure in cattle grazing at high altitude have allowed selection of breeding stock that are less susceptible to bovine high-altitude disease. Some species, such as sheep and llama, can tolerate the hypoxic conditions at high altitude because their pulmonary endothelial cells produce more of the vasodilator nitric oxide. Cattle not only have more muscular pulmonary arteries, but they also produce little nitric oxide to counteract the vasoconstriction induced by hypoxia.



• **Fig. 46.8** Change in mean pulmonary arterial pressure (*squares, solid line*) and pulmonary vascular resistance (*circles, broken line*) in calves 2 and 4 weeks after being transported from sea level to 3400 m. Both vascular resistance and arterial pressure increase when the calves are exposed to the hypoxia of high altitude. Pressure and resistance continue to increase while at high altitude because of the proliferation of smooth muscle in the small pulmonary arteries. Pressure units are in mm Hg; resistance units in dyne-s/cm^5; inspired O_2 partial pressure (Pio_2) in mm Hg. (From data in Ruiz AV, Bisgard GE, Will JA. Hemodynamic response to hypoxia and hyperoxia in calves at sea level and altitude. *Pflugers Arch*. 1973; 344(4):275–286.)

The Pulmonary Circulation Accommodates a Large Increase in Blood Flow During Exercise

To transport the extra O_2 required for muscular effort, cardiac output increases six- to eightfold during strenuous exercise. This increase in blood flow passes through the pulmonary circulation, where it collects O_2. To accommodate the increase in blood flow, the pulmonary blood vessels dilate; that is, PVR decreases. This dilation is in part passive as a result of the increase in intravascular pressure, which is a result of the increased blood flow. In addition, flow-induced release of nitric oxide from the endothelium also causes relaxation of smooth muscle and vessel dilation.

In most species, pulmonary arterial pressure during strenuous exercise is about 35 mm Hg, but in horses it increases to more than 90 mm Hg. The latter increase is attributable in large part to a very high left atrial pressure (50 mm Hg or more), which is probably necessary for rapid left ventricular filling when the heart rate exceeds 200 beats/min. When left atrial pressure is high, pulmonary arterial and capillary pressures must necessarily be even greater to maintain blood flow through, and prevent backup within, the lung. These high exercise-associated intravascular pressures cause leakage of erythrocytes from the pulmonary capillaries when horses exercise strenuously, a phenomenon known as *exercise-induced pulmonary hemorrhage*.

Pulmonary Microvasculature Filters Small Particles From the Blood

Along with the conventional pulmonary arteries that follow the bronchial tree, there are also supernumerary arteries that provide auxiliary blood flow to gas exchange airways. This extensive pulmonary vasculature, with additional support from the bronchial vasculature, provides collateral channels for blood flow to the lungs. As such, the pulmonary vasculature is ideally suited to trap small particles, including thromboemboli, fat particles, and tumor cells that form in or enter the circulation, without substantially compromising the required blood flow to the lungs. If such emboli were to lodge in systemic vascular beds, which typically have little collateral circulation, tissue hypoxia would result in most cases. In addition to its collateral circulation, the pulmonary vasculature also has means of removing small emboli lodged in the pulmonary circulation via well-developed proteolytic and anticoagulant enzymatic systems.

Bronchial Circulation

The Bronchial Circulation Provides a Blood Supply to Airways and Large Vessels and Aids in Conditioning Inhaled Air

The bronchial circulation is part of the systemic circulation. The origin of the bronchial circulation, which receives less than 2% of the output of the left ventricle, varies between species but typically arises from the aorta or intercostal arteries. Arteries of the bronchial circulation follow the tracheobronchial tree to the terminal bronchioles and form a *peribronchial plexus* in the connective tissue along the length of the airways. Branches from this plexus penetrate the smooth muscle layer of the bronchial wall and form a *subepithelial vascular plexus* that serves to warm inhaled air. Branches are also given off to form the *vasa vasorum* (nutrient blood vessels) of the pulmonary vessels and the thoracic aorta. The bronchial vessels also supply mediastinal structures such as the lower third of the esophagus and the pericardium. In cattle, sheep, pigs, and horses, the bronchial arteries provide blood flow to the visceral pleura; in dogs, cats, and monkeys, they do not.

The venous drainage of the bronchial circulation is unique in that it returns to the heart by both the systemic and pulmonary circulations. Specifically, about 20% of bronchial venous blood returns to the right atrium by means of the azygous vein. In this manner, deoxygenated blood from the bronchial veins is returned to the right heart, which is typical of the systemic circulation. In contrast, the remaining approximately 80% of deoxygenated bronchial venous blood drains into the pulmonary veins and then the left atrium; this unique and seemingly illogical configuration results in deoxygenated blood from the bronchial veins mixing with the oxygenated blood leaving the lungs in the pulmonary veins and returning to the left heart. Vascular anastomoses between the bronchial capillaries or veins and pulmonary capillaries or veins allow deoxygenated blood from the bronchial circulation to bypass the alveoli and return to the left atrium without having participated in gas exchange. This vascular configuration is one form of an anatomic *shunt* because the deoxygenated blood leaving the bronchial circulation does not flow past alveoli and become oxygenated but rather is shunted directly into the left side of the heart (left atrium) for distribution in the systemic arterial circulation. The small amount of deoxygenated bronchial blood mixing with the oxygenated blood leaving the lungs contributes to the small difference between the alveolar and arterial O_2 partial pressures that occurs under normal conditions (see Chapter 47).

The pulmonary circulation is unique in that it has little capacity for *angiogenesis* (the growth of new blood vessels), which other vascular beds employ to restore perfusion to tissues in the face of vascular obstruction. Because of this, when parts of the pulmonary circulation are obstructed, it is the bronchial circulation that proliferates and maintains blood flow to the affected part of the lung. Because much of the bronchial circulation drains into the pulmonary veins, the blood supplying the lung parenchyma via bronchial neovascularlization, which occurs in response to pulmonary artery obstruction, has a path for returning to the left heart. The bronchial circulation also proliferates when the airways are inflamed. In acute inflammatory conditions, this bronchial vascular remodeling can support tissue healing; however, chronically, it can lead to *hemoptysis* (coughing up blood).

Inflow pressure to the bronchial circulation is systemic arterial pressure, but outflow pressure varies, depending on whether venous drainage is through the azygos or pulmonary veins. Changes in pressure in both the systemic and the pulmonary vascular beds affect the magnitude of bronchial blood flow. Increasing systemic pressure increases the bronchial flow, but increasing pulmonary vascular pressures (downstream pressure) reduce the flow and may even reverse it. Under hypoxic conditions, bronchial arteries dilate; in contrast, pulmonary arteries constrict under these conditions.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Bovine High-Altitude ("Brisket") Disease in a Heifer

History

A 2-year-old Hereford heifer was kept during the winter on a farm in the foothills of the Rocky Mountains outside Denver, Colorado. In the late spring, the heifer was transported to Climax, Colorado (altitude 3460 m; 11,360 ft), for summer grazing. After 3 weeks, the owners noticed that the animal was having some difficulty breathing, was reluctant to move around the pasture, and had developed an enlarged pendulous brisket and swelling underneath the lower jaw.

Clinical Examination

Inspection of the heifer reveals a lethargic animal in poor condition. The respiratory and heart rates are elevated. Air seems to be moving well through the nostrils. The most noticeable observation is an enlarged and pendulous brisket; the swelling extends up the neck and into the submandibular area. The jugular veins are distended.

Palpation of the swollen brisket reveals that it is heavy; when it is squeezed, the imprints of the fingers remain for some time. The submandibular swelling behaves in a similar manner when palpated. The mucous membranes of the heifer are a normal color, and the lung sounds are unremarkable.

Comment

The swelling in the brisket and submandibular regions, which pits on palpation, is evidence of accumulation of interstitial edema in the (gravity-) dependent areas of the heifer, in which there is loose connective tissue. Accumulation of edema in these regions is an indication of an increase in systemic venous pressure, which is also causing jugular distention. Both are caused by *right-sided heart failure.*

The most likely cause of right-sided heart failure in a heifer grazing at high altitude is diffuse vasoconstriction of the pulmonary circulation as a result of chronic hypoxia (inspired O_2 partial pressure (Pi_{O_2}) at 3460 m elevation is 96 mm Hg, compared with 149 mm Hg at sea level). The smooth muscle in the pulmonary arteries contracts in response to hypoxia; if this response is maintained for several weeks, the amount of smooth muscle in the pulmonary arteries increases. Furthermore, the animal produces extra erythrocytes in an attempt to transport more O_2. These extra erythrocytes increase the hematocrit and make the blood more viscous and difficult to pump through the capillary beds. Maintenance of cardiac output in the presence of the elevated PVR and increased blood viscosity requires an increase in pulmonary arterial and right ventricular pressures. Persistently working against the increased pressure causes right-sided heart failure. Because bovine high-altitude disease has a genetic component, this heifer should not be used for breeding.

Treatment

For immediate treatment, this animal should be moved to an area of lower altitude and could be given O_2 to further relieve the hypoxic stimulus. The vasospasm in the pulmonary circulation and the hematocrit diminish once the hypoxic stimulus is removed. This would cause a reduction in pulmonary arterial pressure, which would cause some relief. However, pulmonary arterial pressure may not decrease to normal levels because of the increased amount of smooth muscle now present in the pulmonary arteries. In the early stages of the disease, this process may be reversible. When cardiac signs develop, however, the prognosis is guarded. Additional supportive treatments, including diuretics, should be implemented.

Stallion With Bilateral Epistaxis Postracing

History

A 4-year-old thoroughbred gelding has been racing frequently. It has performed well at the beginning of the races but has been fading toward the end of the races. The last time the horse raced, the trainer noted blood coming out of both nostrils at the end of the race. The horse has not had any other problems thus far.

Clinical Examination

Physical examination parameters are normal. A rebreathing examination is performed to detect subtle changes in the lungs; none are identified. Endoscopic examination performed at rest reveals a small amount of dark blood in the distal trachea. Bronchoalveolar lavage (BAL) is performed, and the cytology of the sample is characterized by predominantly erythrocytes with a small percentage of neutrophils. Culture for bacteria is negative.

Comment

This horse has a clinical case of exercise-induced pulmonary hemorrhage (EIPH). Although the majority of racehorses leak blood into their air passages during intense exercise, bleeding from the nostrils occurs in less than 0.5%. Because some pulmonary hemorrhage is common following the intense exertion of racing, it is important to rule out other causes of poor performance in racehorses with evidence of EIPH.

The exercising horse has a very high O_2 consumption that necessitates a high cardiac output to transport O_2 from the lungs to the muscles. The heart beats more than 200 times per minute and pumps about 1.5 liters of blood per beat. For left ventricular filling to occur in the short time available, pulmonary venous pressure must be very high. Maintenance of blood flow requires that pulmonary arterial pressure exceeds pulmonary venous pressure, so pressures throughout the pulmonary circulation are all elevated. In fact, pulmonary capillary pressure can be as high as 100 mm Hg. The high pulmonary capillary pressures create a large pressure difference between the alveoli and the pulmonary capillary bed, leading to a breakdown of the epithelial and endothelial tight junctions, resulting in bleeding into the alveoli. Recently, narrowing of small pulmonary veins due to wall thickening has also been identified in horses with EIPH. By obstructing venous outflow, this would further contribute to high capillary pressure.

Treatment

Furosemide is a diuretic that, when used prophylactically (4 hours before strenuous exercise), decreases EIPH. By increasing urination, furosemide decreases blood volume, which decreases pulmonary capillary pressure. This drug can be legally used on race day in the United States and some other countries.

Practice Questions

1. Which of the following statements accurately describes the pulmonary circulation?

- **a.** Pulmonary arteries carry oxygenated blood from the right ventricle to the alveolar capillaries.
- **b.** The medial layer of the main pulmonary arteries is composed of a thick layer of smooth muscle.
- **c.** The pulmonary veins return blood to the right atrium.
- **d.** The pulmonary circulation receives the total output of the right ventricle.
- **e.** All the above.

2. During exercise, cardiac output can increase fivefold, but pulmonary arterial pressure may not even double. This occurs because:

- **a.** Pulmonary vascular resistance decreases during exercise.
- **b.** Unperfused capillaries are recruited during exercise.
- **c.** Previously perfused vessels are distended during exercise.
- **d.** Factors that dilate the pulmonary arteries are released by the endothelium during exercise.
- **e.** All the above.

3. Which of the following will cause the greatest increase in pulmonary arterial pressure?
   a. Exposure of a cow to the hypoxia of high altitude
   b. A twofold increase in pulmonary blood flow
   c. Stimulation of the vagus nerve (parasympathetic system) in a sheep
   d. Inhalation of a tidal volume in a horse
   e. None of the above

4. The bronchial circulation:
   a. Receives the total output of the right ventricle.
   b. Drains into the pulmonary circulation and azygos vein.
   c. Vasoconstricts in response to hypoxia.
   d. Supplies nutrient blood flow only to bronchi and no other structures.
   e. Has a bronchial arterial pressure of the same magnitude as pulmonary arterial pressure.

5. Pulmonary vascular resistance (PVR) is calculated as:
   a. PVR = (right ventricle pressure – left ventricle pressure)/cardiac output
   b. PVR = (aortic pressure – right atrial pressure)/cardiac output
   c. PVR = (pulmonary artery pressure – left atrial pressure)/cardiac output
   d. PVR = (left atrial pressure – pulmonary venous pressure)/cardiac output
   e. None of the above

Bibliography

Bailie EM. The anatomy and physiology of the bronchial circulation. *J Aerosol Med*. 1996;9(1):1–8.

Boron WF. Ventilation and perfusion of the lungs. In: Boron WF, Boulpaep EL, eds. *Medical Physiology: A Cellular and Molecular Approach*. 3rd ed. Philadelphia: Saunders; 2017.

Deffebach ME, Charan NB, Lakshminarayan S, et al. State of art. The bronchial circulation. Small, but a vital attribute of the lung. *Am Rev Respir Dis*. 1987;135(2):463–481.

Hinchcliff KW, Couetil LL, Knight PK, et al. Exercise induced pulmonary hemorrhage in horses: American College of Veterinary Medicine consensus statement. *J Vet Intern Med*. 2015;29(3):743–758.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Holt TN. *Overview of Bovine High-Mountain Disease, Merck Veterinary Manual*. Kenilworth, NJ: Merck & Co., Inc; 2016.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lekeux P, Art T. The respiratory system: anatomy, physiology and adaptations to exercise and training. In: Hodgson DR, McGowan CM, McKeever K, eds. *The Athletic Horse: Principles and Practice of Equine Sports Medicine*. 2nd ed. Philadelphia: Saunders; 2014.

Lumb AB. *Nunn's Applied Respiratory Physiology*. 7th ed. London: Churchill Livingstone; 2010.

Merkus D, de Beer VJ, Houweling B, et al. Control of pulmonary vascular tone during exercise in health and pulmonary hypertension. *Pharmacol Ther*. 2008;119(3):242–263.

Mitzner W, Wagner EM. Vascular remodeling in the circulations of the lung. *J Appl Physiol*. 2004;97(5):1999–2004.

Murray JF. *The Normal Lung*. 2nd ed. Philadelphia: Saunders; 1986.

Robinson NE. Some functional consequences of species differences in lung anatomy. *Adv Vet Sci Comp Med*. 1982;26:1–33.

Ruiz AV, Bisgard GE, Will JA. Hemodynamic response to hypoxia and hyperoxia in calves at sea level and altitude. *Pflugers Arch*. 1973;344(4):275–286.

Suresh K, Shimoda LA. Lung circulation. *Compr Physiol*. 2016;6:897–943.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

Wolin MS, Gupte SA, Mingone CJ, et al. Redox regulation of responses to hypoxia and NO-cGMP signaling in pulmonary vascular pathophysiology. *Ann N Y Acad Sci*. 2010;1203:126–132.

47

Gas Exchange

SUSAN L. EWART

KEY POINTS

1. The composition of a gas mixture can be described by the fractional composition or partial pressures.
2. Alveolar gas composition is determined by alveolar ventilation and the exchange of oxygen and carbon dioxide.
3. Impaired gas exchange results in hypoxemia.
4. Alveolar hypoventilation is synonymous with elevated arterial carbon dioxide partial pressure (P_aCO_2).
5. Exchange of oxygen and carbon dioxide between the alveoli and pulmonary capillary blood occurs by diffusion.
6. The exchange of gases between the blood and tissues also occurs by diffusion.
7. The amount of alveolar ventilation in relation to pulmonary capillary blood flow—the ventilation/perfusion ratio ($\dot{V}_A/\dot{Q}$)—determines the adequacy of pulmonary gas exchange.
8. The gas partial pressures of the systemic arterial blood are determined by the gas partial pressures of the capillary blood passing each alveolus.
9. Right-to-left shunts contribute to hypoxemia by allowing blood to bypass ventilated alveoli.
10. Part of each breath ventilates dead space and does not participate in gas exchange.
11. Partial pressures of arterial oxygen (P_aO_2) and carbon dioxide (P_aCO_2) are measured to evaluate gas exchange.

The Composition of a Gas Mixture Can Be Described by Its Fractional Composition or Partial Pressures

Understanding gas exchange requires an understanding of gas composition and the forces causing gas movement within the lungs, blood, and tissues. For convenience, physiologists use many abbreviations when describing gas exchange (Table 47.1). Ambient air consists of 21% oxygen (O_2) (the *fraction* of O_2 in ambient air, FO_2, is 0.21) and 78% nitrogen (N_2) with other gases, including carbon dioxide (CO_2) and water vapor, together making up the remaining small fraction. At elevated altitudes, such as in the mountains, the air still contains 21% O_2, but because the atmospheric pressure decreases as altitude increases, the *total amount* of O_2 available for respiration is reduced. At lower atmospheric pressure, the O_2 molecules are less densely packed, and therefore the *partial pressure* of O_2 (PO_2) in the air is decreased. It is this partial pressure that is important in O_2 movement into and out of the blood.

The PO_2 in a dry gas mixture is determined by *atmospheric pressure* (P_{atm}) and the *fraction of O_2* (FO_2) in the gas mixture, as follows:

$$PO_2 = P_{atm} \times FO_2$$

In the atmosphere, FO_2 is 0.21, so PO_2 in dry air at sea level where P_{atm} = 760 mm Hg is approximately 160 mm Hg:

$$PO_2 = 760 \text{ mm Hg} \times 0.21 = 160 \text{ mm Hg}$$

Just as the P_{atm} can vary with altitude, the FO_2 can also vary under certain circumstances. Although the *fraction* of O_2 in the air that is *inspired* (F_iO_2) when breathing normal atmospheric air is 0.21, the value of F_iO_2 can range up to 100%. For example, when a patient is supplemented with nasal insufflation of 100% O_2, the F_iO_2 will increase.

During inhalation, air is warmed to body temperature and humidified to saturation in the larger air passages. The concentration of O_2 and other gases is reduced by the presence of water vapor molecules in this humidified air; therefore PO_2 is less in humidified air than in dry air. The PO_2 of humidified gas in the airways is calculated as follows:

$$PO_2 = (P_{atm} - P_{H_2O}) \times F_iO_2$$

where P_{H_2O} is the *partial pressure of water vapor* at body temperature. The P_{H_2O} is determined by the temperature and percentage saturation of the air with water. In a mammal with a body temperature of 38°C, P_{H_2O} in saturated air equals 50 mm Hg; therefore at sea level where P_{atm} = 760 mm Hg, the P_iO_2 of warmed, completely humidified gas in the conducting airways is approximately 149 mm Hg:

$$P_iO_2 = (760 \text{ mm Hg} - 50 \text{ mm Hg}) \times 0.21 = 149 \text{ mm Hg}$$

Alveolar Gas Composition Is Determined by Alveolar Ventilation and the Exchange of Oxygen and Carbon Dioxide

Oxygen is consumed and CO_2 is produced during aerobic metabolism. Carbon dioxide generated in the tissues by metabolism is transferred to the lungs in the blood. Despite the fact that CO_2 is constantly produced, its amount in inspired air is negligible. For this reason, the *alveolar partial pressure of CO_2* (P_ACO_2) is determined by the rate of CO_2 production ($\dot{V}CO_2$) in relation to the amount of *alveolar ventilation* ($\dot{V}_A$):

$$P_ACO_2 = (P_{atm} - P_{H_2O}) \times (\dot{V}CO_2 / \dot{V}_A)$$

TABLE 47.1 Common Abbreviations Used in Gas Exchange

Abbreviation	Definition
$A\text{-}aDO_2$	Alveolar-arterial oxygen difference
F_iO_2	Fraction of oxygen in inspired air
FO_2	Fraction of oxygen in a gas mixture
P_aCO_2	Arterial carbon dioxide partial pressure
P_ACO_2	Alveolar carbon dioxide partial pressure
P_aO_2	Arterial oxygen partial pressure
P_AO_2	Alveolar oxygen partial pressure
P_{atm}	Atmospheric pressure
$P_{cap}CO_2$	Capillary carbon dioxide partial pressure
$P_{cap}O_2$	Capillary oxygen partial pressure
PCO_2	Carbon dioxide partial pressure
P_{H_2O}	Partial pressure of water vapor
P_iO_2	Inspired oxygen partial pressure
PO_2	Oxygen partial pressure
P_vCO_2	Venous carbon dioxide partial pressure
P_vO_2	Venous oxygen partial pressure
$\bar{V}$	Mixed venous blood
$\dot{Q}$	Perfusion
R	Respiratory exchange ratio
$\dot{V}$	Ventilation
$\dot{V}_A$	Alveolar ventilation
$\dot{V}CO_2$	Rate of carbon dioxide production
$\dot{V}_A/\dot{Q}$	Ratio of alveolar ventilation to blood flow

It follows from this equation that, if $\dot{V}CO_2$ increases as occurs during the increased metabolism of exercise, $\dot{V}_A$ must also increase if P_ACO_2 is to remain constant. If $\dot{V}_A$ does not increase sufficiently, P_ACO_2 rises. Similarly, if $\dot{V}CO_2$ remains constant and $\dot{V}_A$ halves, then P_ACO_2 doubles.

Because O_2 is constantly consumed, the partial pressure of O_2 in alveolar air (P_AO_2) is lower than the partial pressure of O_2 in inspired air (P_iO_2). During breathing, P_AO_2 fluctuates around an average value, increasing during inhalation and decreasing during exhalation. The average partial pressure of O_2 in the alveoli is the difference between the O_2 inspired minus the O_2 consumed, which can be calculated from the *alveolar gas equation,* a simplified version of which is:

$$P_AO_2 = P_iO_2 - O_2 \text{ consumed}$$

The P_iO_2 is a function of the fraction of O_2 in the inspired air, the atmospheric pressure, and water pressure added by humidification during inspiration, as previously described. The O_2 consumed is a function of the CO_2 produced and the type of fuel metabolized, as follows:

$$O_2 \text{ consumed} = P_ACO_2/R$$

• BOX 47.1 Respiratory Exchange Ratio (R) = CO_2 Produced/O_2 Consumed

- Glucose fuel, R = 1.0
- Free fatty acid fuel, R = 0.7
- Typical mix of glucose and free fatty acid fuel, R = 0.8

Note: Protein is not completely oxidized in vivo, so its R value is difficult to quantify, but is estimated to be 0.82.

where R, the *respiratory exchange ratio* (also referred to as *respiratory quotient,* RQ), is the ratio of the amount of CO_2 produced to the amount of O_2 consumed. The respiratory exchange ratio is determined by the substrates being metabolized by the animal (Box 47.1).

Putting this together, we can calculate the partial pressure of O_2 in alveolar air as:

$$P_AO_2 = [(P_{atm} - P_{H_2O}) \times F_iO_2] - P_ACO_2/R$$

The following modified form of the alveolar gas equation, in which *arterial* CO_2 partial pressure (P_aCO_2) is substituted for *alveolar* CO_2 partial pressure (P_ACO_2), is used when P_aCO_2 values from blood gas analyses are available but P_ACO_2 values, which require capnography, are unavailable (unless otherwise noted, in this chapter the terms *arterial* and *Pa* refers to *systemic* arteries):

$$P_AO_2 = [(P_{atm} - P_{H_2O}) \times F_iO_2] - P_aCO_2/R$$

Assuming an average R of 0.8 and a P_aCO_2 of 40 mm Hg, P_AO_2 averages approximately 100 mm Hg under normal conditions at sea level.

$$P_AO_2 = [(760 \text{ mm Hg} - 50 \text{ mm Hg}) \times 0.21] - 40 \text{ mm Hg}/0.8 = 99 \text{ mm Hg}$$

P_aCO_2 is measured by arterial blood gas analysis, whereas P_ACO_2 is measured noninvasively by *capnography* as end tidal CO_2 ($ETCO_2$, see Chapter 48). A *capnograph* is a monitoring device that measures CO_2 concentration in expired air; in veterinary medicine, capnography is most often used in intubated patients during anesthesia but can also be used in an awake animal. Because CO_2 is highly soluble and equilibrates readily across the alveolar epithelium, alveolar and arterial CO_2 partial pressures are very similar under normal conditions. When P_ACO_2 and P_aCO_2 values diverge, the gradient between them can be a useful diagnostic aid for various respiratory and circulatory disturbances.

In summary, the alveolar gas equation demonstrates that P_AO_2 is determined by the P_iO_2 and the exchange of O_2 for CO_2. It also shows that whenever P_ACO_2 increases, P_AO_2 decreases, and vice versa. Comparing calculated values of alveolar and arterial O_2 partial pressures is clinically useful for differentiating causes of hypoxemia.

Impaired Gas Exchange Results in Hypoxemia

Hypoxemia is the state of low partial pressure of O_2 in arterial blood (P_aO_2) and is one of the major causes of *hypoxia,* which is the state of inadequate levels of O_2 reaching the tissues. Low levels of inspired O_2 or impairment of gas exchange by the following mechanisms can result in hypoxemia: alveolar hypoventilation, diffusion impairment, low ventilation to perfusion ratios,

TABLE 47.2 Causes of Hypoxemia and Their Distinguishing Features

Cause	A-aDO_2	P_aCO_2	Response to O_2 Therapy	Comment
Alveolar hypoventilation	Normal	Increased	Improved	
Low P_IO_2	Normal	Normal (or reduced)[a]	Improved	Situation specific to high altitude or anesthesia
Low $\dot{V}_A/\dot{Q}$	Increased	Normal (or increased or reduced)[b]	Improved	Common cause of hypoxemia
Right-to-left shunt	Increased	Normal (or reduced)[a]	Minimal improvement	
Diffusion impairment	Increased	Normal (or reduced)[a]	Improved	Uncommon cause of hypoxemia; often requires exercise to induce symptoms
Intense exertion of elite athlete	Increased	Increased	Not applicable	Blood gas values return to normal after exercise

[a]P_aCO_2 may be reduced by reflexive (hypoxemia-induced) increased ventilation.
[b]In most cases, P_aCO_2 is normal or reduced by reflexive (hypoxemia-induced) increased ventilation; however, in some cases, the additional work of this compensatory increased ventilation is not sustained and P_aCO_2 increases.

right-to-left shunts, and intense exercise in elite athletes, such as race horses (Table 47.2). The normal physiological processes for these aspects of gas exchange and how they may be impaired are described in the following sections.

Determining the cause of hypoxemia in clinical patients is important for appropriate case management. Comparing P_AO_2, as calculated by the alveolar gas equation, to P_aO_2, as measured by blood gas analysis, can help distinguish the cause of hypoxemia (see Table 47.2). The difference between these two values is termed the *alveolar-arterial O_2 difference* (*A-aDO_2*, also known as *alveolar-arterial O_2 gradient*) and averages 5 to 10 mm Hg in normal patients. The A-aDO_2 remains normal in hypoxemic patients when both the lungs and blood experience decreased oxygenation, as occurs in alveolar hypoventilation or when the P_IO_2 is substantially decreased as occurs at altitudes of about 1500 meters (5000 ft) or greater. In contrast, the A-aDO_2 is increased in hypoxemic patients when delivery of O_2 to the lungs is normal, but its uptake is impaired by diffusion abnormalities, low ventilation-perfusion ratios, or right-to-left shunt.

Alveolar Hypoventilation Is Synonymous With Elevated Arterial Carbon Dioxide Partial Pressure (P_aCO_2)

Alveolar hypoventilation is defined as a decrease in alveolar ventilation in relationship to CO_2 production; this relative decrease in ventilation results in accumulation of CO_2 in the alveoli, with subsequent increase in P_aCO_2 as the alveolar and arterial partial pressures of CO_2 equilibrate across the alveolar epithelium. Thus *alveolar hypoventilation is clinically diagnosed by the presence of increased P_aCO_2*. Furthermore, as CO_2 accumulates in the alveoli during hypoventilation, it displaces alveolar O_2 with a resultant decrease in P_AO_2 and subsequently a decrease in P_aO_2. Fig. 47.1 shows the causes of alveolar hypoventilation. It occurs when (1) the central nervous system respiratory centers are diseased, injured, or depressed by drugs (anesthetics, opioids); (2) there is disease or injury to the nerve conduction pathways or their origins such as spinal cord injury (e.g., motor ventral horns), or insult to the phrenic nerves that supply the diaphragm; (3) the neuromuscular junction is diseased (anticholinesterase toxicity, myasthenia gravis); (4) there is disease or damage to the thorax (hit by car), pleural space, or respiratory muscles; (5) there is substantial airway obstruction (foreign body obstruction, laryngeal hemiplegia in dogs with geriatric onset laryngeal paralysis polyneuropathy [GOLPP], or horses with left recurrent laryngeal neuropathy); or (6) there is lung disease that severely decreases lung compliance.

The converse of alveolar hypoventilation, *alveolar hyperventilation,* causes a decrease in P_ACO_2 because ventilation is increased in relation to CO_2 production. Therefore according to the alveolar gas equation, as P_ACO_2 decreases, P_AO_2 increases. Hyperventilation occurs when the need to ventilate is increased by stimuli such as hypoxemia, hypoxia, acidosis, or an increase in body temperature, as well as due to pain, fear, or anxiety.

Exchange of Oxygen and Carbon Dioxide Between the Alveoli and Pulmonary Capillary Blood Occurs by Diffusion

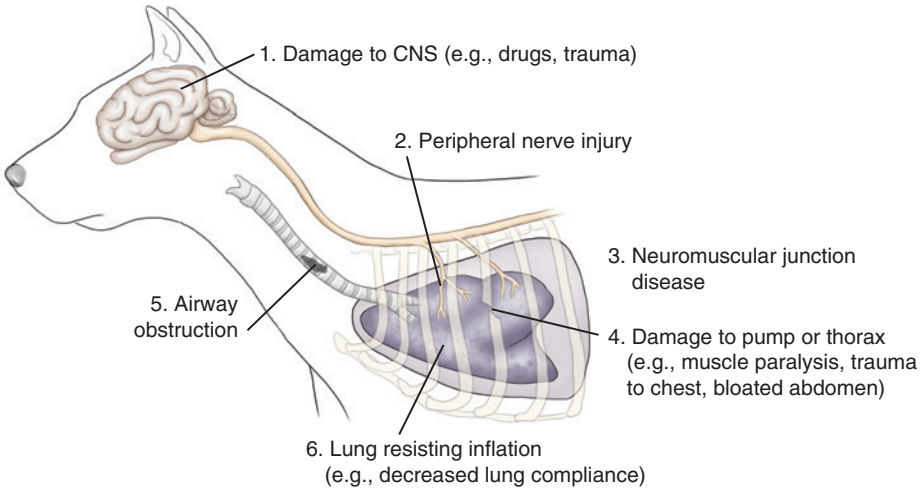
Diffusion is the passive movement of gases down a concentration (partial pressure) gradient. The rate of gas movement between the alveoli and the blood is determined by the physical properties of the gas (D), the surface area available for diffusion (A), the thickness of the air-blood barrier (T), and the driving pressure gradient of the gas between the alveolus and capillary blood. The equations for diffusion of O_2 and CO_2 are written as follows:

$$O_2 \text{ diffusion rate} = DO_2 \times A \times (P_AO_2 - P_{cap}O_2)/T$$

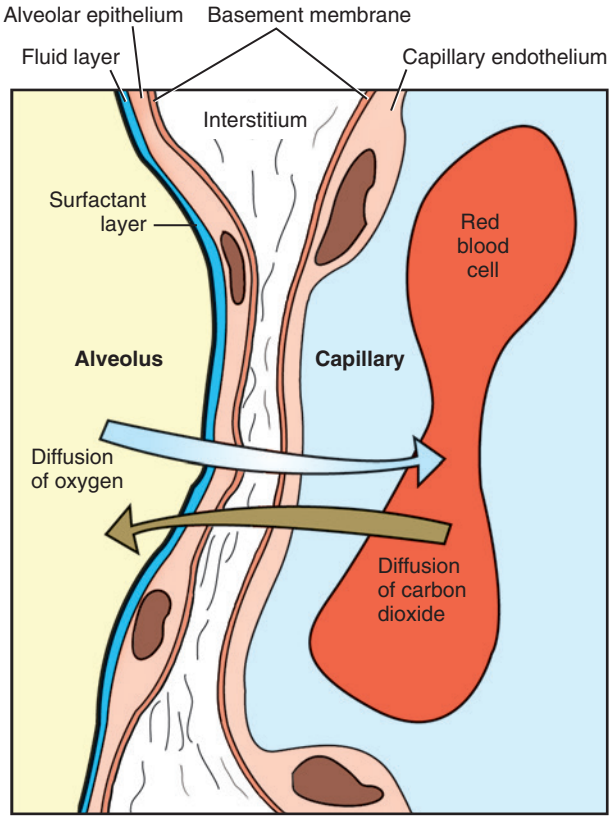
$$CO_2 \text{ diffusion rate} = DCO_2 \times A \times (P_ACO_2 - P_{cap}CO_2)/T$$

The *diffusion coefficient (D)* of a gas is determined by its molecular weight and solubility. Although the molecular weight of CO_2 is slightly more than that of O_2, the solubility of CO_2 is much greater than that of O_2, which results in CO_2 diffusing through the alveolar septa approximately 20 times faster than O_2.

The *surface area (A)* available for diffusion is the area of ventilated alveoli adjacent to perfused pulmonary capillaries. During exercise, more capillaries become perfused by blood, and thus the alveolar surface area available for diffusion increases. In contrast, dehydration



• **Fig. 47.1** Diagrammatic representation of the causes of alveolar hypoventilation. *CNS,* Central nervous system.



• **Fig. 47.2** Diagrammatic representation of the air-blood barrier within the lung, showing the pathway for diffusion of O_2 and CO_2 between the alveolus and the erythrocyte within the pulmonary capillary.

and certain pulmonary vascular diseases decrease the surface area of perfused capillaries.

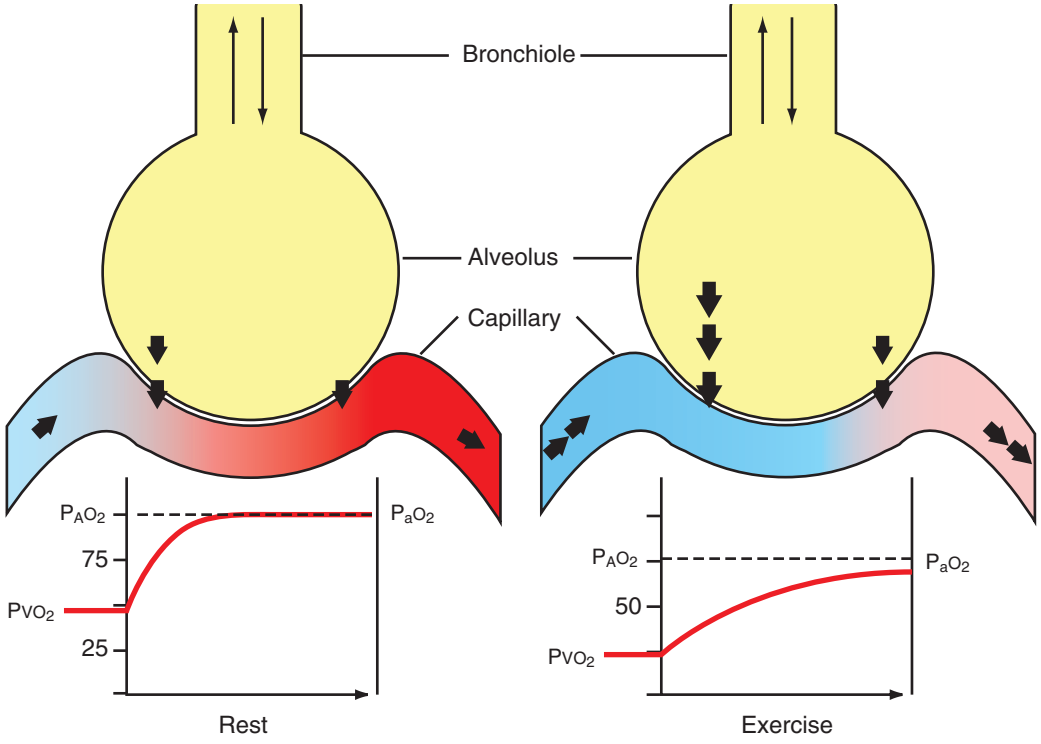
The *thickness (T) of the barrier* separating air and blood is that of the alveolar septum, which is less than 1 µm thick (Fig. 47.2). Although very thin, the alveolar septum consists of multiple layers, including a layer of liquid and surfactant lining the alveolar surface; an epithelial layer, usually formed by type I alveolar epithelial cells; an alveolar basement membrane; variable-thickness interstitium; an endothelial basement membrane; and a layer of capillary endothelium. In mammals, these six layers constitute the "thick side" of the alveolar septum, whereas the "thin side" has only four layers as it has a single basement membrane shared by the alveolar epithelium and capillary endothelium and the absence of an interstitium (see Chapters 46 and 50). In addition to moving gases through this air-blood barrier, diffusion also moves gases within the plasma, allowing O_2 to gain access to erythrocytes and hemoglobin.

The *driving pressure* for gas diffusion is the difference in gas partial pressure between the alveolus and the capillary blood. Blood entering the alveolar capillaries from the small pulmonary arteries is known as *mixed venous blood* because it has returned to the right side of the heart in veins from all parts of the systemic circulation. In a resting animal, the mixed venous blood entering the alveolar capillary has an O_2 partial pressure (P_vO_2) of approximately 40 mm Hg; the P_AO_2 averages close to 100 mm Hg. Thus the driving pressure gradient of 60 mm Hg (100–40) causes rapid diffusion of O_2 into the capillary, where it combines with hemoglobin. Hemoglobin takes up O_2 from the plasma, which helps maintain the pressure gradient for O_2 diffusion.

The CO_2 partial pressure in mixed venous blood (P_vCO_2) returning to the lungs averages 46 mm Hg, and P_ACO_2 is 40 mm Hg. Thus the driving pressure for diffusion of CO_2 from the blood into the alveolus is only 6 mm Hg. Despite this small driving pressure, the amount of CO_2 that diffuses per minute between the capillaries and the alveoli is similar to the amount of O_2 moving in the opposite direction. The 20-fold greater solubility of CO_2 compared with O_2 compensates for the small driving pressure gradient. For the same reason, CO_2 diffusion between the blood and the alveoli is rarely affected by lung disease.

Normally, in the resting animal, it takes approximately 0.75 second for an erythrocyte to travel through a pulmonary capillary; equilibration between alveolar and capillary O_2 partial pressures occurs early in that transit time (within 0.25 second) (Fig. 47.3). During strenuous exercise, muscles extract a large amount of O_2 from the blood so the mixed venous blood returning to the lung contains little O_2. In addition, during exercise the cardiac output is high and the *velocity of blood flow* through the capillaries is rapid. More O_2 must therefore be transferred in less time than in the resting animal. Under these strenuous conditions, diffusion equilibrium may not occur, and the P_aO_2 may decrease during intense exercise (Fig. 47.3). This exercise-associated arterial hypoxemia is observed in racing horses and working or racing dogs.

In diseased lungs, diffusion of O_2 across the alveolar septum may be impeded as a result of edema, inflammation, or fibrosis



• **Fig. 47.3** Schematic representation of two alveoli and their pulmonary capillaries, showing the increase in P_{O_2} that occurs as blood passes through the capillaries. The number of the *arrows* between the alveolus and the capillary represents the magnitude of the O_2 fluxes. In the resting animal, P_{vO_2} is approximately 40 mm Hg, and blood and air equilibrate rapidly. In the exercising animal, P_{vO_2} is reduced and even though O_2 fluxes are high the blood does not equilibrate with the P_AO_2 before it leaves the alveolus. P_AO_2, Alveolar oxygen partial pressure; P_aO_2, arterial oxygen partial pressure; P_{vO_2}, mixed venous oxygen partial pressure.

that thickens the air-blood barrier (T) or reduces the surface area (A) available for gas exchange. Although these changes to the airway architecture impede gas diffusion across the alveolar septa, there is great reserve capacity for diffusion such that hypoxemia due to diffusion impairment is unusual in the resting animal. In contrast, diffusion impairment is notable for causing hypoxemia during exercise; this occurs because exercise increases the velocity of blood flow, which reduces the contact time between gases and blood in the pulmonary capillaries and further taxes the diffusion reserve capacity. The administration of O_2 is useful to treat hypoxemia due to diffusion impairment as the increased P_AO_2 provides a greater driving pressure to deliver O_2 into the blood (see Table 47.2).

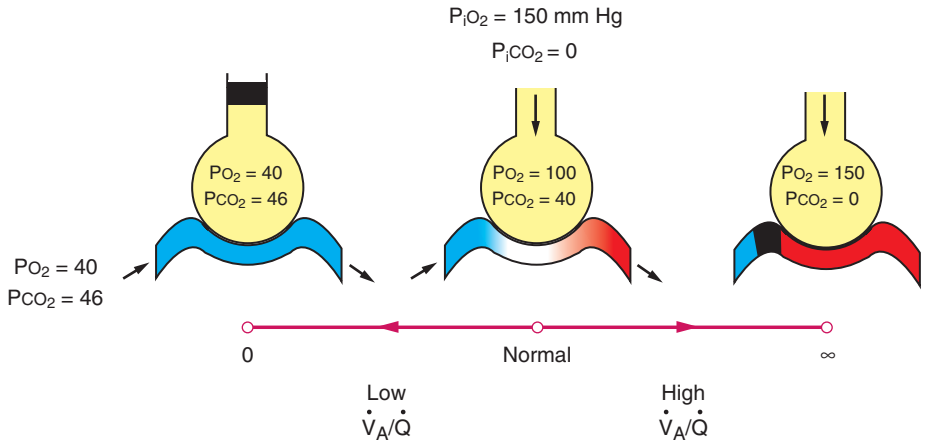
The Exchange of Gases Between the Blood and Tissues Also Occurs by Diffusion

The P_aO_2 of blood entering the tissue capillaries from the systemic arteries is 85 to 100 mm Hg, and the P_aCO_2 averages 40 mm Hg. As blood passes through these capillaries, it is exposed to the tissues consuming O_2 and producing CO_2. *Tissue O_2 partial pressure* is determined by the rate of delivery of O_2 in relation to its rate of consumption; it averages 40 mm Hg. Similarly, *tissue CO_2 partial pressure* is determined by the rate of tissue production of CO_2 in relation to its rate of removal by the blood; it averages 46 mm Hg. As a result of the partial pressure differences between the tissues and capillaries, O_2 diffuses out of the blood and into the tissues whereas CO_2 diffuses out of the tissues and into the blood until the partial pressures of each gas across blood and tissue are equal. Tissues with a high O_2 demand have more capillaries per gram of tissue. This provides a larger surface (A) for diffusion and also means that the maximal distance between the tissue and the nearest capillary is less than in the poorly vascularized tissues.

During exercise, muscle blood flow increases, in part, as a result of recruitment of capillaries that are not perfused in the resting animal. *Capillary recruitment* brings blood closer to the metabolizing tissues and slows the rate of blood flow, which allows more time for diffusion equilibrium. In addition, the increased O_2 utilization and CO_2 production by muscles during exercise lowers the P_{O_2} and increases the P_{CO_2} of the muscle, which increases the driving pressure gradients for diffusion.

The Amount of Alveolar Ventilation in Relation to Pulmonary Capillary Blood Flow—the Ventilation/Perfusion Ratio ($\dot{V}_A/\dot{Q}$)—Determines the Adequacy of Pulmonary Gas Exchange

In the lungs, gas exchange is accomplished by the close apposition of alveolar air and capillary blood. Ideally, each of the millions of alveoli would receive air and blood in amounts optimal for gas exchange; that is, alveolar ventilation ($\dot{V}_A$) and perfusion ($\dot{Q}$) would be perfectly matched. However, even in young healthy animals, there is some mismatching of ventilation and perfusion. In disease, this mismatching becomes more extreme and leads to *hypoxemia,*



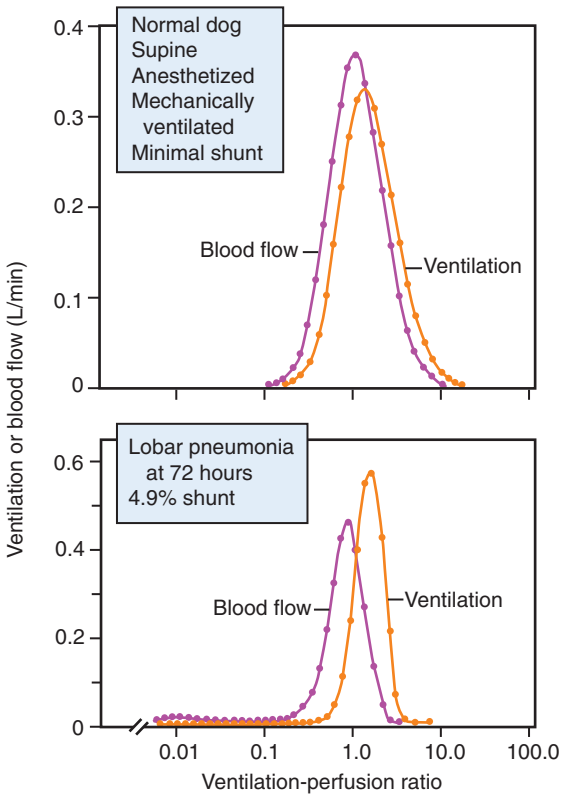
• **Fig. 47.4** Diagrammatic representation of three alveoli and their capillaries showing the effect of differing ventilation to perfusion ratios ($\dot{V}_A/\dot{Q}$) on O_2 and CO_2 partial pressures (P_{O_2} and P_{CO_2}, respectively). See text for explanation. (Adapted from West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016, Fig. 5.7, page 72.) *P_iO_2,* inspired oxygen partial pressure; *P_iCO_2,* inspired carbon dioxide partial pressure.

a low P_aO_2. The matching of ventilation with perfusion in bipedal animals (including humans) is further complicated by their upright posture and the greater effect of gravity on blood (because of its greater mass) than on air. Specifically, in bipedal species, the top of the lungs receives a greater proportion of ventilation, and the bottom of the lungs receives a gravity-driven greater proportion of perfusion, contributing to mismatching of ventilation and perfusion even under normal conditions.

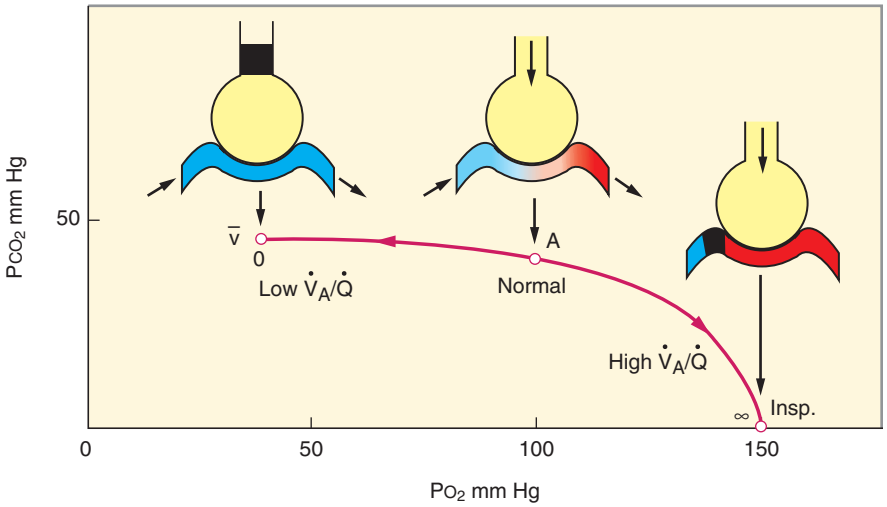
Fig. 47.4 shows schematic alveoli and capillaries with three ventilation/perfusion ratios ($\dot{V}_A/\dot{Q}$). The alveolus in the center is ideal (close to normal); it receives ventilation and blood flow with a $\dot{V}_A/\dot{Q}$ of about 1.0. Mixed venous blood arrives at each alveolus with a P_{O_2} of 40 and a P_{CO_2} of 46 mm Hg, equilibrates with alveolar gas partial pressures, and leaves the normal alveolus with an end-capillary P_{O_2} and P_{CO_2} of 100 and 40 mm Hg, respectively. The alveolus on the left is served by an obstructed bronchiole so it receives no ventilation. Its $\dot{V}_A/\dot{Q}$ is therefore low (in this extreme case, zero), and blood passes through without participating in gas exchange; this is termed right-to-left shunt because blood is flowing ("shunted") from the right side of the heart (right ventricle) to the left side of the heart (left atrium) without passing ventilated alveoli (see Chapter 46 for other causes of shunts). By contrast, the alveolus on the right receives no blood flow but continues to ventilate; its $\dot{V}_A/\dot{Q}$ is high (infinity). In this alveolus, the composition of the alveolar gas approaches that of inspired air, but because there is no blood flow, the alveolus makes no contribution to gas exchange; this is alveolar dead space.

The millions of gas exchange units in the lungs can have $\dot{V}_A/\dot{Q}$ ranging from zero to infinity. In normal individuals, the range of $\dot{V}_A/\dot{Q}$ is actually fairly small and averaged across the lungs is approximately 0.8, meaning that, on average, there is slightly more perfusion than ventilation reaching the alveoli (Fig. 47.5). Any alveoli with low $\dot{V}_A/\dot{Q}$ are relatively overperfused and/or underventilated. Fig. 47.6 shows that the P_{O_2} of blood leaving these low $\dot{V}_A/\dot{Q}$ units is decreased substantially, but the overall P_{CO_2} is less affected and may even be in the normal range (see end of next section). Low $\dot{V}_A/\dot{Q}$ units occur frequently in lung disease when ventilation is impeded by airway obstruction or related pathologies.

The alveolus to the right of the normal unit in Fig. 47.6 has a high $\dot{V}_A/\dot{Q}$; ventilation is high in relation to blood flow. This can



• **Fig. 47.5** Distribution of ventilation (orange circles) and blood flow (red circles) as a function of the ventilation to perfusion ratio ($\dot{V}_A/\dot{Q}$). *Top,* In the normal dog, most of the blood flow and ventilation are received by gas exchange units with a $\dot{V}_A/\dot{Q}$ close to 1.0. No blood flow and no ventilation are received by units with extremely high or extremely low $\dot{V}_A/\dot{Q}$. *Bottom,* In the dog with pneumonia, an increased portion (4.9%) of the blood flow is received by units with low $\dot{V}_A/\dot{Q}$, that is, units with little to no ventilation (shunt). (From Wagner PD, Laravuso RB, Goldzimmer E, et al. Distributions of ventilation-perfusion ratios in dogs with normal and abnormal lungs. *J Appl Physiol.* 1975;38(6):1099–1109.)



• **Fig. 47.6** Effect of ventilation to perfusion ratios ($\dot{V}_A/\dot{Q}$) on alveolar O_2 and CO_2 partial pressures. Three alveoli are shown, but $\dot{V}_A/\dot{Q}$ can range from zero to infinity. When an alveolus receives no ventilation ($\dot{V}_A/\dot{Q} = 0$, *left*), alveolar gas equilibrates with mixed venous blood and there is no gas exchange. When an alveolus receives no blood flow but ventilation continues ($\dot{V}_A/\dot{Q} = \infty$, *right*), alveolar gas composition approaches that of inspired air. All alveoli between these two $\dot{V}_A/\dot{Q}$ extremes have some gas exchange and their gas composition is shown by the *red line.* A normal (ideal) alveolus *(center)* with a $\dot{V}_A/\dot{Q} = 0.8$ has P_{O_2} and P_{CO_2} of 100 and 40 mm Hg, respectively. $\bar{V}$, Mixed venous blood; *Insp.,* inspired air. (Adapted from West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016, Fig. 5.8, page 73.)

occur when blood flow to part of the lungs is reduced as occurs with vascular obstruction or hypoxic pulmonary vasoconstriction. The blood leaving such units has an elevated P_{O_2} and decreased P_{CO_2}, however, this does not offset the decreased P_aO_2 of blood from low $\dot{V}_A/\dot{Q}$ units as described later.

The Gas Partial Pressures of the Systemic Arterial Blood Are Determined by the Gas Partial Pressures of the Capillary Blood Passing Each Alveolus

Blood that returns from the lungs to the left heart for distribution to the tissues comes from capillaries associated with millions of alveoli, each with its own $\dot{V}_A/\dot{Q}$. The content of O_2 and CO_2 in blood that leaves each alveolus varies because of these differing $\dot{V}_A/\dot{Q}$ (Fig. 47.6). Thus the gas partial pressures of arterial blood are determined by distribution of $\dot{V}_A/\dot{Q}$ in the lungs.

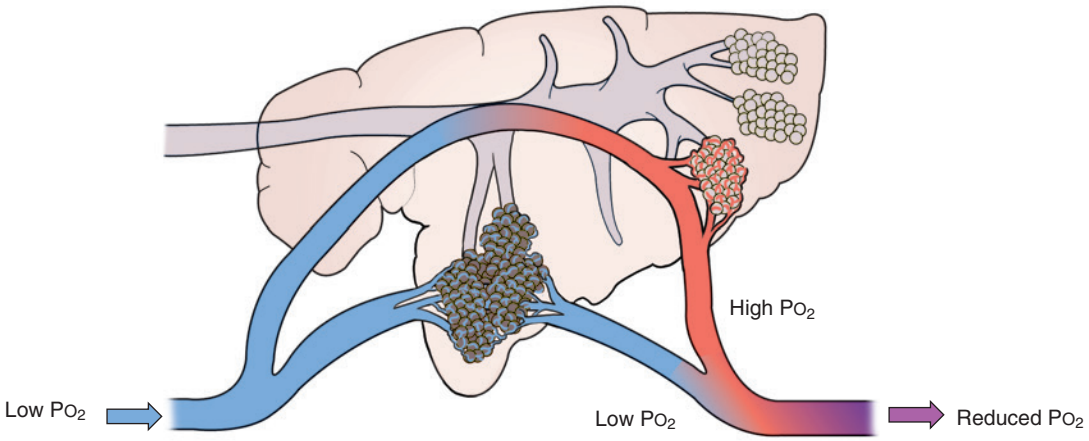
Lung diseases accentuate mismatching of ventilation and perfusion because of obstruction of airways, flooding of alveoli with exudates, and obstructions to blood flow. This mismatching of ventilation and perfusion has a major effect on O_2 exchange and is the most common cause of hypoxemia. Low $\dot{V}_A/\dot{Q}$ regions have a much greater effect than high $\dot{V}_A/\dot{Q}$ regions on the overall P_aO_2 values for several reasons. First, high $\dot{V}_A/\dot{Q}$ regions have—by definition—low blood flow, thus the volume of blood with increased P_aO_2 values is limited. Second, the extent to which P_aCO_2 values can rise is constrained by the P_iO_2, which is maximally 149 mm Hg when breathing room air. Finally, the total O_2 content (CaO_2) of blood from high $\dot{V}_A/\dot{Q}$ regions is only minimally increased; because of the sigmoidal shape of the *oxyhemoglobin dissociation curve* (see Chapter 48), increases in P_{O_2} above about 80 mm Hg add very little O_2 to the blood. Thus even if the redistribution of ventilation and/or perfusion that creates low $\dot{V}_A/\dot{Q}$ regions is accompanied by the creation of high $\dot{V}_A/\dot{Q}$ regions, the relatively overventilated (high $\dot{V}_A/\dot{Q}$) alveoli with a high P_AO_2 cannot add enough O_2 to the blood to compensate for the deficiency that arises from the underventilated (low $\dot{V}_A/\dot{Q}$) alveoli with a low P_AO_2. Therefore hypoxemia occurs to varying degrees in most lung diseases.

In contrast to O_2, inequalities of ventilation and perfusion have less effect on the exchange of CO_2 between alveoli and blood. Carbon dioxide is very soluble, and because its dissociation curve (see Chapter 48) is almost linear, the overventilated alveoli can compensate for those that are underventilated. Furthermore, hypoxemia—from $\dot{V}_A/\dot{Q}$ inequalities or other mechanisms—stimulates the peripheral chemoreceptors to increase alveolar ventilation (see Chapter 49), which serves to reduce P_aCO_2. For these reasons, *hypercapnia,* also called *hypercarbia* (increased P_aCO_2), as a feature of lung disease is less common than hypoxemia (see Table 47.2).

As the degree of $\dot{V}_A/\dot{Q}$ mismatching increases and O_2 exchange becomes less efficient, the difference between the average P_AO_2 and P_aO_2 increases, thus the A-aDO_2 increases (see Table 47.2).

Right-to-Left Shunts Contribute to Hypoxemia by Allowing Blood to Bypass Ventilated Alveoli

In a *right-to-left shunt,* blood from the right ventricle bypasses ventilated alveoli and returns poorly oxygenated blood to the left atrium. An example is provided in Fig. 47.7 where the blood flowing to a lung lobe that is consolidated (not ventilated) due to severe bronchopneumonia does not become oxygenated; as a result, the O_2 and CO_2 partial pressures of the capillary blood leaving these diseased alveoli are the same as the deoxygenated blood that entered the lungs. When this blood mixes with blood that has perfused the healthy lung, it dilutes the O_2 content so that the blood going to systemic arteries has a lower P_aO_2 than normal. Right-to-left shunts have a $\dot{V}_A/\dot{Q}$ of zero and are formed when alveoli are collapsed *(atelectasis),* are unventilated because of complete *airway obstruction,* or are filled with exudates, as in *pneumonia.* Right-to-left shunts can also result from complex congenital cardiac defects, such as *tetralogy of Fallot,* which allow blood to flow directly from the right to the left chambers of the heart, bypassing the lungs. These pathological right-to-left shunts are a major impairment to O_2 exchange and thus cause hypoxemia. Interestingly, a small amount of physiologically normal shunting of blood is present in normal, healthy animals due to the normal anatomy of the bronchial and coronary circulations (see Chapter 46); the deoxygenated blood from the *bronchial veins* travels through bronchopulmonary anastomoses to the *pulmonary veins* that carry oxygenated blood leaving the lungs. Similarly, a portion of the blood in the coronary veins drains directly into the left heart.



• **Fig. 47.7** Diagrammatic representation of a region of pneumonia in an otherwise healthy lung. The pneumonic region receives no ventilation so blood flow through that region does not pick up any O_2 (i.e., it forms a right-to-left shunt). When this poorly oxygenated blood mixes with well-oxygenated blood coming from the healthy lung, the result is a lower P_{O_2} than normal in blood returning to systemic arteries for delivery to the tissues.

These normal anatomical shunts contribute to the small difference between alveolar and arterial P_{O_2} under normal conditions. In the presence of pathological right-to-left shunts, the A-aDO_2 widens (see Table 47.2). A unique feature of right-to-left shunting is that treatment with supplemental O_2 does not substantially improve the shunt-induced hypoxemia. Because the shunt fraction is not passing ventilated alveoli (and in some cases not passing alveoli at all), increasing the P_iO_2, and thus P_AO_2, does not improve the oxygenation of the shunted blood.

Part of Each Breath Ventilates Dead Space and Does Not Participate in Gas Exchange

Dead space ventilation consists of the gas that does not participate in gas exchange. This includes both anatomical dead space and alveolar dead space (see Chapter 45). The latter consists of alveoli that receive ventilation but no blood flow; that is, they have a $\dot{V}_A/\dot{Q}$ of infinity (Fig. 47.6, *right*). *Alveolar dead space* can form when the precapillary blood flow to the lungs is limited, as occurs in pulmonary embolism. The alveoli in these regions have elevated P_AO_2 and depressed P_ACO_2 levels, approaching those of inspired air, because they do not participate in gas exchange with the blood. The reduced flow of blood past ventilated alveoli reduces the opportunity for the CO_2 elimination, which in most cases results in increased minute ventilation in response to stimulation of chemoreceptors by the retained CO_2; the increased minute ventilation is accompanied by an increase in the work of breathing. Under these circumstances, the P_aCO_2 is not increased, so long as the increased work of breathing can be maintained. However, in some patients, for example those with reduced lung compliance, this increased work of breathing cannot be sustained and their P_aCO_2 values eventually rise.

Partial Pressures of Arterial Oxygen (P_aO_2) and Carbon Dioxide (P_aCO_2) Are Measured to Evaluate Gas Exchange

A systemic arterial blood sample is used to evaluate pulmonary gas exchange because this blood has just passed through the lungs. In contrast, a venous blood sample is inadequate, especially for assessing oxygenation, because its composition varies depending on the blood flow to metabolism ratio of the tissue of origin. Arterial blood gas partial pressures are the end result of the individual processes involved in gas exchange and thus are affected by the *composition of inspired air, alveolar ventilation, gas diffusion from the alveoli to capillaries,* and *matching of ventilation and perfusion.* All these factors must be considered when evaluating blood gas values.

Inspired air usually contains 21% O_2 ($F_iO_2 = 0.21$), but during anesthesia or O_2 therapy, the administration of O_2 increases F_iO_2, which causes an increase in P_iO_2. Although the daily fluctuations in atmospheric pressure cause only trivial changes in P_iO_2, the decrease in atmospheric pressure that occurs at higher altitudes results in a major decrease in P_iO_2. As a result, there is a decrease in P_AO_2 and thus a decrease in P_aO_2 as animals ascend in altitude. In the appropriate geographical locations, altitude-induced changes in P_aO_2 must always be considered when blood gas partial pressures are evaluated.

Adequacy of alveolar ventilation is assessed by examination of P_aCO_2; it is elevated above the normal value of 40 mm Hg when animals hypoventilate and is decreased during hyperventilation. Furthermore, hypoventilation decreases P_AO_2 and P_aO_2 in the face of a normal A-aDO_2, whereas hyperventilation increases these O_2 partial pressures.

Diffusion abnormalities and $\dot{V}_A/\dot{Q}$ mismatching impair the transfer of O_2 from the alveolus to arterial blood, increase the A-aDO_2, and reduce P_aO_2. P_aCO_2 is rarely elevated by these problems for two reasons. First, the high solubility of CO_2 allows easy diffusion across the remaining healthy lung. Second, the hypoxemia caused by lung disease stimulates ventilation. The resultant increase in alveolar ventilation keeps P_aCO_2 normal or even reduces it below normal.

In animals with normal lungs, administering O_2 (increasing F_iO_2) elevates P_aO_2. As $\dot{V}_A/\dot{Q}$ mismatching becomes more extreme, O_2 administration increases P_aO_2 only modestly, especially in the presence of right-to-left shunts. Concurrently, the A-aDO_2 widens. The P_aO_2 response to O_2 is a good way to evaluate the type and severity of lung disease (see Table 47.2).

P_aO_2 tends to be lower in newborn animals than in adults. This is because there is greater mismatching of ventilation and blood flow in the immature lungs.

It is worth noting that the equipment to measure blood gas values is typically limited to larger referral veterinary practices; instead, many practices use pulse oximetry to assess O_2 saturation of hemoglobin (SO_2), which is a function of (but not equivalent to) P_aO_2 (see Chapter 48 for discussion of SO_2). For CO_2 assessment in the absence of a blood gas analyzer, veterinary practices may use capnography to measure end tidal CO_2 ($ETCO_2$), which gives a close approximation of P_ACO_2. $ETCO_2$ is about 2 to 5 mm Hg lower than P_aCO_2 under normal conditions; the gradient (difference) between these two measures is increased by alveolar dead space.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Hypoventilation in an English Bulldog

History

A 2-year-old English Bulldog is presented to you because it refuses to exercise. Normally the dog is willing to go for short, slow walks. Over the past 6 months, the dog has been making an increasing amount of noise when it breathes. When it is awake, it makes a rattling sound during inhalation; when it sleeps, it snores loudly and wakes frequently, standing up, turning around, and then lying down again. On one occasion, the owner tried to get the dog to run, but the dog collapsed, making a loud noise in its throat as it struggled to inhale.

Clinical Examination

The dog is in good condition, but even as you walk into the room, you notice the loud rattling noises being made by the dog during breathing. You also observe that the mucous membranes of the pendulous lips have a bluish tinge. The dog is standing when you walk into the room, but while you talk to the owner, the dog lies down and apparently goes to sleep. This causes the breathing noises to become louder.

Examination of the dog reveals no abnormalities in the heart or the digestive tract, but examination of the respiratory tract reveals multiple abnormalities. The external nares of the dog are unusually small, and it is difficult to introduce a speculum to examine the nasal cavity. When the dog's mouth is opened, an excessive amount of loosely folded tissue is observed in the pharynx, and it is impossible to move this aside to examine the larynx. Listening to the lungs is not helpful because the sounds generated by the loose, vibrating tissue in the upper airway are transmitted to the lungs. Radiographs reveal no abnormalities in the lungs, but the trachea is quite narrow. An arterial blood sample is taken for blood gas analysis. P_aO_2 is 50 mm Hg (normal, 85–100), and P_aCO_2 is 75 mm Hg (normal, 35–45). The A-aDO_2 is calculated to be 5.35 mm Hg (normal 5–10 mm Hg).

Comment

This dog's condition is consistent with *brachycephalic syndrome,* which is seen in brachycephalic (shorter skull configuration) dog breeds, particularly bulldogs. The syndrome usually includes stenosis (narrowing) of the external nares and obstruction of the pharynx by pendulous folds of excessive soft tissue. In some of these dogs, the trachea is also narrow. Affected dogs have difficulty breathing, particularly during inhalation, when the subatmospheric pressure within the upper airway sucks the loose folds of tissue into the airway lumen, causing obstruction to airflow. In general, these dogs make a lot of noise during inhalation as the loose folds of tissue vibrate. Exhalation presents less difficulty because the higher-than-atmospheric pressure in the pharynx tends to push back the loose tissue and open the airway. Over time, the chronic, excessively subatmospheric pressure during inhalation can cause deformity of the larynx.

The upper airway obstruction in this dog is limiting ventilation so severely that the dog is suffering from alveolar hypoventilation, as indicated by the elevated P_aCO_2. An elevation in P_aCO_2 occurs when alveolar ventilation is not sufficient to remove the CO_2 being produced by the body. The accumulating CO_2 in the alveoli and the lack of ventilation also depress the P_AO_2, which leads to a decrease in P_aO_2 (hypoxemia), as in this dog. The hypoxemia then leads to an increase in desaturated hemoglobin, which accounts for the bluish color (cyanosis) of the dog's mucous membranes.

Treatment

The treatment for this dog is surgical removal of some of the excessive tissues of the upper airway and enlargement of the external nares. This will alleviate some of the obstruction and improve ventilation. However, with the narrowing of the trachea observed in this dog, it is unlikely that the dog will ever be able to exercise to a significant degree, although its condition may be improved sufficiently that it can make a suitable pet.

Hypoxemia in an Anesthetized Clydesdale Horse

History

A 2-year-old, 750-kg Clydesdale horse is presented for removal of a testicle that has been retained in the abdomen, a procedure that requires anesthesia. You know that anesthesia of heavy draft horses can lead to gas exchange problems, and therefore you have an anesthesia machine available to provide ventilation and to supplement the horse with extra O_2. The horse is anesthetized with a short-acting intravenous anesthetic, positioned in dorsal recumbency and an endotracheal tube is inserted. The horse is connected to the anesthesia machine and administered 100% O_2 containing isoflurane for anesthesia. Ventilation is not assisted.

Thirty minutes after the induction of anesthesia, the veterinary technician takes an arterial blood sample to monitor the horse's gas exchange. P_aO_2 is 70 mm Hg and P_aCO_2 is 65 mm Hg. Are you satisfied with the results of the blood gas analysis? If not, what can be done to improve gas exchange?

Comment

Several factors are likely contributing to the blood gas abnormalities in this case. The elevation of P_aCO_2 shows that the horse is suffering from alveolar hypoventilation; that is, the ventilation received by the alveoli is insufficient to remove the CO_2 produced by the horse. This is probably a result of depression of the central nervous system by the anesthetic gases, so that the drive to breathe is reduced. In addition, the positioning of the horse on its back causes the heavy viscera to push on the diaphragm, making it difficult for the horse to ventilate. Alveolar hypoventilation in an anesthetized animal can be corrected by the use of positive-pressure ventilation. You have a ventilator as part of the anesthesia machine and choose to ventilate the horse to return the P_aCO_2 to acceptable levels.

The P_aO_2 of 70 mm Hg shows that the horse has considerable problems in exchanging O_2. Although a P_aO_2 of 70 mm Hg is sufficient to saturate most hemoglobin, it is very low in an animal breathing 100% O_2. Under normal conditions when animals breathe 100% O_2, the alveolar and arterial O_2 partial pressures will both rise to greater than 600 mm Hg (see the alveolar gas equation). In this horse, P_AO_2 is calculated to be 628.75 mm Hg, so the A-aDO_2 (628.75 – 70) has widened to 558.75 mm Hg (normal 5–10 mm Hg).

This large A-aDO_2 is not unusual in large, anesthetized mammals. The positioning of the horse on its back with the consequent weight of the viscera pushing forward on the diaphragm and compressing the lungs can lead to severe $\dot{V}_A/\dot{Q}$ inequalities and shunts. Parts of the dependent lung are unable to ventilate, although they continue to receive blood flow and therefore become right-to-left shunts. These right-to-left shunts result in arterial hypoxemia. As long as the P_aO_2 is sufficient to saturate hemoglobin, the horse is in no danger. The dangerous point is during recovery from anesthesia. The horse must be supplemented with O_2 until it is sufficiently conscious to be able to rest on its sternum unaided and eventually to stand. Returning to these postures eliminates right-to-left shunts, restores the $\dot{V}_A/\dot{Q}$ distribution to normal, and improves gas exchange.

Practice Questions

1. Calculate the P_AO_2 of an anesthetized cow when the atmospheric pressure is 750 mm Hg, P_{H_2O} at body temperature = 50 mm Hg, and P_aCO_2 = 80 mm Hg. The cow is breathing a mixture of 50% O_2 and 50% nitrogen. Assume the respiratory exchange ratio is 0.8.
- **a.** 250 mm Hg
- **b.** 620 mm Hg
- **c.** 275 mm Hg
- **d.** 195 mm Hg
- **e.** 670 mm Hg

2. Which of the following will decrease the rate of O_2 transfer between the alveolar air and the pulmonary capillary blood?
- **a.** Increasing P_AO_2 from 100 to 500 mm Hg
- **b.** Perfusing previously unperfused pulmonary capillaries
- **c.** Decreasing the P_vO_2 from 40 to 10 mm Hg
- **d.** Destruction of alveolar septa and pulmonary capillaries by a disease known as *alveolar emphysema*
- **e.** None of the above

3. During exercise, recruitment of muscle capillaries that are unperfused in the resting animal results in all the following except:
- **a.** An increase in the velocity of capillary blood flow.
- **b.** An increase in the surface area for gas diffusion between tissues and blood.
- **c.** A decrease in distance between tissue capillaries.
- **d.** Maintenance of tissue PO_2 in the presence of increased demand for O_2.
- **e.** A shorter distance for gas diffusion.

4. Which of the following could potentially result in more low $\dot{V}_A/\dot{Q}$ regions within the lung?
- **a.** Atelectasis of one lobe of a dog lung
- **b.** Obstruction of both pulmonary arteries
- **c.** Doubling the ventilation to the right cranial lobe while its blood flow remains constant
- **d.** Vasoconstriction of the pulmonary arteries of the left lung in a cow
- **e.** None of the above

5. Which of the following statements is correct?
- **a.** Right-to-left shunts represent an extremely high $\dot{V}_A/\dot{Q}$ ratio.
- **b.** Right-to-left shunts are not a cause of elevated A-aDO_2.
- **c.** An increase in the alveolar dead space can result from an increase in the number of high $\dot{V}_A/\dot{Q}$ units in the lung.
- **d.** The shape of the oxyhemoglobin dissociation curve means that low $\dot{V}_A/\dot{Q}$ units in the lung are not a cause of hypoxemia (low P_aO_2).
- **e.** Totally occluding the right pulmonary artery increases the right-to-left shunt fraction by 50%.

6. A horse has difficulty inhaling, especially during exercise. Arterial blood gas partial pressures at rest are P_aO_2 = 55 mm Hg and P_aCO_2 = 70 mm Hg. After you give the horse O_2 to breathe, P_aO_2 increases to 550 mm Hg, and P_aCO_2 remains unchanged. The cause of these gas partial pressures is:
- **a.** Right-to-left shunt through a complex cardiac defect.
- **b.** Alveolar hyperventilation.
- **c.** A large number of alveoli with high $\dot{V}_A/\dot{Q}$ ratios.
- **d.** Alveolar hypoventilation.
- **e.** None of the above.

Bibliography

Boron WF, Boulpaep EL. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lekeux P, Art T, Hodgson DR. The respiratory system: anatomy, physiology and adaptations to exercise and training. In: Hodgson DR, McGowan CM, eds. *The Athletic Horse: Principles and Practice of Equine Sports Medicine*. St. Louis: Saunders; 2014.

Petersson J, Glenny RW. Gas exchange and ventilation-perfusion relationships in the lung. *Eur Respir J*. 2014;44:1023–1041.

Poole DC, Erickson HH. Highly athletic terrestrial mammals: horses and dogs. *Compr Physiol*. 2011;1(1):1–37.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

West JB, Luks AM. *West's Pulmonary Pathophysiology: The Essentials*. 9th ed. Philadelphia: Wolters Kluwer; 2017.

48

Gas Transport in the Blood

SUSAN L. EWART

KEY POINTS

Oxygen Transport

1. Most oxygen is transported in combination with hemoglobin, but a small amount is in solution in plasma.
2. A molecule of hemoglobin can reversibly combine with four molecules of oxygen.
3. The oxyhemoglobin dissociation curve describes the relationship between oxygen partial pressures and hemoglobin saturation.
4. The affinity of hemoglobin for oxygen varies with blood temperature, pH, carbon dioxide partial pressure, and 2,3-BPG concentration.
5. Cyanosis results from increased desaturated hemoglobin.
6. Oximetry is a clinically useful measure of oxygen saturation.
7. Hemoglobin has much greater affinity for carbon monoxide than for oxygen.
8. Methemoglobinemia occurs in certain toxic states, notably nitrite poisoning.
9. Veterinarians frequently manage oxygen transport.

Carbon Dioxide Transport

1. Carbon dioxide is transported in the blood both in solution in plasma and in chemical combinations.
2. Carbon dioxide content of blood depends on P_{CO_2}, P_{O_2}, and buffering capacity.
3. Capnography is a clinically useful indicator of P_{CO_2}.

Gas Transport During Exercise

1. Oxygen demands of exercise are met by increases in blood flow, hemoglobin levels, and oxygen extraction from blood.

Oxygen Transport

Most Oxygen Is Transported in Combination With Hemoglobin, but a Small Amount Is in Solution in Plasma

Oxygen (O_2) is transported in the blood to tissues throughout the body to fuel aerobic metabolism. Animals consume approximately 5 milliliters (mL) of O_2 from every 100 mL of blood as it passes through capillaries. However, O_2 is poorly soluble and thus does not dissolve well in the plasma component of blood. Because of its low solubility, less than half a milliliter of O_2 is dissolved per deciliter (dL, 100 mL = 1 dL) of arterial blood under normal atmospheric conditions at sea level, so an additional method of transporting O_2 is required; the pigmented protein *hemoglobin* (Hb) is the primary transporter of O_2 in blood. Without hemoglobin, the cardiac output would have to be inordinately high to maintain the O_2 supply to the body organs. As a pigment, hemoglobin is bright red in color when its binding sites are saturated with O_2 and becomes purplish in color as it binds less O_2.

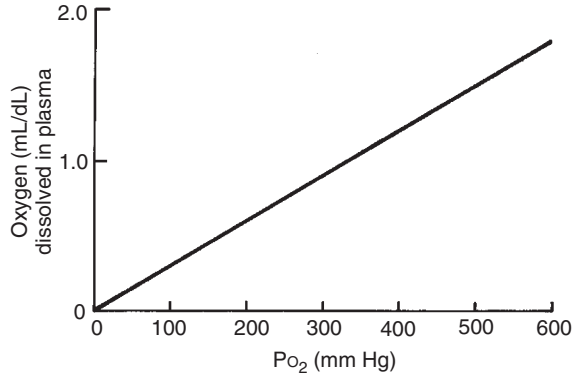
When blood in the pulmonary capillaries flows past ventilated alveoli, O_2 passively diffuses from the alveoli into the blood until the partial pressures equilibrate, that is, when there is no further difference between O_2 partial pressures in the alveoli and in the blood.

The partial pressure of O_2 (P_{O_2}) in a blood sample is a measure of O_2 *dissolved* in plasma. Even though the amount of O_2 dissolved in plasma is small, it increases directly as the P_{O_2} increases; 0.003 mL of O_2 dissolves in each 1 dL of plasma for every 1 mm Hg of O_2 partial pressure (Fig. 48.1). Under normal atmospheric conditions at sea level, the partial pressure of O_2 in arterial blood (P_aO_2) is approximately 100 mm Hg; therefore 0.3 mL of O_2 is dissolved in each deciliter of blood. If an animal breathes pure (100%) O_2, its P_aO_2 increases to at least 600 mm Hg and thus 1.8 mL of O_2 (600 mm Hg × 0.003 mL/dL/mm Hg) dissolves in each deciliter of plasma. Although this is a substantial increase in the amount of O_2 dissolved in the blood, it is still insufficient to meet the tissues' need for approximately 5 mL of O_2 per dL of blood.

A Molecule of Hemoglobin Can Reversibly Combine With Four Molecules of Oxygen

Mammalian hemoglobin consists of four units, each containing one *heme* and its associated protein, *globin.* The four globins, together, consist of two pairs of polypeptide chains whose amino acid sequences define the different types of mammalian hemoglobin. Adult hemoglobin contains two alpha (α) chains and two beta (β) chains; fetal hemoglobin contains two α chains and two gamma (γ) chains. *Heme* is a protoporphyrin consisting of four pyrroles with a ferrous iron (Fe^{2+}) at the center. Each ferrous iron can combine reversibly with a single molecule of O_2. The complete hemoglobin molecule has four hemes, each with its associated globin, and thus can combine reversibly with up to four molecules of O_2 (Fig. 48.2). When all four hemes of a hemoglobin molecule bind an O_2 molecule, that hemoglobin molecule is *fully saturated.* Under normal atmospheric conditions at sea level, the P_aO_2 is 100 mm Hg and the hemoglobin in arterial blood is nearly fully saturated. Mixed venous blood under these normal conditions has approximately a P_{vO_2} of 40 mm Hg and hemoglobin that is 75% saturated; this means that under normal conditions, on average, only a single O_2 molecule is removed from each hemoglobin to satisfy the tissues' need for O_2.

The *reversible* combining of O_2 with hemoglobin is essential to the *loading* of O_2 onto hemoglobin as blood passes alveoli in the lungs, and to *unloading* of O_2 from hemoglobin as blood travels through tissue capillaries. Loading and unloading of O_2 are each four-step processes, and the O_2 affinity of each heme is influenced by the oxygenation of the other three hemes in a hemoglobin molecule. This means that when the first heme unit is oxygenated, a conformational change in the hemoglobin occurs that increases the affinity of hemoglobin for the next O_2, such that the next O_2 loads onto hemoglobin more easily and binds to it with tighter affinity. Similarly, the third and then fourth O_2 each bind to hemoglobin with increasing ease and affinity; this stepwise process facilitates the loading of O_2 onto hemoglobin in alveolar capillaries. These *heme-heme* interactions are responsible for the sigmoidal shape of the oxyhemoglobin dissociation curve (Fig. 48.3). The reverse occurs at the tissues, where the unloading of the first O_2 molecule from hemoglobin results in decreased affinity of heme for the remaining O_2 so that the second O_2 molecule unloads more easily. Likewise, the third and then the fourth O_2 unload with increasing ease. However, the unloading of more than a single O_2 is unusual as under normal conditions only a single O_2 molecule from each hemoglobin needs to be unloaded at the tissues to meet metabolic demands. Therefore the second, third, and fourth O_2 molecules bound to each hemoglobin act as a reserve for when metabolic demands increase or when O_2 supplies are limited.

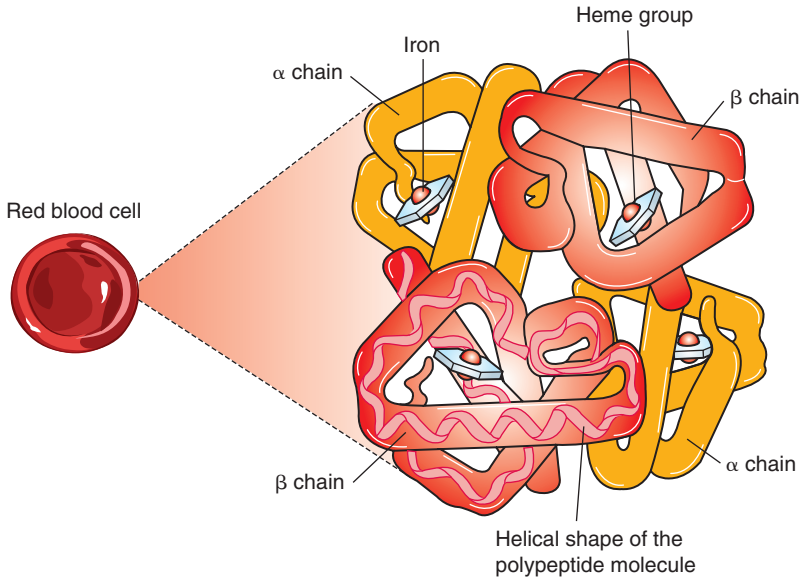


• **Fig. 48.1** Amount of O_2 dissolved in plasma as a function of P_{O_2}.

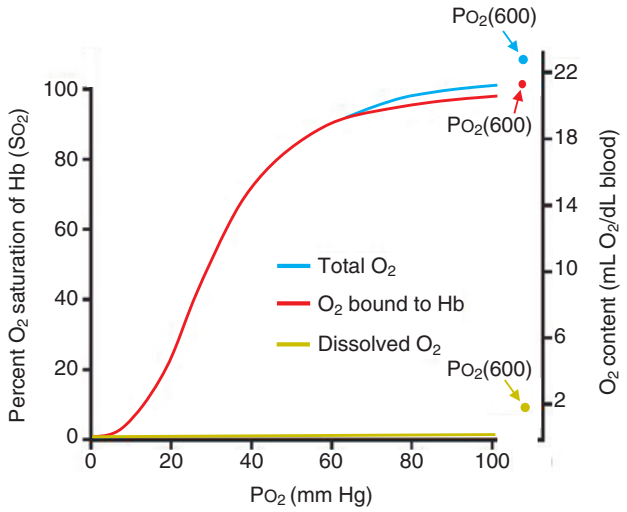
The Oxyhemoglobin Dissociation Curve Describes the Relationship Between Oxygen Partial Pressures and Hemoglobin Saturation

The oxyhemoglobin dissociation curve is a valuable tool for understanding plasma oxygenation in clinical patients (Fig. 48.3). It shows that, as P_{O_2} increases, the hemoglobin saturation with O_2 increases but in a nonlinear manner. The vertical axis of the oxyhemoglobin dissociation curve plots *oxygen saturation* (S_{O_2}), also termed *percent oxygen saturation of hemoglobin,* which is a measure of the overall percentage of binding sites on hemoglobin occupied by O_2. Although the S_{O_2} of any single hemoglobin molecule is limited to values of 0%, 25%, 50%, 75%, or 100% (corresponding to 0–4 O_2 molecules bound), the overall S_{O_2} values for arterial blood (S_{aO_2}) under normal conditions ranges from 96% to 98% and for mixed venous blood (S_{vO_2}) is approximately 75%. The nonlinear relationship between P_{O_2} and S_{O_2}, as seen in the sigmoidal shape of the oxyhemoglobin dissociation curve, is due to the heme-heme interactions that increase affinity for binding O_2 to hemoglobin with each successive O_2 molecule bound, as previously described. Blood passing oxygenated alveoli in the lungs encounters the physiological conditions of the oxyhemoglobin dissociation curve at its plateau, as P_{O_2} approaches 100 mm Hg and S_{O_2} approaches 100%, due to the facilitated binding of O_2 by hemoglobin in its close-to-saturated state. Conversely, the release of O_2 from hemoglobin at the tissues is facilitated by the progressively weaker binding between O_2 and hemoglobin at lower P_{O_2} values. This facilitated unbinding of O_2 is depicted by the altered slope of the oxyhemoglobin curve as it becomes significantly steeper at P_{O_2} values of about 40 mm Hg and below. Consequently, there is more unloading of O_2 from hemoglobin (hence lower S_{O_2} values) at P_{O_2} values in the range found in tissues and subsequently in mixed venous blood.

Oxygen saturation of hemoglobin is a function of P_{O_2} and is *independent of hemoglobin concentration* of blood. This independence



• **Fig. 48.2** Structure of the hemoglobin molecule showing the two pairs of globin polypeptide chains. Within each globin is a heme, four in all. The ferrous iron in the center of each heme provides the binding site for molecular O_2. (From Mader SS. *Inquiry into Life*. 8th ed. New York, NY: McGraw-Hill; 1997.)



• **Fig. 48.3** Oxyhemoglobin dissociation curve at normal physiological conditions of pH 7.4, P_{CO_2} 40 mm Hg, and 37°C body temperature. The slope of the curve is steep across P_{O_2} values typical of mixed venous blood (~40 mm Hg) and fairly flat across P_{O_2} values above approximately 70 mm Hg, which reflects facilitated hemoglobin unloading and loading of O_2, respectively (see text for further explanation). Modified from Boron WF, Boulpaep EL: Medical Physiology, 3rd ed, Philadelphia, 2017, Elsevier, with data from Severinghaus JW: Simple, accurate equations for human blood O_2 dissociation computations, J Appl Physiol. 1979;46(3):599–602.

• BOX 48.1 Partial Pressure of Oxygen in Inspired Air at Different Altitudes

$P_iO_2 = (P_{atm} - PH_2O) \times F_iO_2$
F_iO_2 = 21% in ambient air
P_{atm} = 760 mm Hg at sea level
PH_2O = 50 mm Hg at 38° C body temperature
$P_iO_2 = (760 - 50) \times 0.21 = 149$ mm Hg at sea level
P_{atm} = 508 mm Hg at 11,360 ft of altitude
$P_iO_2 = (508 - 50) \times 0.21 = 96$ mm Hg at 11,360 ft altitude

from hemoglobin concentration has important clinical implications as P_aO_2 and SaO_2 values are unaffected by the lower hemoglobin concentration in anemia, whereas the O_2 carrying capacity and O_2 content (described later) are decreased in anemia. Note that in mixed venous blood, O_2 saturation and partial pressure (PvO_2) may be decreased in states of severe anemia when an increased amount of O_2 must be unloaded at the tissues to meet metabolic needs.

At a P_{O_2} of more than approximately 70 mm Hg, the oxyhemoglobin dissociation curve plateaus, which indicates that further increases in P_{O_2} add little O_2 to hemoglobin (see Fig. 48.3). At this point, the hemoglobin is *fully saturated* with O_2 because each heme is binding an O_2 molecule. The fact that hemoglobin becomes virtually saturated with O_2 at a P_{O_2} of more than 70 mm Hg has important clinical implications. Many animals live at altitudes considerably above sea level, where the lower barometric pressure results in a low partial pressure of inspired O_2 (P_iO_2). Although these animals have a lower P_aO_2 than their sea level-dwelling counterparts, they are still able to transport sufficient O_2 to their tissues because their hemoglobin is sufficiently saturated with O_2. However, at extreme altitudes, P_iO_2 falls to the extent that hemoglobin cannot be adequately saturated with O_2 (Box 48.1); for example, at an altitude of 11,360 ft the atmospheric pressure is 508 mm Hg, so the P_iO_2 drops to 96 mm Hg, resulting in an alveolar O_2 partial pressure (P_AO_2) of 46 mm Hg (see Chapter 47) and an S_{O_2} of approximately 80% in the normal animal.

When the P_{O_2} falls to less than approximately 60 mm Hg, as occurs in tissues, the oxyhemoglobin dissociation curve has a steeper slope, which reflects the decreasing affinity of hemoglobin for O_2 as it begins to unload at the tissues. Tissue P_{O_2} varies in accordance with the blood flow to metabolism ratio of each tissue, but average tissue P_{O_2} is 40 mm Hg. Blood exposed to such a P_{O_2} loses approximately 25% of its O_2 to the tissues. In rapidly metabolizing tissues in which tissue P_{O_2} is lower, more O_2 is unloaded from the blood. The O_2 remaining in combination with hemoglobin forms a reserve that can be utilized in the event that metabolism increases, such as occurs during exercise.

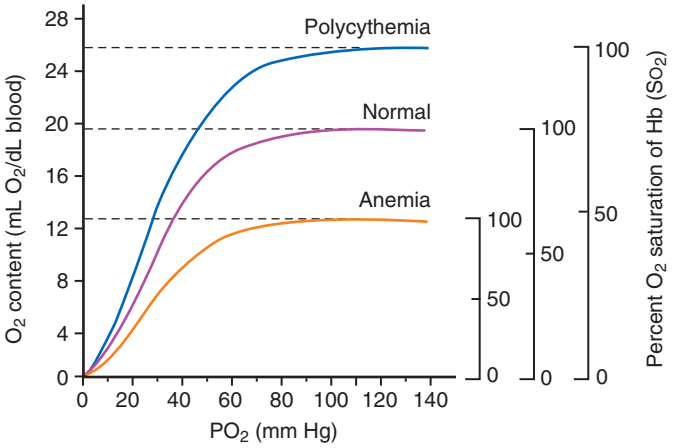
Along with O_2 saturation, there are several other indices of O_2 transport used to assess the different aspects of this process. *Oxygen carrying capacity* (or *oxygen capacity*) describes the amount of O_2 that can be "carried" by hemoglobin per deciliter of blood; it is a function of the amount of hemoglobin per deciliter of blood and the amount of O_2 that each hemoglobin can bind when fully saturated. Under normal conditions, each deciliter of blood contains approximately 15 g of hemoglobin, and each gram of hemoglobin binds 1.35 mL of O_2 (range is 1.34–1.39 mL/g), in which case, the O_2 carrying capacity is 15 g/dL × 1.35 mL/g = 20.25 mL O_2/dL blood. The O_2 carrying capacity is an indication of the maximum amount of O_2 that *can be* carried, and yet the amount that *is* carried also depends on how many of the four hemes in each hemoglobin have bound an O_2 molecule; this is termed *oxygen saturation* (S_{O_2}) as described earlier. Finally, the *oxygen content* (or *oxygen concentration*) of blood is a measure of the *total* amount of O_2 present in each deciliter of blood. Oxygen content is generally measured in arterial blood (CaO_2) and is determined by the O_2 carrying capacity, the O_2 saturation, and the amount of O_2 dissolved in blood, as follows:

$$\begin{aligned}&\text{Oxygen content of arterial blood}\\&= (O_2 \text{ carrying capacity} \times \%\text{ hemoglobin saturation})\\&\quad + \text{mL } O_2 \text{ in solution}\\&CaO_2 = ([\text{Hb g/dL} \times 1.35 \text{ mL } O_2/\text{g Hb}] \times SaO_2\%)\\&\quad + (0.003 \times PaO_2 \text{ mm Hg})\end{aligned}$$

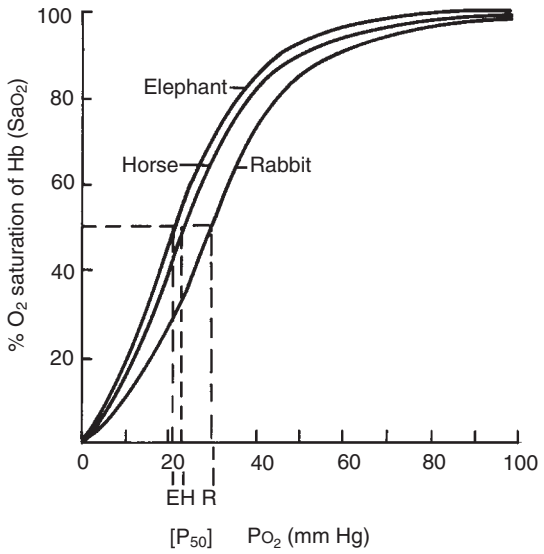
Using the normal arterial blood values of 15 g Hb/dL blood, 1 g hemoglobin carries 1.35 mL O_2; SaO_2 = 98%; 0.003 mL of O_2 dissolves/dL of plasma for every 1 mm Hg of O_2 partial pressure; and P_aO_2 = 100 mm Hg, yields the following results:

$$\begin{aligned}CaO_2 &= ([15 \text{ g/dL} \times 1.35 \text{ mL } O_2/\text{g Hb}] \times 98\%)\\&\quad + (0.003 \text{ mL} \times 100 \text{ mm Hg})\\&= 20.145 \text{ mL } O_2/\text{dL blood}\end{aligned}$$

Because the vast majority of O_2 is transported bound to hemoglobin, and under normal circumstances the hemoglobin in arterial blood is nearly saturated with O_2, the O_2 carrying capacity and CaO_2 are typically nearly the same in a normal individual. Fig. 48.4 shows the effect of both hemoglobin concentration and P_{O_2} on the O_2 content of blood. Hemoglobin concentration is a function of the erythrocyte concentration, termed *hematocrit* or *packed cell volume,* such that hemoglobin concentration is approximately one-third of the hematocrit. The reduced hematocrit of *anemia* results in decreased hemoglobin concentration and subsequent reductions in both O_2 carrying capacity and O_2 content.



• **Fig. 48.4** Effects of O_2 carrying capacity (hemoglobin concentration) and P_{O_2} on O_2 content (left vertical axis) in *normal* (Hb = 15 g/dL), *anemic* (Hb = 10 g/dL), and *polycythemic* (Hb = 20 g/dL) blood. O_2 saturation (right vertical axis) is determined by P_{O_2} but is not altered by O_2 carrying capacity.



• **Fig. 48.5** Oxyhemoglobin dissociation curves of three species. Although the curves have similar shapes in all mammals, they are not superimposed. The differences between the curves can be expressed by the partial pressure at which hemoglobin is 50% saturated with O_2 (P_{50}). P_{50} for each species is indicated as *E* (elephant), *H* (horse), and *R* (rabbit).

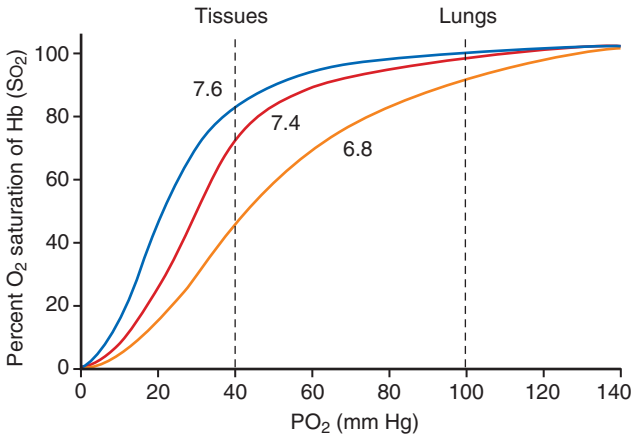
The opposite occurs when the hematocrit increases as in *polycythemia,* with the resultant increased O_2 carrying capacity and O_2 content. In some species, most notably horses and to a lesser extent dogs, the spleen stores red blood cells and contracts during exercise, releasing these additional erythrocytes and thereby increasing the O_2 carrying capacity, as well as the O_2 content, of blood.

In summary, changes in hemoglobin concentration will affect the O_2 content via altered O_2 carrying capacity, whereas changes in plasma P_{O_2} will affect O_2 saturation, which in turn affects O_2 content. Despite the fact that most O_2 is carried in blood bound to hemoglobin, it is only the O_2 dissolved in plasma that determines the P_{O_2}. It is important to understand that P_{O_2} values measured in a blood gas analysis do not indicate hemoglobin concentration.

The Affinity of Hemoglobin for Oxygen Varies With Blood Temperature, pH, Carbon Dioxide Partial Pressure, and 2,3-BPG Concentration

Although all mammals have similarly shaped oxyhemoglobin dissociation curves, the position of the curve with regard to P_{O_2} varies with species (Fig. 48.5) and physiological conditions (Fig. 48.6). The *position* of the curve can be described by measurement of P_{50}, the O_2 partial pressure at which hemoglobin is 50% saturated with O_2. The P_{50} is a measure of the *affinity* of hemoglobin for O_2. Normal values for P_{50} in domestic animals range from a P_{O_2} of approximately 24 to 34 mm Hg, with larger species having somewhat lower values (Fig. 48.5). An increased P_{50} indicates that a higher P_{O_2} is necessary to maintain 50% O_2 saturation, which is reflected in a shift of the oxyhemoglobin saturation curve to the right, termed *right shift.* A right shift in the curve reflects facilitated unloading of O_2 at the tissues. Conversely, a decreased P_{50} indicates that 50% O_2 saturation is maintained at a lower plasma P_{O_2}, with a corresponding *left shift* in the oxyhemoglobin dissociation curve. A left shift of the curve reflects facilitated loading of O_2 onto hemoglobin in the lungs.

The position of the oxyhemoglobin curve is dynamic as it continually responds to local physiological conditions. An increase in tissue metabolism produces heat, which elevates blood temperature and shifts the oxyhemoglobin dissociation curve to the right (increases P_{50}). Such a shift reflects facilitated dissociation of



• **Fig. 48.6** Effect of pH on the oxyhemoglobin dissociation curve. A decrease in pH shifts the dissociation curve to the right, and an increase in pH shifts the curve to the left. The shifts in the dissociation curve reflect much less effect on the percent saturation of hemoglobin when O_2 is being loaded into the blood in the lungs (i.e., P_{O_2} = 100 mm Hg) than when O_2 is being unloaded in the tissues (i.e., P_{O_2} = 40 mm Hg).

O_2 from hemoglobin and release of O_2 to the tissues; hemoglobin is then said to have "less affinity" for O_2. Conversely, cooling of the blood, as occurs in the lungs (relative to exercising muscle) or in the tissues in states of hypothermia, shifts the dissociation curve to the left (decreases P_{50}). Because of this increased affinity of hemoglobin for O_2, tissue P_{O_2} must be lower than usual to release O_2 from hemoglobin.

Changes in carbon dioxide (CO_2) partial pressure (P_{CO_2}) and pH also affect the affinity of hemoglobin for O_2. The shift in the oxyhemoglobin dissociation curve resulting from a change in P_{CO_2} is called the *Bohr effect.* This shift results in part from the combination of CO_2 with hemoglobin but is mostly due to the decrease in pH from the production of hydrogen ions, which follows the chemical reaction of CO_2 in the bicarbonate buffer system (see later and Chapter 52). A change in pH alters the O_2 binding by

changing the structure of hemoglobin. As a result, a higher, more alkaline pH shifts the oxyhemoglobin dissociation curve to the left, and a lower, more acidic pH shifts the curve to the right (Fig. 48.6). The Bohr effect is not constant among species; a given change in pH produces a greater shift in the oxyhemoglobin dissociation curve for small mammals, which have higher basal metabolic rates in relation to their body mass than for large mammals. This ensures the delivery of adequate O_2 to species with high metabolic activity.

The final normal physiological metabolite that affects the affinity of hemoglobin for O_2 is 2,3-bisphosphoglycerate (2,3-BPG, also known as 2,3-diphosphoglycerate [2,3-DPG]). Erythrocytes lack mitochondria and thus rely on glycolysis for energy production; 2,3-BPG is a byproduct of glycolysis. 2,3-BPG binds to hemoglobin in a way that stabilizes the hemoglobin molecule in the absence of O_2 and thus decreases the affinity of hemoglobin for O_2 as reflected in an increased P_{50} and right shift of the oxyhemoglobin curve. Glycolysis is stimulated by reduced P_{O_2}, therefore 2,3-BPG production is increased by chronic hypoxia, anemia, and acclimation to high altitude with the resultant facilitation of unloading O_2 at the tissues. In contrast, a reduction in 2,3-BPG levels, which occurs in stored blood, shifts the dissociation curve to the left. Not all forms of hemoglobin bind 2,3-BPG equally; the hemoglobin of horses, dogs, rabbits, guinea pigs, and rats contain high levels of bound 2,3-BPG whereas ruminants and cats have low levels of bound 2,3-BPG. There appears to be some further species variation in terms of whether higher levels of 2,3-BPG necessarily equate to increased action of modifying hemoglobin-O_2 affinity. Finally, the gamma chains of fetal hemoglobin do not bind 2,3-BPG, making fetal hemoglobin largely unresponsive to 2,3-BPG. This results in left shift of the fetal oxyhemoglobin curve, which is an important adaptation of the fetus to the normal in utero conditions of lower P_{O_2} compared with adult P_{O_2} values (see Chapter 51).

Cyanosis Results From Increased Desaturated Hemoglobin

The change in the color of hemoglobin from red to purple as desaturated hemoglobin increases is known as *cyanosis.* It is observed in the mucous membranes of animals when the blood in the underlying capillaries contains an increased amount of desaturated hemoglobin, which typically occurs as O_2 saturation decreases. Because cyanosis results from the purplish color of desaturated hemoglobin rather than from a reduction in the number of saturated hemoglobin molecules, it may be more pronounced in patients with polycythemia (where there are more hemoglobin molecules that can become desaturated) and less evident in anemic patients. Cyanosis can be a feature of a variety of respiratory and circulatory abnormalities.

Oximetry Is a Clinically Useful Measure of Oxygen Saturation

Oxygen saturation can be measured directly from arterial blood (Sa_{O_2}) or indirectly at peripheral locations such as the tongue, lip, pinnae, or hairless minimally pigmented skin fold by means of pulse oximetry (Sp_{O_2}). Oximeters distinguish oxygenated from desaturated (partially oxygenated) hemoglobin based on differences in their absorption of red and infrared wavelength light, which is emitted from a diode and transmitted to a detector. Because pulse oximetry is noninvasive, it is clinically useful for measuring O_2 saturation and, by means of the oxyhemoglobin dissociation curve, predicting P_aO_2. However, oximeters based on dual (red and infrared) wavelength technology do not provide accurate readings in the presence of *dyshemoglobinemias,* such as carboxyhemoglobin and methemoglobin. Dyshemoglobins are hemoglobin derivatives that do not reversibly bind O_2. Cooximetry is a type of oximetry that measures multiple wavelength absorptions to distinguish the various forms of hemoglobin including oxyhemoglobin, deoxyhemoglobin, as well as the dyshemoglobins carboxyhemoglobin and methemoglobin, as described later. Cooximetry is performed on blood and pulse cooximetry is performed noninvasively, like pulse oximetry.

Hemoglobin Has Much Greater Affinity for Carbon Monoxide Than for Oxygen

Carbon monoxide (CO) is a gas generated by the burning of carbon-based fuels. In veterinary species, CO toxicity is often associated with smoke inhalation. Carbon monoxide binds to the same sites on hemoglobin as does O_2, forming *carboxyhemoglobin.* The affinity of hemoglobin for CO far exceeds the affinity of hemoglobin for O_2—more than 200-fold greater—and as a result exposure to CO levels of less than 1% can saturate hemoglobin and displace O_2, leading to death. Carbon monoxide not only displaces O_2 bound to hemoglobin, it also shifts the oxyhemoglobin dissociation curve far to the left. This exaggerated left shift in the presence of CO means that hemoglobin has a much higher affinity for the remaining bound O_2 such that O_2 is not readily released at the tissues, resulting in tissue hypoxia. Carboxyhemoglobin is not distinguishable from oxyhemoglobin by standard pulse oximetry, resulting in Sp_{O_2} readings that are falsely high in CO toxicity. Thus cooximetry or pulse cooximetry is superior at detecting CO toxicity. Treatment of CO toxicity requires removal of the CO source and administration of O_2 to displace the CO from hemoglobin.

Methemoglobinemia Occurs in Certain Toxic States, Notably Nitrite Poisoning

When the normal ferrous iron (Fe^{2+}) of hemoglobin is *oxidized* by *nitrites* and other toxins to ferric iron (Fe^{3+}), brown-colored *methemoglobin* is formed. Methemoglobin does not bind O_2, thus the O_2 saturation and O_2 carrying capacity of the blood are functionally reduced. Methemoglobin reductase is an enzyme present in erythrocytes that functions to reduce the normally small amounts of methemoglobin back to hemoglobin; this enzyme is inadequate in the presence of excessive oxidation, at which point methemoglobinemia results. Methemoglobin production in companion animals can result from ingestion of a variety of common items including acetaminophen, garlic, onions, naphthalene (moth balls), paradichlorobenzene (toilet bowl cleaner), and in ruminants from ingestion of spoiled feeds or nitrate-rich feeds such as Sudan grass. Pulse oximetry is inaccurate in the presence of methemoglobin because methemoglobin absorbs red and infrared light wavelengths equally, resulting in default Sp_{O_2} readings of 85%, regardless of the true O_2 saturation of hemoglobin values. Thus cooximetry and pulse cooximetry are superior in distinguishing methemoglobin from oxy- and deoxyhemoglobin.

Veterinarians Frequently Manage Oxygen Transport

Use of inhalant anesthesia and managing anemia, polycythemia, lung disease, and dyshemoglobinemias all require an understanding

of O_2 transport and the oxyhemoglobin dissociation curve. In a healthy animal (blood Hb = 15 g/dL) breathing air containing 21% O_2 (fraction of inspired O_2 [F_iO_2] = 0.21), the hemoglobin is virtually saturated with O_2 at the normal P_aO_2 of 85 to 100 mm Hg (Table 48.1). Anesthetizing a healthy animal with a gas mixture containing almost 100% O_2 (F_iO_2 = 1.0) raises the P_aO_2 to over 600 mm Hg, but this has no effect on O_2 carrying capacity (no more hemoglobin has been added to the blood) and only raises O_2 content slightly; no more O_2 can be added to hemoglobin but a little more is dissolved in solution in plasma. In contrast, breathing less than 21% O_2 (for example, when F_iO_2 is 0.15) decreases P_aO_2 and reduces both O_2 saturation and O_2 content. The same situation arises when P_aO_2 decreases because of lung disease or ascent to altitudes of about 1500 m (5000 ft) or greater.

Anemic animals have hemoglobin deficiency and therefore have reduced O_2 carrying capacity. Unless they also have lung disease, their P_aO_2 is normal. Their O_2 content is decreased because of the low O_2 carrying capacity, but because P_aO_2 is normal, all the hemoglobin is saturated so SaO_2 and SpO_2 are normal. Polycythemia is simply the converse situation; because of their higher hematocrit and hemoglobin concentration, the O_2 carrying capacity and O_2 content are increased in polycythemic animals. However, because P_aO_2 and SaO_2 are essentially at maximal values under normal conditions, they do not change in polycythemia.

It takes only a very small fraction of CO in air to cause poisoning yet this is not sufficient to measurably reduce the F_iO_2, consequently P_aO_2 is not decreased. CO does not reduce the amount of hemoglobin but by binding to it so avidly, it reduces the amount of hemoglobin available for O_2 binding.

Carbon Dioxide Transport

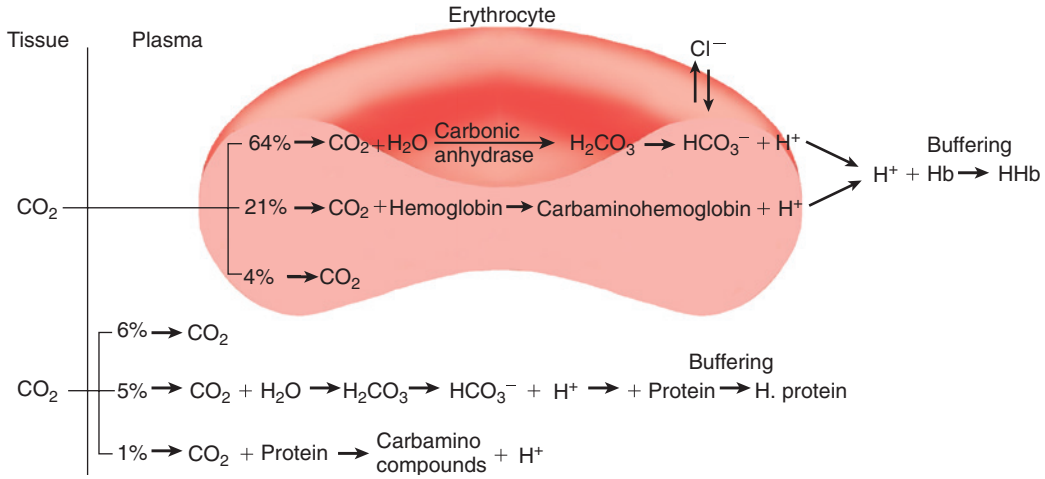
Carbon Dioxide Is Transported in the Blood Both in Solution in Plasma and in Chemical Combinations

Unlike O_2, which is bound only to hemoglobin, CO_2 is transported in the blood in several forms (Fig. 48.7). CO_2 is produced in tissues; therefore tissue PCO_2 is higher than the PCO_2 of the blood arriving in the capillaries. CO_2 diffuses down a concentration gradient from the tissues into the blood. When the blood leaves the tissues, its PCO_2 has risen from approximately 40 mm Hg to 46 mm Hg, with exact values depending on the tissue-specific ratio of blood flow to metabolism.

The majority of CO_2 diffuses into erythrocytes, where it undergoes one of two chemical reactions to form either carbonic acid or carbaminohemoglobin. Most (64%) of the CO_2 combines

TABLE 48.1 Effect of Some Common Situations on Arterial Oxygen Partial Pressure (P_aO_2), Oxygen Carrying Capacity, Percent Saturation of Hemoglobin (SaO_2), and Oxygen Content

	P_aO_2 (mm Hg)	O_2 Carrying Capacity (mL O_2/dL blood)	SaO_2 (%)	O_2 Content (mL O_2/dL blood)
Breathing ambient air (21% O_2)	Normal (85–100)	Normal (~20.25)	~98	Normal (~20.145)
Breathing 100% O_2	Increased (~600) Less rise with lung disease	Normal	~99	Slightly increased (~22)
Breathing 15% O_2	Reduced (~55)	Normal	Reduced (~88)	Reduced (~18)
Anemia	Normal	Reduced	Normal	Reduced
Polycythemia	Normal	Increased	Normal	Increased



• **Fig. 48.7** Forms of CO_2 transport in the blood. All reactions displayed in this diagram are reversed when the blood reaches the lung and CO_2 diffuses into the alveoli. *H_2CO_3*, Carbonic acid; *HCO_3^-*, bicarbonate.

with water and forms *carbonic acid* (H_2CO_3), which then dissociates into *bicarbonate* (HCO_3^-) and *hydrogen ion* (H^+), as follows:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

The hydration of CO_2 to form H_2CO_3 occurs very rapidly in erythrocytes due to the presence of the catalyst *carbonic anhydrase* (CA), which accelerates the reaction rate several hundredfold. Ionization of H_2CO_3 occurs rapidly to produce H^+ and HCO_3^- within the erythrocytes. These reversible reactions are kept moving to the right because H^+ is *buffered,* that is, removed from solution by binding to hemoglobin (HHb). Similarly, the gradient that drives the movement of CO_2 from tissues to the capillary blood must be maintained; it is facilitated by the fact that deoxyhemoglobin is a better buffer than oxyhemoglobin, by nature of it being a weaker acid (see Chapter 52). Thus deoxyhemoglobin, which is generated as blood passes through the tissue capillaries, combines readily with H^+ and facilitates the formation of HCO_3^- from CO_2, which decreases CO_2 concentration in the blood. Most of the HCO_3^- that is produced in the erythrocyte diffuses out of cells down a concentration gradient into the plasma. Chloride ion (Cl^-) diffuses into the erythrocyte as HCO_3^- diffuses out to maintain the electrochemical neutrality.

The reactions $CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$ also occur in plasma, accounting for another 5% of CO_2 transport, but the reaction is much slower due to the decreased proximity to carbonic anhydrase.

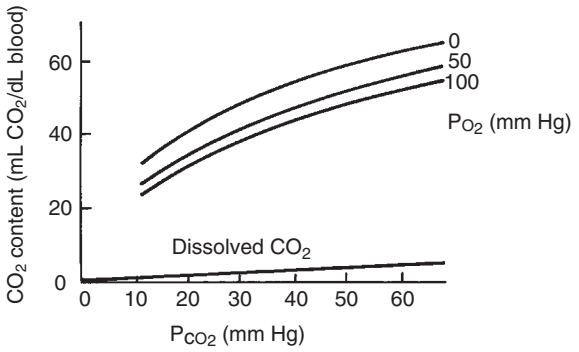
The *carbamino compounds* are the second most common form in which CO_2 is transported in the blood. Carbamino compounds are formed by coupling of CO_2 to the amino (–NH) groups of proteins, particularly hemoglobin. Note that CO_2 in its binding to amino groups does *not* bind to the same position on hemoglobin as does O_2, which binds the heme moiety. Approximately 21% of CO_2 is transported as carbaminohemoglobin within erythrocytes, whereas another 1% of CO_2 is transported as other carbamino proteins that form in the plasma.

Finally, approximately 6% of the CO_2 entering the blood is transported dissolved in the plasma, whereas the remaining 4% of CO_2 is transported dissolved in the cytosol of erythrocytes.

When venous blood reaches the lungs, CO_2 diffuses down its concentration gradient from the plasma and erythrocytes into the alveoli where it is ultimately expired, thus causing the reactions shown in Fig. 48.7, which are all reversible, to move to the left. Simultaneously, the oxygenation of hemoglobin releases H^+ ions, which combine with HCO_3^- to form H_2CO_3 followed by dehydration to $CO_2 + H_2O$.

Carbon Dioxide Content of Blood Depends on P_{CO_2}, P_{O_2}, and Buffering Capacity

The blood *content of carbon dioxide* as a function of P_{CO_2} is depicted in the CO_2 equilibrium (a.k.a. dissociation) curves shown in Fig. 48.8. Curves are shown for oxygenated blood (P_{O_2} = 100 mm Hg), for partially deoxygenated blood (P_{O_2} = 50 mm Hg), and for deoxygenated blood (P_{O_2} = 0 mm Hg). In contrast to the oxyhemoglobin dissociation curve, the CO_2 curves are almost linear and have no plateau in the physiological range; thus CO_2 can be added to the blood as long as buffering capacity is available. As P_{O_2} or hemoglobin saturation falls, the blood CO_2 content increases. This higher CO_2 content of deoxygenated blood results from the *higher buffering capacity* of deoxyhemoglobin, compared with O_2-saturated hemoglobin.



• **Fig. 48.8** CO_2 equilibration curves showing the amount of CO_2 contained in the blood (CO_2 content) as a function of P_{CO_2}. Curves are shown for dissolved CO_2 and for total CO_2 content at P_{O_2} of 0, 50, and 100 mm Hg.

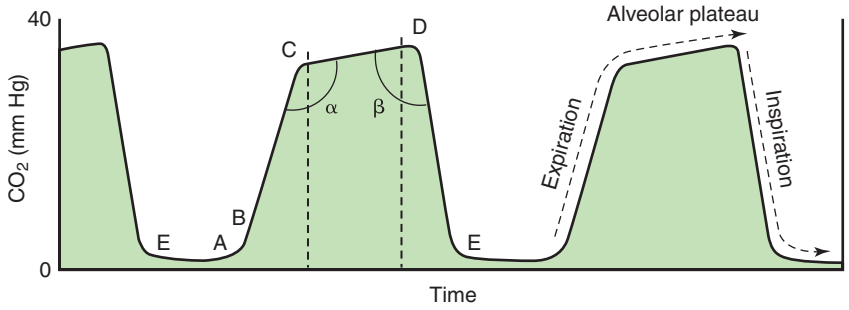
Capnography Is a Clinically Useful Indicator of P_{CO_2}

Capnography is a noninvasive measure of P_{CO_2} in breath used to assess *end tidal* CO_2 (Fig. 48.9); this is the partial pressure of CO_2 in the last portion of expired air of each breath, which contains alveolar air. Under normal circumstances, end tidal CO_2 values are very similar to but slightly lower than P_aCO_2 (up to 5 mm Hg difference). The normal difference between end tidal CO_2 and P_aCO_2 is due to the small amount of alveolar dead space ventilation, in which ventilated alveoli are not perfused so CO_2 does not enter the air of the affected alveoli (see Chapters 45 and 47). Furthermore, the gradient between end tidal CO_2 and P_aCO_2 can be a useful diagnostic aid for various respiratory and circulatory abnormalities. For example, in uncomplicated hypoventilation where CO_2 production exceeds alveolar ventilation, the CO_2 increase in the blood results in a similar increase in CO_2 in alveolar air; thus the end tidal CO_2 to P_aCO_2 difference remains unchanged. In contrast, the end tidal CO_2 to P_aCO_2 difference is elevated by increases in alveolar dead space ventilation such as result from pulmonary thromboembolism, ventilation to perfusion mismatching, hypovolemia, and decreased cardiac output, among other conditions. Obstructive lung diseases that limit exhalation also increase the end tidal CO_2 to P_aCO_2 gradient.

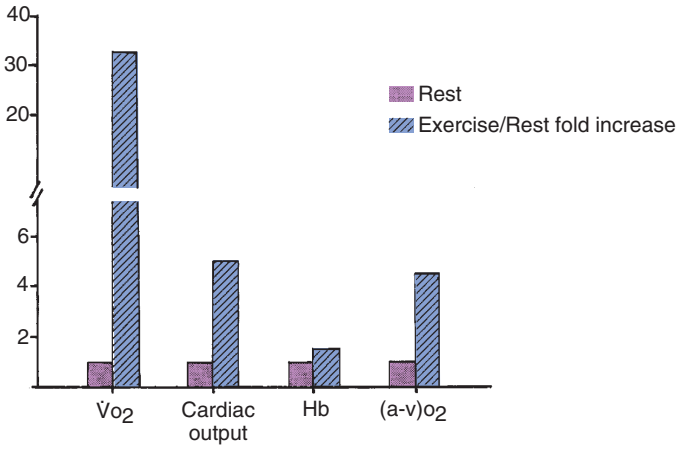
Gas Transport During Exercise

Oxygen Demands of Exercise Are Met by Increases in Blood Flow, Hemoglobin Levels, and Oxygen Extraction From Blood

The demands for gas transport in the blood are not constant but vary with metabolism. Strenuous exercise represents the most extreme demand placed on the gas transport mechanisms. In the galloping horse, O_2 consumption can increase to 30 times resting levels. Fig. 48.10 shows how this extra demand for O_2 is met. Part of the demand is provided by an increase in *cardiac output,* which causes the amount of blood flowing through the lungs per minute to increase. This allows an increased uptake of O_2 from the lungs. The cardiac output is also redistributed, with an increased fraction of the blood going to the exercising muscles. The increase in cardiac output and redistribution increases muscle blood flow by 20-fold.



• **Fig. 48.9** Normal capnogram waveforms. Capnography measures CO_2 at the airway opening throughout the respiratory cycle. During inspiration capnography CO_2 levels approach 0, reflecting the minimal CO_2 in inhaled air. During expiration CO_2 levels increase as air that has participated in gas exchange is exhaled. CO_2 falls quickly in early inspiration (D to E) due to the mixing of residual alveolar gas and fresh inspired air. CO_2 remains low for the remainder of inspiration (E to A). CO_2 levels rise quickly during early expiration (B to C) as dead space air that did not participate in gas exchange is exhaled followed by exhalation of alveolar gas that has equilibrated with pulmonary capillary P_{CO_2}. The rise in CO_2 then slows until the very end of expiration (D) during which undiluted alveolar air is exhaled. End tidal CO_2 is the value of CO_2 measured at the end of expiration of a tidal (normal resting) breath (D), as this portion of expired air best reflects alveolar gas. (Figure courtesy Dr. Andrew Claude, Department of Small Animal Clinical Sciences, Michigan State University).



• **Fig. 48.10** Oxygen consumption ($\dot{V}O_2$), cardiac output, hemoglobin level (Hb), and arteriovenous O_2 difference [(a-v)O_2] in a horse at rest and during strenuous exercise at a gallop. The thirtyfold increase in $\dot{V}O_2$ is accomplished by a fivefold increase in cardiac output, a 50% increase in hemoglobin, and a fourfold increase in (a-v)O_2.

Exercising horses also meet the increased O_2 demand with an increase in the number of circulating erythrocytes, and therefore an increased amount of hemoglobin. Contraction of the muscular spleen forces stored erythrocytes into the circulation and can increase the hematocrit from 35% to 50%. This provides almost 50% more binding sites for O_2, which raises the O_2 carrying capacity of the blood. The usefulness of an increase in hematocrit is limited because it increases blood viscosity, which tends to slow the flow of blood through the capillaries and increase the work of the heart. The increase in muscle blood flow and hematocrit together increase the delivery of O_2 to the muscle. An exercising muscle extracts a larger percentage of the O_2 from the blood than does a muscle at rest. This is accomplished as follows: (1) the diffusion gradient for O_2 is increased by the decrease in muscle P_{O_2}, which results from the increase in aerobic metabolism, and (2) the affinity of hemoglobin for O_2 is decreased (as seen by a rightward shift in the oxyhemoglobin dissociation curve) by the higher temperature of the exercising muscle and by the lower pH that results from release of CO_2 and hydrogen ions from the muscle. As a result of the increased extraction of O_2 from capillary blood as it perfuses the tissues, the arteriovenous O_2 content difference is increased.

Muscle itself contains the O_2-binding pigment, *myoglobin,* which provides a small store of O_2. However, myoglobin's main function is the transfer of O_2 within the muscle cell. Myoglobin, like hemoglobin, is an iron-containing pigment, but unlike hemoglobin, it contains only one heme group. As a result, the dissociation curve is not sigmoid but rather hyperbolic. The affinity of myoglobin for O_2 is high at P_{O_2} values typically found in tissues; for example, myoglobin is about 75% saturated with O_2 at a P_{O_2} of 20 mm Hg, whereas hemoglobin is only about 25% saturated with O_2 at P_{O_2} of 20 mm Hg. In contrast, myoglobin releases O_2 readily in the presence of very low P_{O_2} due to the steep slope of its O_2 dissociation curve between P_{O_2} values of 0 and about 5 mm Hg. As a result of these dissociation characteristics, myoglobin releases O_2 only when intracellular P_{O_2} is very low. Myoglobin is more plentiful in slow-twitch aerobic muscle fibers, which are capable of sustained aerobic work, than in fast-twitch fibers, which are generally used for short bursts of anaerobic activity. The amount of myoglobin is increased by exercise training.

In summary, in exercise the increased demand for O_2 is met by increases in blood flow, hematocrit, O_2 extraction from blood, and, to a small degree, by O_2 release from myoglobin. These mechanisms are available whenever unusual demands for gas exchange arise. In anemia, for example, O_2 carrying capacity is reduced, but O_2 delivery to the tissues can be preserved somewhat by an increase in cardiac output and increased extraction of O_2 from hemoglobin.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Flea Infestation in a Cat

History

A cat is presented to you because the owner notices that the cat seems weak and recently has been staggering when it walks around the house. The cat's appetite is good and besides the weakness, the owner reports the cat is normal.

Clinical Examination

Inspection of the animal, which is resting quietly on the examination table, shows that it is in reasonably good condition. The respiratory rate does not appear to be elevated and from a distance the cat shows no obvious signs of disease. When you place your hands on the cat's back, you immediately notice a gritty material in the fur. Further examination of the skin shows accumulations of a reddish-brown material deep within the coat and you notice many fleas scurrying around when the coat is parted. When you moisten some of the gritty material, it produces a red liquid. The cat's mucous membranes are almost white, and the examination of the mucous membranes produces sufficient struggling that the cat begins to breathe rapidly. The cat's pulse rate is elevated, but the lung sounds are normal. On physical examination, all the body systems appear to be normal. You take a blood sample; the hematocrit is 10% (normal, 30%–45%).

Comment

This cat has a severe infestation of fleas. The gritty material in the fur is flea feces, which contain blood products that become red when wet. The infestation is further confirmed by the observation of many fleas in the coat. By their blood-sucking method of feeding, fleas can produce anemia when they are present in large numbers, as with this cat. If the flea infestation develops gradually, the anemia is slow in onset, and the host animal may show few clinical signs until the infestation and anemia become severe. The anemia is confirmed in this case by the paleness of the mucous membranes and the low hematocrit. The rapid heart rate is a response to the anemia. To deliver sufficient O_2 to the tissues, the cardiac output needed to be increased by increasing the heart rate. When the cat is stressed by your examination, it shows signs of respiratory distress because there is inadequate O_2 delivery to the tissues; this causes production of lactic acid as a result of anaerobic metabolism. The resultant decrease in pH stimulates the chemoreceptors, causing the signs of respiratory distress.

Treatment

Your treatment approach to this cat is twofold. First, you administer blood to increase the cat's hematocrit and provide it with sufficient O_2 carrying capacity until it can generate new erythrocytes. Second, you treat the flea infestation and instruct the owner on how to remove fleas from the house.

Several weeks later the owner returns with the cat and notes that she has had no further problems. She is diligent about regular vacuuming to remove flea eggs and larvae from the house.

Atrial Fibrillation in a Horse

History

The owner of a 3-year-old Standardbred gelding is concerned because the horse is no longer able to complete its training program. Up until a week ago, the horse had been performing well during its daily training. In the past 2 days, the horse has been reluctant to exercise and, if pushed to do so, begins to stagger and appears weak in the rear legs.

Clinical Examination

Inspection of the horse reveals a normal appearing Standardbred in excellent condition. It is standing in its box stall, eating, and looks alert when you enter the stall. Clinical examination reveals normal colored mucous membranes, no abnormality of lung sounds, and no abnormalities in the gastrointestinal, urinary, or nervous system. When you take the pulse, you note that it is irregular in both amplitude and rate. Several pulses follow one another rapidly and then there are prolonged pauses with no consistent pattern to the irregularity. Auscultation of the heart reveals a similar irregularity in the heart sounds.

You take a blood sample for measurement of the hematocrit, which is normal. You also obtain an electrocardiogram (ECG), which reveals a continuous irregular pattern of multiple P waves with fewer and irregularly occurring QRS complexes.

Comment

The history, heart rhythm, and ECG findings in this horse are typical of atrial fibrillation. The multiple P waves observed on the ECG are a result of circuitous depolarization of the atria. In atrial fibrillation, the atria contract and relax in an uncoordinated manner. The atrioventricular node is activated at intervals that vary from cycle to cycle; thus there is no constant interval between ventricular contractions. The variable time between ventricular contractions allows for variable degrees of ventricular filling and therefore results in uneven stroke volume; consequently, the pulse varies in amplitude as well as frequency.

The irregular ventricular rhythm may be sufficient to maintain cardiac output in the resting animal, but during exercise sufficient increase in cardiac output cannot be generated. As a result, O_2 delivery to the muscles is inadequate to sustain exercise. This is an example of a failure of O_2 delivery resulting from inadequate blood flow.

Treatment

Treatment for atrial fibrillation in the horse can be pursued with either electrical cardioconversion or the administration of quinidine sulfate. Quinidine sulfate has a negative inotropic effect on the myocardium and slows atrioventricular node conduction time, which allows the reestablishment of normal atrial and ventricular rhythm. With successful treatment, the horse's heart rate will likely return to normal if there is no underlying cardiac disease. This horse was treated with quinidine sulfate and rested for a week, at which time training was reinstituted. Several months later, the owner reports that the horse is doing well.

This horse is a young adult with no evidence of cardiac disease (e.g., murmur, signs of heart failure). Treating a horse with underlying cardiac disease has associated risks and the horse may not return to normal heart rhythm. Therefore in most cases, an echocardiogram is recommended before treatment to determine that the horse has no underlying cardiac disease. If present, any disease should be addressed first. In some horses the underlying disease is more significant and treatment of fibrillation will not be attempted.

Practice Questions

1. If 1 g of hemoglobin has an O_2 capacity of 1.36 mL of O_2, what is the O_2 content of blood containing 10 g of hemoglobin when the blood P_{O_2} is 70 mm Hg?
- **a.** 13.6 mL O_2/dL blood
- **b.** 9.5 mL O_2/dL blood
- **c.** 6.8 mL O_2/dL blood
- **d.** 21 mL O_2/dL blood
- **e.** Cannot be calculated from the information provided

2. An increase in pH of blood will:
- **a.** Shift the oxyhemoglobin dissociation curve to the right.
- **b.** Decrease P_{50}.
- **c.** Decrease the affinity of hemoglobin for O_2.
- **d.** Decrease the O_2 capacity of the blood.
- **e.** All the above.

3. Which of the following decreases O_2 content but does not alter P_aO_2 or percentage saturation of hemoglobin?
- **a.** Ascent to an altitude of 3500 m
- **b.** Polycythemia
- **c.** Breathing 50% O_2
- **d.** Anemia
- **e.** Development of a large right-to-left shunt

4. All the following shift the oxyhemoglobin dissociation curve to the right except:
- **a.** An increase in pH
- **b.** An increase in P_{CO_2}
- **c.** An increase in 2,3-BPG
- **d.** An increase in temperature

5. Quantitatively, the most important form of CO_2 in blood is:
- **a.** HCO_3^- produced in plasma
- **b.** CO_2 dissolved in plasma
- **c.** HCO_3^- produced in the erythrocyte
- **d.** CO_2 dissolved in the erythrocyte
- **e.** CO_2 combined with plasma proteins

6. Oxygenation of hemoglobin in the lungs assists with the release of CO_2 from the blood because:
- **a.** O_2 combines with the –NH groups on hemoglobin and displaces CO_2 from carbamino compounds.
- **b.** O_2 combines with HCO_3^- and produces CO_2.
- **c.** O_2 facilitates the movement of chloride ions out of the erythrocyte.
- **d.** O_2 combines with hemoglobin, making it a better buffer, which retains H^+.
- **e.** None of the above.

Bibliography

Bartels H. Comparative physiology of oxygen transport in mammals. *Lancet*. 1964;2(7360):601–604.

Bliss S. Anemia and oxygen delivery. *Vet Clin North Am Small Anim Pract*. 2015;45:917–930.

Boron WF. Transport of oxygen and carbon dioxide in the blood. In: Boron WF, Boulpaep EL, eds. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Brauner CJ, Wang T. The optimal oxygen equilibrium curve: a comparison between environmental hypoxia and anemia. *Amer Zool*. 1997;37:101–108.

Bunn HF. Differences in the interaction of 2,3-diphosphoglycerate with certain mammalian hemoglobins. *Science*. 1971;172(3987):1049–1050.

Burggren W, McMahon B, Powers D. Respiratory functions of blood. In: Prosser CL, eds. *Comparative Animal Physiology: Environmental and Metabolic Animal Physiology*. 4th ed. New York: Wiley-Liss; 1991.

Collins J-A, Rudenski A, Gibson J, Howard L, O'Driscoll R. Relating oxygen partial pressure, saturation and content: the haemoglobin-oxygen dissociation curve. *Breathe*. 2015;11(3):194–201.

Haskin SC. Hypoxemia. In: Silverstein D, Hopper K, eds. *Small Animal Critical Care Medicine*. 2nd ed. Elsevier; 2015.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Kitchen H, Brett I. Embryonic and fetal hemoglobin in animals. *Ann N Y Acad Sci*. 1974;241(0):653–671.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lekeux P, Art T, Hodgson DR. The respiratory system: anatomy, physiology and adaptations to exercise and training. In: Hodgson DR, McGowan CM, eds. *The Athletic Horse: Principles and Practice of Equine Sports Medicine*. St. Louis: Saunders; 2014.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

West JB, Luks AM. *West's Pulmonary Pathophysiology: The Essentials*. 9th ed. Philadelphia: Wolters Kluwer; 2017.

49
Control of Ventilation

SUSAN L. EWART

KEY POINTS

1. Ventilation is regulated to meet the metabolic demands for delivery of oxygen and removal of carbon dioxide.

Central Control of Ventilation

1. Respiratory rhythmicity originates in the brainstem.

Pulmonary and Somatic Respiratory Muscle Receptors

1. Stretch receptors, irritant receptors, and J receptors modify the rhythm of breathing.
2. Somatic respiratory muscle mechanoreceptors monitor the effort exerted by respiratory muscles.

Chemoreceptors

1. Hypoxemia, hypercapnia, and acidosis stimulate ventilation through chemoreceptors.
2. Peripheral chemoreceptors monitor partial pressures of oxygen (P_{O_2}) and carbon dioxide (P_{CO_2}), and hydrogen ion concentration.
3. Central chemoreceptors monitor P_{CO_2} and hydrogen ion concentration.
4. Under some extreme conditions, the chemoreceptor response seems paradoxical.
5. Integrated breathing involves the central pattern generator and inputs from the various respiratory receptors.
6. Ascent to high altitude is accompanied by a decrease in inspired oxygen partial pressure (P_{iO_2}) and consequently by hypoxemia, which leads to an increase in ventilation.
7. During exercise, ventilation must increase because the tissues demand more oxygen and produce more carbon dioxide.

Ventilation Is Regulated to Meet the Metabolic Demands for Delivery of Oxygen and Removal of Carbon Dioxide

Animals constantly vary their level of activity, breathing air of varying composition and purity. To allow the respiratory system to respond to these different situations, control mechanisms monitor (1) the chemical composition of the blood, (2) the effort exerted by the respiratory muscles on the lungs, and (3) the presence of foreign materials in the respiratory tract. This information is integrated with the other nonrespiratory activities, such as thermoregulation, vocalization, eructation ("belching"), and parturition to produce patterns of breathing that maintain gas exchange.

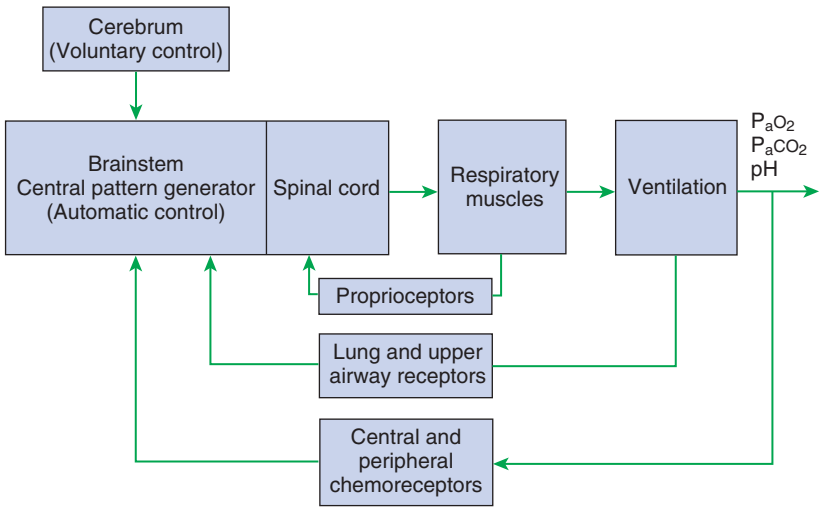
Fig. 49.1 provides a diagram of feedback control for the respiratory system. The neurons of the central pattern generator produce signals that regulate the activity of the respiratory muscles, which by contracting give rise to alveolar ventilation. Changes in alveolar ventilation affect blood gas partial pressures and pH, which are monitored by the peripheral and central chemoreceptors. Other receptors located in the lung and upper airways sense mechanical stretching of the airways during inspiration, airway irritants, and inflammatory mediators or chemicals in the circulation. Finally, mechanoreceptors in somatic respiratory muscles monitor the effort of breathing. These various receptors send signals back to the central pattern generator so that necessary adjustments can be made to ventilation (Fig. 49.2).

Central Control of Ventilation

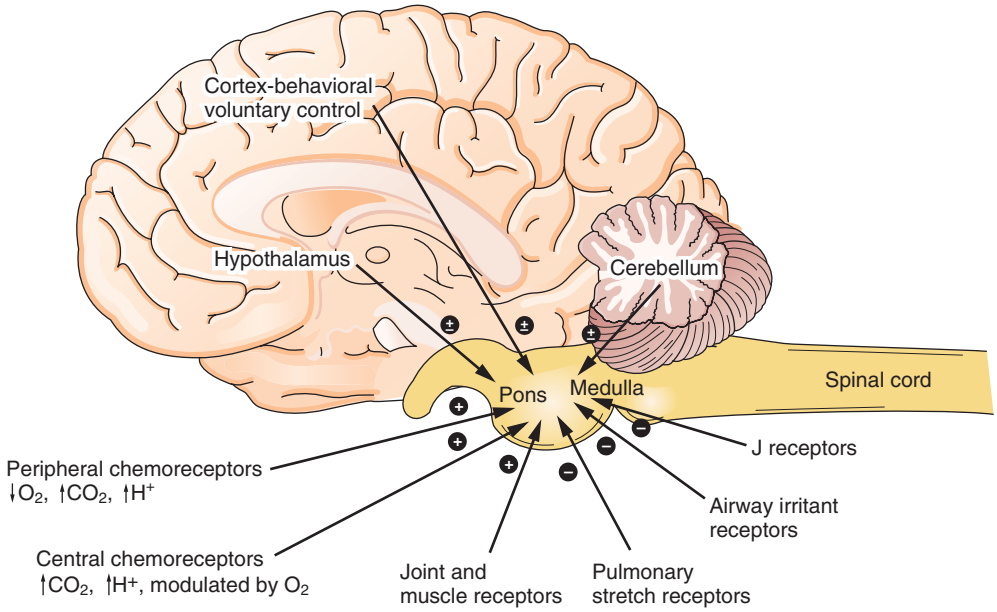
Respiratory Rhythmicity Originates in the Brainstem

The seemingly simple rhythmic in and out nature of breathing involves complex neural networks that include the respiratory *central pattern generator* (CPG) located in the pons and medulla of the brainstem. Although understanding of the CPG is incomplete, there is agreement that respiratory neurons include those in the pontine respiratory group (lateral parabrachial and Kolliker-Fuse areas) and several medullary areas, especially the retrotrapezoid nucleus (RTN), the Bötzinger and pre-Bötzinger complexes, and the rostral ventral and caudal ventral respiratory groups (rVRG and cVRG, respectively) (Fig. 49.3). These nuclei are arranged in a column extending from the lateral pons to the ventrolateral medulla and are thus referred to as the lateral respiratory column (LRC). A dorsal respiratory group of neurons located in the nucleus tractus solitarius of the medulla relays information to the CPG from chemoreceptors, stretch receptors, and irritant receptors.

It appears that the rhythmicity of normal resting breathing (*eupnea*) originates from the inhibitory interactions between several excitatory neuronal microcircuits, producing the rhythmic respiratory phases of inspiration, postinspiration (passive expiration), and expiration (late stage or active). During inspiration, there is an increase in activity in the inspiratory neurons of the pre-Bötzinger group and rVRG associated with phrenic nerve and intercostal nerve activation, which signals diaphragm and external intercostal muscle contraction, respectively. This increased activity can be amplified by an increase in the chemical respiratory drive, such as hypercapnia or hypoxemia. Termination of inspiration can be a



• **Fig. 49.1** Feedback control diagram for the regulation of ventilation. The central pattern generator in the brainstem drives the respiratory muscles that bring about ventilation and can be overridden by voluntary control from the cerebrum. Changes in ventilation can cause changes in blood gas partial pressures (P_aO_2, P_aCO_2) and pH that are monitored by central and peripheral chemoreceptors. Receptors in the lung detect the stretch in the lung tissues and the presence of materials in the lungs and airways. Proprioceptors in the respiratory muscles monitor the amount of effort being applied by the muscles. *P_aO_2*, Arterial O_2 partial pressure; *P_aCO_2*, arterial carbon dioxide partial pressure.

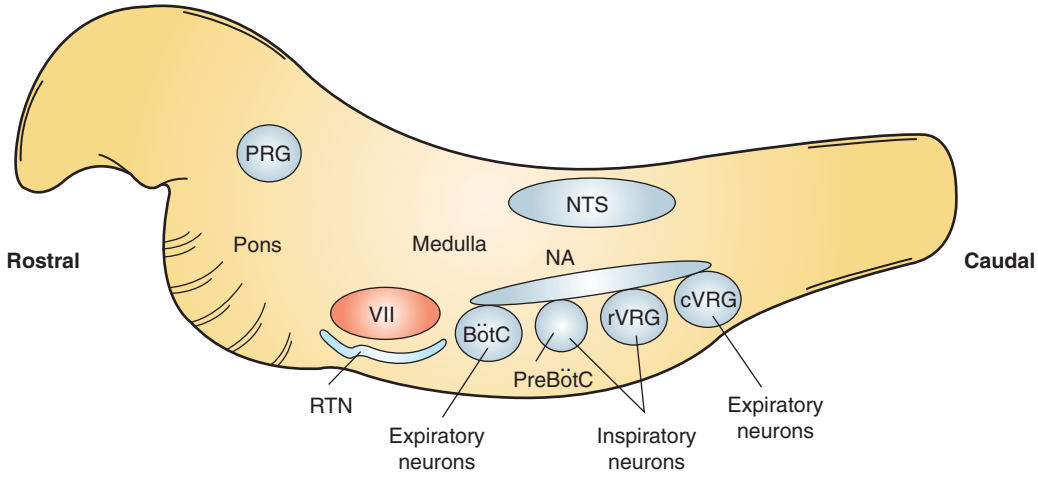


• **Fig. 49.2** Overview of ventilatory control mechanisms. Respiratory rhythmicity originating in the medulla and pons is modulated by multiple inputs. Symbols: *+*, stimulatory; *–*, inhibitory; *±*, stimulatory or inhibitory. Motor output to diaphragm, intercostal, and upper airway muscles not shown. (After Carroll JL, Agarwal A. Development of ventilatory control in infants. *Paediatr Respir Rev.* 2010;11[4]:199–207.)

result of vagal inputs from pulmonary stretch receptors or from a central pontine off switch. There is a complex interaction between tidal volume and the duration of inspiration related to these pulmonary stretch receptor signals relayed to the brain. This interaction leads to a larger tidal volume and more rapid respiratory frequency when the chemical drive to breathe is increased.

When inspiratory neurons are inhibited, inspiration is terminated and exhalation occurs passively as a result of the elastic recoil of the lung. Activity in some inspiratory neurons (pre-Bötzinger complex) early in exhalation leads to inspiratory muscle activity, which provides a "brake" on exhalation and regulates the rate of expiratory airflow. Later in exhalation, the braking is removed. When respiratory drive is low, this second phase of exhalation is initiated later than when drive is increased. If exhalation is active, as occurs during exercise, expiratory neurons (Bötzinger complex and cVRG) are activated in the latter part of expiration leading to contraction of expiratory muscles (abdominal and internal intercostal muscles).

Opioids are a class of analgesics that may induce respiratory depression, with potentially fatal consequences. Opioid receptors



• **Fig. 49.3** Diagram of the brainstem showing structures in the medulla and pons involved in the control of breathing. *BotC,* Bötzinger complex; *preBotC,* pre-Bötzinger complex; *cVRG,* caudal ventral respiratory group; *rVRG,* rostral ventral respiratory group; *NA,* nucleus ambiguus; *NTS,* nucleus tractus solitarius; *RTN,* retrotrapezoid nucleus/parafacial respiratory group; *PRG,* pontine respiratory group. VII, facial nucleus is shown for anatomical reference.

are abundant in the brainstem, carotid bodies, and on mechanoreceptors throughout the respiratory system. The primary mechanism by which opioids depress ventilation appears to be through suppression of rhythm generation.

The rhythmicity of breathing is frequently overridden by demands from higher brain regions for activities such as vocalization, swallowing, defecation, and parturition, which require the active participation of the respiratory system.

Pulmonary and Somatic Respiratory Muscle Receptors

Stretch Receptors, Irritant Receptors, and J Receptors Modify the Rhythm of Breathing

The vagus nerve includes both myelinated and nonmyelinated afferent axons conveying *sensory* information from the slowly adapting pulmonary stretch receptors, irritant receptors, and J receptors to the respiratory neurons of the pons and medulla.

Slowly adapting pulmonary stretch receptors are mechanoreceptors that have myelinated axons and nerve endings associated with smooth muscle in the trachea and main bronchi, and to a lesser degree in the smaller intrapulmonary airways. They are stimulated by deformation of the smooth muscle surrounding airway walls, as when intrathoracic airways are stretched during lung inflation. Firing rates from these receptors increase progressively as the lung inflates, resulting in negative feedback on inspiration (*Hering-Breuer reflex*). In this manner they modulate respiratory frequency and prevent overinflation. Termination of input from these receptors by vagotomy leads to a slowing of respiration and an increase in tidal volume. Slowly adapting pulmonary stretch receptors may be responsible, in part, for the important task of minimizing the work of breathing; this is done by adjusting the rate and depth of respiration so that adequate alveolar ventilation is achieved with minimal effort from the respiratory muscles. With each breath, the respiratory muscles have to work to overcome the inward recoil of the lungs during inspiration. During active expiration, as occurs in exercise or in the presence of some airway diseases, the respiratory muscles also have to overcome the outward recoil of the thoracic wall (see Chapter 45 and Fig. 45.13). Feedback from the slowly adapting pulmonary stretch receptors may also affect the tone of airway smooth muscle and thus airway resistance.

Irritant receptors, also known as *rapidly adapting pulmonary stretch receptors,* are myelinated nerve endings branching among epithelial cells from the nasopharynx to the bronchi. They are responsive to noxious stimuli including irritant gases, dusts, cold air, and some inflammatory mediators. They are also stimulated by mechanical deformation of the airways, such as that which results from mucus accumulations. Stimulation of these pulmonary irritant receptors leads to cough, bronchoconstriction, mucus secretion, and rapid, shallow breathing (*tachypnea*), all of which are protective responses to clear irritant materials from the respiratory system. These receptors may also initiate the sighs thought to redistribute pulmonary surfactant over the alveolar surface. Stimulation of irritant receptors located in the nasal cavity causes sniffing and sneezing, whereas stimulation of laryngeal and pharyngeal receptors may cause cough, apnea, or bronchoconstriction. Temperature receptors in the pharynx cooled by airflow alert the animal if there is insufficient airflow so that appropriate adjustments can be made by the inspiratory muscles to increase flow.

Pulmonary C fibers are unmyelinated nerve fibers that originate peripherally in *juxtacapillary* (or *J*) *receptors* located in the pulmonary interstitium close to both alveoli and pulmonary capillaries, as well as in the walls of conducting airways. They monitor blood composition and respond to chemicals in the circulation. They are also responsive to mechanical stimulation as they detect the degree of distention of the interstitium. As with pulmonary irritant receptors, stimulation of J receptors results in rapid, shallow breathing, mucus production, and bronchoconstriction. These protective responses accompany allergic, infectious, or vascular diseases.

Somatic Respiratory Muscle Mechanoreceptors Monitor the Effort Exerted by Respiratory Muscles

Muscle spindles and *tendon organs* are proprioceptive mechanoreceptors activated by changes in somatic muscle length and tension,

respectively (see Chapter 8). These receptors sense the increased strength of contractions needed for ventilation when respiratory muscles are unable to achieve their proper length or tension due to the presence of airway narrowing or obstruction, and modify the strength of respiratory muscle contraction accordingly. For the muscle spindles of respiratory muscles, the expected stretch during breathing movements can be regulated by the γ motor neuron system (see Chapters 8 and 10). The density of muscle spindles and tendon organs varies across the different respiratory muscles, and the effects of stimulating these receptors can vary with the anatomical location of the muscle group. The diaphragm has many tendon organs and few muscle spindles, whereas intercostal muscles are well supplied with both tendon organs and muscle spindles.

Chemoreceptors

Hypoxemia, Hypercapnia, and Acidosis Stimuli Ventilation Through Chemoreceptors

Chemoreceptors monitor hydrogen ion concentration [H^+] and partial pressures of carbon dioxide (P_{CO_2}) and oxygen (P_{O_2}) at several sites in the body. The levels of these substances detected by chemoreceptors is fed back to the respiratory control neurons in the pons and medulla and is used to modulate ventilation so as to maintain blood gas values in their normal range. Chemoreceptors provide some tonic drive to respiration during normal breathing. As blood composition departs from normal, changes in pH, P_{CO_2}, and P_{O_2} can produce major changes in ventilation.

There are two types of chemoreceptors: peripheral and central. The central chemoreceptors are responsible for more than two-thirds of the chemically-mediated control of ventilation with peripheral chemoreceptors responsible for the remainder. The response of the central chemoreceptors is limited to P_{CO_2} and pH; it is slower but more robust than the response to hypoxemia detected by peripheral chemoreceptors. The primary chemical regulator of ventilation is H^+, which is derived from arterial CO_2 and detected by central chemoreceptors.

Peripheral Chemoreceptors Monitor Partial Pressures of Oxygen (Po_2) and Carbon Dioxide (Pco_2), and Hydrogen Ion Concentration

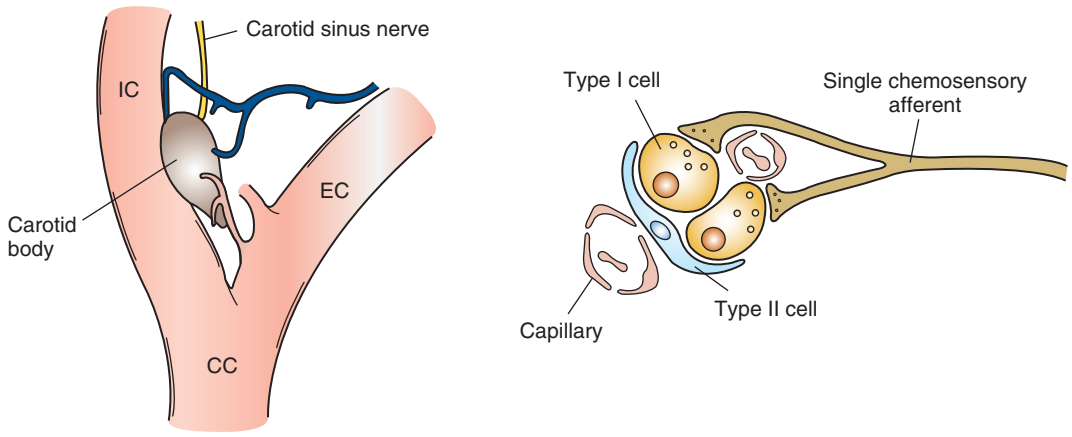
Peripheral chemoreceptors include the carotid bodies and aortic bodies. These structures are unique from the carotid sinus and aortic sinus baroreceptors present in these same locations. Carotid and aortic bodies respond to changes in arterial P_{O_2}, P_{CO_2}, and [H^+] by signaling through their afferent nerve fibers to the respiratory nuclei in the pons and medulla to modify ventilation so as to maintain blood gases in the normal range.

The *aortic bodies* are located around the aortic arch and are supplied by the vagus nerve. They appear to be most active in the fetus and are of less understood importance in the adult.

The *carotid bodies* (Fig. 49.4A) are located close to the bifurcation of the internal and external carotid arteries and are supplied by a branch of the glossopharyngeal nerve, called the carotid sinus nerve. They are small, nodular structures with extremely high blood flow per gram of tissue. This high blood flow-to-metabolism ratio allows the carotid bodies to obtain their O_2 from what is *dissolved* in the blood. Consequently, there is no arteriovenous difference in hemoglobin saturation and only a small difference between P_aO_2 and P_vO_2 as blood passes the carotid bodies.

Carotid bodies contain several cell types. Type I cells, or *glomus cells,* synapse with afferent axons that transmit information back to the brain (Fig. 49.4B). These glomus cells contain a variety of neurotransmitters, including catecholamines, especially dopamine. Glomus cells are most likely responsible for the chemosensitivity of the carotid bodies because they depolarize when P_{O_2} decreases. Alternatively, they may modify chemosensitivity of the afferent nerve terminals. Type II cells, or *sustentacular cells,* provide structural support to the axons and blood vessels that ramify within the carotid body.

When the carotid bodies are perfused with blood that has a low P_{O_2}, high P_{CO_2}, or low pH, firing rates in the carotid sinus nerve afferent fibers increase. The primary trigger for carotid body afferent signaling is decreased P_{O_2}. Modest increases in firing rate, and consequently ventilation, occur as P_{O_2} decreases from



• **Fig. 49.4** (A) Cartoon of the carotid body located at the bifurcation of the common carotid artery (*CC*) into its external (*EC*) and internal (*IC*) branches. (B) Cartoon of the basic cellular arrangement within the carotid body. Chemosensory afferent fibers are in synaptic contact with transmitter-filled type I (glomus) cells. Clusters of type I cells are encapsulated by type II (sustentacular) cells, and the organ receives a rich capillary blood supply. (From Peers C, Wyatt CN, Evans AM. Mechanisms for acute O_2 sensing in the carotid body. *Respir Physiol Neurobiol.* 2010;174[3]:292–298.)

nonphysiological levels of 500 mm Hg down to 70 mm Hg. Further P_{O_2} decreases cause a greater increase in ventilation, particularly at a P_{O_2} of less than 60 mm Hg, which is the level at which the slope of the oxyhemoglobin dissociation curve becomes steep and hemoglobin begins to desaturate (see Chapter 48). It is important to note that ventilation does not increase in response to either modest anemia or carbon monoxide (CO) poisoning, conditions that decrease the *O_2 content* of blood (see Chapter 48) but do not alter P_{O_2}. This is because P_{O_2}, rather than O_2 content, is the stimulus to the carotid bodies.

Along with responding to reduced P_{O_2}, the carotid bodies also increase their firing as P_{CO_2} increases and/or pH decreases; the presence of hypercapnia or acidosis results in an almost linear increase in ventilation. Peripheral chemoreceptors respond to both increased P_{CO_2} and decreased P_{O_2} such that the response is greater if both disturbances are present. Furthermore, hypercapnia serves to increase the sensitivity of peripheral chemoreceptors to reduced P_{O_2}. Finally, carotid body chemoreceptors have an important role in the respiratory response to metabolic acid-base disturbances; because central chemoreceptors have limited access to peripherally-derived H^+ (as described later), peripheral chemoreceptor detection of metabolically-driven changes in $[H^+]$ is important for the respiratory compensation for metabolic acid-base disorders (see Chapter 52).

Fig. 49.5 shows possible mechanisms of chemosensitivity of the carotid bodies. In either mechanism, glomus cells are depolarized by tissue hypoxia in the carotid body. Depolarization involves potassium (K^+) channels and leads to an increase in intracellular calcium (Ca_i^{2+}). The latter causes release of neurotransmitters, primarily dopamine and acetylcholine, which activate the afferent nerve terminals. Hypercapnia and changes in blood pH may also release neurotransmitters by decreasing the pH in the glomus cells.

Central Chemoreceptors Monitor P_{CO_2} and Hydrogen Ion Concentration

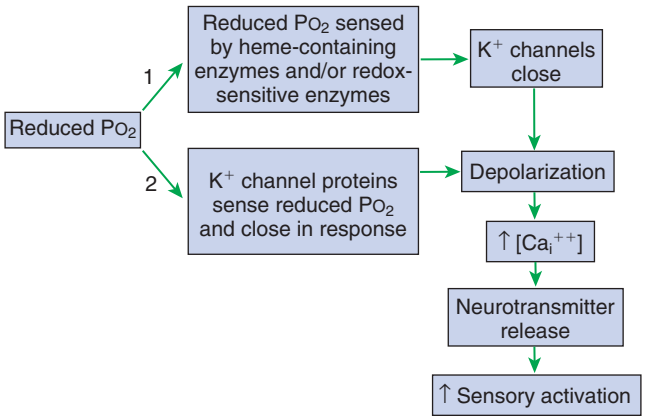
Central chemoreceptors detect changes in P_{CO_2} and $[H^+]$ within the cerebrospinal fluid and brain extracellular fluid, but they are *insensitive to P_{O_2}.* Carbon dioxide in particular is a potent stimulus to the central chemoreceptors. Increases in brain P_{CO_2} of approximately 5 mm Hg result in a 50% increase in ventilation; even greater increases in P_{CO_2} further stimulate ventilation and also cause a sensation of respiratory distress. Although hypercapnia is the primary driver for central chemoreceptors, its action is through H^+ as follows: CO_2 has ready access to central chemoreceptors because, unlike ions such as H^+, it is highly soluble and thus readily diffuses across the *blood-brain barrier.* Once it has crossed the blood-brain barrier, the CO_2 exerts its main effects on central chemoreceptors via conversion to H^+ in the reactions of the bicarbonate buffer system:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

As $[H^+]$ increases, the decrease in brain pH activates central chemoreceptors to stimulate ventilation. Conversely, decreased P_{CO_2} results in less CO_2 crossing the blood-brain barrier, with resultant lower $[H^+]$ and higher pH, which act on central chemoreceptors to reduce ventilation.

Metabolic acidosis can also stimulate ventilation by signaling through the central chemoreceptors, but this process is slower because the increased $[H^+]$ generated by retained fixed acids or loss of buffer base does not readily cross the blood-brain barrier, thus the ventilatory response to altered $[H^+]$ is primarily from peripheral chemoreceptors.

The precise location of the central chemoreceptors that affect breathing is unclear. Increases in ventilation can be provoked by local acid probing of several parts of the brain, and this stimulus exerts a particularly strong effect on the ventral surface of the medulla, especially in the region of the RTN. This is a small group of medullary glutaminergic neurons located under the facial motor nucleus and just rostral to the structures involved in the CPG. The RTN may be most critical for detecting the small CO_2-induced changes in brain interstitial fluid pH that regulate ventilation minute-by-minute (Fig. 49.6). This hypothesis is strongly supported by evidence from children and mice with a mutation in the transcription factor PhoxB2. The children develop central hypoventilation syndrome, and the mice, in which there is loss of specific RTN neurons, lose respiratory sensitivity to CO_2, have disordered breathing, and die early. The cells of the RTN are highly sensitive to acid and also vigorously activated by inputs from the carotid body. Under conditions of low respiratory drive, for example during sleep, the RTN may be the primary chemosensitive region, but it is also an integrating region for inputs from other chemosensitive tissues in the brain and especially the carotid body, and from higher brain centers, particularly the hypothalamus. During sleep, the RTN provides the major chemosensory drive but during wakefulness its function is more as a relay station for information from higher centers and other chemoreceptors.

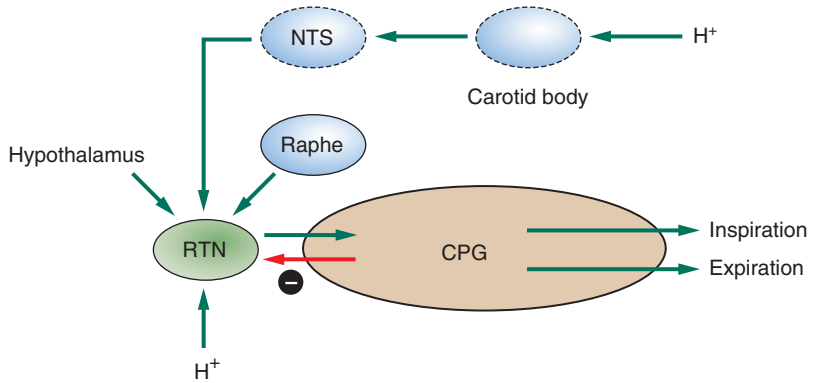


• **Fig. 49.5** Hypothesized mechanisms of chemosensitivity in the carotid body. *1,* Reduced P_{O_2} is sensed by enzymes in the glomus cell, leading to closure of potassium (K^+) channels, depolarization, and neurotransmitter release. *2,* Oxygen partial pressure is sensed directly by glomus cell potassium channels, causing their closure. (Modified from Prabhaker NR. Oxygen sensing by the carotid body chemoreceptors. *J Appl Physiol.* 2000;88[6]:2287–2295.)

Under Some Extreme Conditions, the Chemoreceptor Response Seems Paradoxical

The presence of severe respiratory disease and/or environmental conditions characterized by reduced O_2 or increased CO_2 can result in physiological responses that lead to further impairment of blood and tissue oxygenation. These paradoxical responses can occur by the following mechanisms.

Severe lung disease can result in a blunted ventilatory drive. In the presence of chronic hypercapnia that results from ongoing respiratory disorders, there is time for HCO_3^- ions to equilibrate across the blood-brain barrier and neutralize the acidic pH in the tissues surrounding the central chemoreceptors, thus removing



• **Fig. 49.6** The possible role of chemosensitive neurons of the retrotrapezoid nucleus (*RTN*) in the regulation of breathing. RTN neurons regulate breathing according to the surrounding pH but also based on the strength of inputs received from the carotid bodies, other chemoreceptors in the hypothalamus, medullary raphe, and nucleus tractus solitarius (*NTS*). RTN neurons provide a primary drive to the central pattern generator (*CPG*) in the medulla, which can also exert negative feedback on the RTN. *Green,* excitation; *red,* inhibition. (Modified from Guyenet PG, Stornetta RL, Bayliss DA. Central respiratory chemoreception. *J Comp Neurol.* 2010;518[19]:3883–3906.)

the central chemoreceptor stimulus for ventilation. In this case, the primary drive for ventilation is now carried by the peripheral chemoreceptors as they detect low P_aO_2. If a patient in this condition is then given supplemental O_2 such that the P_aO_2 increases to above approximately 60 mm Hg, the remaining peripheral chemoreceptor drive is lost and ventilation is compromised.

Another potentially disadvantageous response relates to vascular activity that may be linked to central chemoreceptors. Because the elevated P_aCO_2 that results from hypoventilation is often accompanied by hypoxemia, when P_aCO_2 rises, the blood supply to the brain vasodilates to ensure a sufficient supply of O_2. By the same token, in the presence of hyperventilation, when P_aCO_2 falls, the vascular supply to the brain vasoconstricts resulting in reduced blood flow to the brain and light headedness.

Finally, respiratory neurons are depressed by extreme conditions of very low P_aO_2 and very high P_aCO_2. These seemingly illogical responses under extreme conditions can occur in closed environments like closed circuit anesthesia systems (as well as submarines and space capsules).

Integrated Breathing Involves the Central Pattern Generator and Inputs From the Various Respiratory Receptors

The minute-by-minute pattern of breathing that veterinarians recognize as either normal or abnormal involves the integrated actions of the CPG, RTN, chemoreceptors, vagal afferents from pulmonary receptors, respiratory skeletal muscle mechanoreceptors, and higher brain centers (Figs. 49.2, 49.3, and 49.6). The complexity of overall integrated control is illustrated by the following examples. (1) Normally, the peripheral chemoreceptors are responsible for 40% to 50% of tonic drive to ventilate. Because of this, administration of O_2 suppresses ventilation in healthy animals. (2) The changes in brain pH detected by central chemoreceptors provide the fine control on breathing, but this is greatly affected by information from the peripheral chemoreceptors, by sleeping and waking, and by body temperature. (3) Metabolic acidosis, which is a consequence of increased production of H^+ without a change in P_aCO_2, increases ventilation by activation of both peripheral and central chemoreceptors. (4) P_aCO_2 remains constant during aerobic exercise, in part because of an anticipatory increase in ventilation that begins at the start of exercise before CO_2 production has increased. (5) The sensation of difficult breathing (*dyspnea*) and the altered respiratory frequency that accompany lung inflammation are consequences of increased input from carotid bodies, pulmonary stretch receptors, C fibers, and respiratory muscle spindles and tendon organs.

Ascent to High Altitude Is Accompanied by a Decrease in Inspired Oxygen Partial Pressure (P_iO_2) and Consequently by Hypoxemia, Which Leads to an Increase in Ventilation

The fraction of O_2 in air is constant worldwide ($F_iO_2 = 0.21$), but barometric pressure decreases as altitude increases. Consequently, the P_iO_2 in air decreases with increasing altitude (see Chapter 47). The ventilatory response to this hypoxemia of altitude varies, depending on whether it lasts for less than an hour or for several days. The acute hypoxemia experienced on first ascending to high altitude causes hyperventilation mediated through activation of the peripheral chemoreceptors. This hyperventilation serves to increase P_aO_2 toward normal, as intended, but it also decreases P_aCO_2 and increases pH, creating a respiratory alkalosis (see Chapter 52). This inadvertent respiratory alkalosis dampens the hyperventilatory response to hypoxemia. Over time (hours to days), the kidneys begin to compensate for the respiratory alkalosis by eliminating HCO_3^- and thus reducing the pH. Similarly, the pH of the cerebrospinal fluid (CSF) is modified by outward movement of bicarbonate in response to the alkalosis. These compensatory mechanisms return the pH toward normal, which subsequently removes the ventilatory-dampening effects induced by the initial alkalosis. At this point, with the respiratory alkalosis partially compensated in the face of ongoing environmentally-induced hypoxemia (when the animal remains at high altitude and the hyperventilation has not been sufficient to restore P_aO_2 to normal), the ventilation again increases. Ventilation may remain somewhat elevated for hours to days after the hypoxic stimulus is removed because the hypoxic stimulus increases the chemosensitivity of the carotid body glomus cells.

Longer-term adjustment to hypoxemia and subsequent tissue hypoxia involves (1) production of more erythrocytes (increased hematocrit) under the influence of erythropoietin, (2) decreased affinity of hemoglobin for O_2 (facilitating unloading at the tissues) because of increased concentrations of 2,3-bisphosphoglycerate (2,3-BPG), (3) an increase in the pulmonary surface area available for diffusion, and (4) increased capillary density in muscle. These adjustments are sufficient to restore maximal O_2 consumption to normal at moderate (<2000 meters) but not extreme altitudes.

During Exercise, Ventilation Must Increase Because the Tissues Demand More Oxygen and Produce More Carbon Dioxide

The increase in ventilation after the onset of exercise is initially rapid, then progresses more slowly, and, provided the workload remains constant, reaches a steady state after about 4 minutes. Although the *ventilatory response to exercise* has been well described, the reasons for the increase in ventilation are still not well understood.

The primary chemical stimuli for ventilation—P_aO_2, P_aCO_2, and pH—do not change in most animals during moderate aerobic exercise. This observation shows that the increase in ventilation is well matched to the needs of the tissues and that other factors besides chemical drive increase ventilation during exercise. The leading hypothesis for exercise-induced ventilatory drive is based on the function of *central command neurons* coupled through the RTN to the CPG neurons. These command neurons are thought to control the responses of the respiratory and cardiovascular systems necessary to maintain O_2 delivery and CO_2 removal during exercise. Other factors that may contribute to ventilatory drive during exercise include (1) reflexes originating from motion of the exercising limbs, (2) increase in cardiac output, (3) the need to maintain thermoregulation, and (4) psychogenic stimuli related to anticipation of the onset of exercise.

When the *anaerobic threshold* is exceeded, the production of lactic acid decreases blood pH. The latter stimulates an increase in ventilation, which leads to a decrease in P_aCO_2. In the horse, the increase in ventilation that can occur during exercise may be limited by respiratory rate being linked, one breath per stride, with stride frequency. During strenuous anaerobic exercise, the horse's arterial pH decreases progressively, although ventilation remains constant. When exercise ceases, ventilation increases further, presumably because the restrictions imposed by locomotion are removed.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Hypoxemia With Hyperventilation in a Samoyed Puppy

History

An 8-month-old, intact female Samoyed is brought to you because it is reluctant to exercise. Ever since the owner obtained the puppy, she has noticed that its behavior is not "puppy-like;" it tires easily and prefers to sleep rather than play.

Clinical Examination

The puppy is small for its age and breed. When standing quietly in the examination room, the dog breathes normally, but when you call and it runs toward you, its respiratory rate increases and the dog begins to pant. At this point, you notice that the dog's tongue and gums have a bluish tinge.

Before examining the dog further, you are already suspicious of a congenital cardiac anomaly. You suspect this because of the dog's age and history, and because it became cyanotic with minimal exercise.

Palpation of the dog shows that, even though it is small, it is not thin. The major abnormalities are the cyanosis of the mucous membranes, an elevated heart rate, and loud, abnormal cardiac sounds. A murmur is audible over the tricuspid valve area during systole. The murmur (grade 4/5) is turbulent enough to produce a palpable vibration on the chest wall. You explain your suspicions of a cardiac defect to the owner, and together you decide to request some studies to determine the nature of the defect.

An echocardiogram, which is performed to determine the type of cardiac anomaly, demonstrates a large opening between the left and right atria, caused by a patent foramen ovale. Before angiographic studies, an arterial blood gas sample is taken to determine the suitability of the dog for anesthesia; P_aO_2 is 61 mm Hg (normal, 85–100 mm Hg) and P_aCO_2 is 23 mm Hg (normal, 35–45 mm Hg).

Angiography is performed successfully. A catheter is floated into the right atrium of the dog, and dye is injected at this site. Some of the dye passes into the right ventricle, but a large portion passes from the right atrium into the left atrium and out into the systemic circulation.

Comment

The blood gas results are fairly typical for an animal with a major O_2 exchange problem; this dog has a right-to-left vascular shunt through a cardiac defect. A large amount of the mixed venous blood returning to the heart is bypassing the lungs, resulting in the low P_aO_2. The P_aO_2 is low enough to cause a major increase in ventilation by stimulating the peripheral chemoreceptors. This increase in ventilation causes excessive elimination of CO_2, resulting in the reduced P_aCO_2. Ventilation has increased but not enough to resolve the hypoxemia; this is because the low P_aCO_2 acting on the central chemoreceptor slows down the ventilatory increase.

The echocardiogram and angiographic studies are indicative of a patent foramen ovale. Normally this would not result in right-to-left shunting of blood because the pressure in the left atrium is usually higher than that in the right atrium. However, this dog probably also has abnormalities of the tricuspid valve that cause a partial obstruction. This is sufficient to increase the pressure in the right atrium and cause blood to flow from right to left through the foramen ovale.

Treatment

The cardiac defect will need to be corrected surgically if the dog is to have any chance of life. Echocardiography is generally performed before angiography because it can better determine the extent of the abnormalities and is safer. An angiogram could then be performed, if necessary. With some limited cardiac anomalies, a transcatheter device can accomplish closure of the foramen ovale. However, if there are other cardiac abnormalities, the method of repair may be altered.

Hypoventilation in an Anesthetized Saint Bernard

History

A 2-year-old, neutered male Saint Bernard is brought to you for treatment of a fractured femur. You elect to place an intramedullary pin in the femur, and the dog will require anesthesia. The dog is anesthetized with a barbiturate, an endotracheal tube is placed, and the dog is administered O_2 containing 2% isoflurane. It is not ventilated but is allowed to breathe the anesthetic mixture spontaneously. The veterinary technician observing the dog notices that the gas reservoir bag on the anesthesia machine is not moving much when the dog breathes. Therefore she draws an arterial blood sample for measurement of blood gases. Measurement reveals a P_aO_2 of 480 mm Hg and P_aCO_2 of 90 mm Hg (normal, 35–45 mm Hg).

CLINICAL CORRELATIONS—cont'd

Comment

This is an example of alveolar hypoventilation. Carbon dioxide is produced by the tissues faster than it is being eliminated by the lungs, so the P_aCO_2 is elevated. The P_aO_2 is sufficient because the dog is on supplemental O_2; the measured P_aO_2 is acceptable for a dog breathing O_2. Hypoventilation is a common occurrence in anesthetized animals, particularly when anesthesia is induced with a barbiturate drug. Barbiturates depress the respiratory control centers and are prone to cause apnea. Hypoventilation also occurs because the elevated P_aO_2 provides no peripheral chemoreceptor drive. Furthermore, the ventilatory response to CO_2 is depressed by anesthesia, and therefore a larger increase in P_aCO_2 than usual is required to trigger an increase in ventilation.

Treatment

The dog needs more alveolar ventilation to decrease P_aCO_2 and prevent respiratory acidosis. The additional ventilation can be supplied by assisting ventilation via squeezing the rebreathing bag on the anesthetic machine or by instituting full mechanical ventilation. When the dog recovers from anesthesia, its own respiratory control mechanisms will regulate alveolar ventilation and return P_aCO_2 to normal.

Practice Questions

1. The rhythmicity of breathing is thought to originate in:
- **a.** The carotid body
- **b.** The central pattern generator
- **c.** The central chemoreceptor
- **d.** Rapidly adapting pulmonary stretch receptors
- **e.** None of the above

2. Which of the following receptors have afferent nerve fibers in the glossopharyngeal nerve?
- **a.** Carotid bodies
- **b.** Slowly adapting pulmonary stretch receptors
- **c.** Aortic bodies
- **d.** Intercostal stretch receptors
- **e.** Rapidly adapting pulmonary stretch receptors

3. Which of the following statements correctly describes the carotid bodies?
- **a.** Carotid bodies can increase ventilation in response to low P_{O_2}, but not in response to an increase in P_{CO_2}.
- **b.** Carotid bodies have a low blood flow/metabolism ratio.
- **c.** Chemoreception is thought to occur in the sustentacular cells.
- **d.** Carotid bodies are located near the bifurcation of the internal and external carotid arteries.
- **e.** All the above.

4. The retrotrapezoidal neurons:
- **a.** Are highly sensitive to increases in hydrogen ion concentration.
- **b.** Receive inputs from the carotid body via the nucleus tractus solitarius.
- **c.** Provide input to the central pattern generator.
- **d.** Receive inputs from higher centers such as the hypothalamus.
- **e.** Are described by all of the above.

5. Which of the following is correct concerning the role of P_aCO_2 in breathing?
- **a.** P_aCO_2 exerts its effects on ventilation by changing the pH of brain interstitial fluid.
- **b.** P_aCO_2 has no effect on the carotid body.
- **c.** P_aCO_2 remains constant when hypoxemia increases ventilation during the ascent to high altitude.
- **d.** P_aCO_2 is much less important than P_aO_2 in regulation of breathing.
- **e.** None of the above.

6. Which of the following receptors are thought to initiate a cough in response to mechanical deformation of the airway?
- **a.** Juxtacapillary receptors
- **b.** Rapidly adapting stretch receptors
- **c.** Slowly adapting stretch receptors
- **d.** Intercostal tendon organs
- **e.** None of the above

Bibliography

Burki NK, Lee LY. Mechanisms of dyspnea. *Chest*. 2010;138(5):1196–1201.

Carroll JL, Agarwal A. Development of ventilatory control in infants. *Paediatr Respir Rev*. 2010;11(4):199–207.

Dutschmann M, Dick TW. Pontine mechanisms of respiratory control. *Compr Physiol*. 2012;2:2443–2469.

Guyenet PG, Bayliss DA. Neural control of breathing and CO_2 homeostasis. *Neuron*. 2015;87(5):946–961.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Richerson GB, Boron WF. Control of ventilation. In: Boron WF, Boulpaep EL, eds. *Medical Physiology*. 3th ed. Philadelphia: Saunders; 2017.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

Widdicombe J. Functional morphology and physiology of pulmonary rapidly adapting receptors (RARs). *Anat Rec A Discov Mol Cell Evol Biol*. 2003;270A:2–10.

Widdicombe JG, Sant'Ambrogio G. *Mechanoreceptors in Respiratory Systems. Advance in Comparative and Environmental Physiology*. Vol 10. Springer-Verlag Berlin Heidelberg. 1992.

50

Nonrespiratory Functions of the Lungs

SUSAN L. EWART

KEY POINTS

Defense Mechanisms of the Respiratory System

1. The lung is protected by innate and adaptive immune defense mechanisms.
2. Particle deposition in the respiratory tract depends on particle size and occurs by impaction, sedimentation, and diffusion.
3. Mucociliary clearance removes particles that are deposited on the walls of the upper airway and tracheobronchial tree.
4. Dendritic cells and alveolar macrophages scavenge particles deposited on the airway surface and bridge innate and adaptive immune responses.
5. Cytokines and chemokines communicate adaptive defense responses.

Pulmonary Fluid Exchange

1. The lung continuously produces lymph as a result of net fluid movement from the pulmonary microvasculature into the pulmonary interstitium.
2. Pleural fluid originates by filtration from capillaries in the parietal pleura and is reabsorbed through stomata that connect to lymphatics.

Metabolic-Related Functions of the Lung

1. The lung metabolizes many biochemical substances, removes toxins from the blood, and its actions affect thermoregulation.

Defense Mechanisms of the Respiratory System

The Lung Is Protected by Innate and Adaptive Immune Defense Mechanisms

The respiratory system is continually exposed to the ambient air and all the organic dust, pathogens, and toxins it may contain, including particles of plant and animal origin; infectious agents such as bacteria, viruses, and fungi; allergens such as spores and pollen; and other agents such as endotoxin. Furthermore, livestock that are intensively housed or being transported are commonly subjected to additional pollutant gases such as ammonia, diesel fumes, oxides of nitrogen, and ozone. Along with the injurious substances inhaled into the lung are many innocuous particles and commensal organisms. The respiratory system has a variety of defense mechanisms to protect it against potentially injurious substances, as well as distinguish pathogens from harmless substances.

The immune response in the lungs is complicated and involves many cells types, signaling molecules, and compounds with enzymatic, oxidant, and other antimicrobial actions. It is generally categorized into *innate* and *adaptive* (or *acquired*) *immune responses* with well-coordinated communication between the two systems.

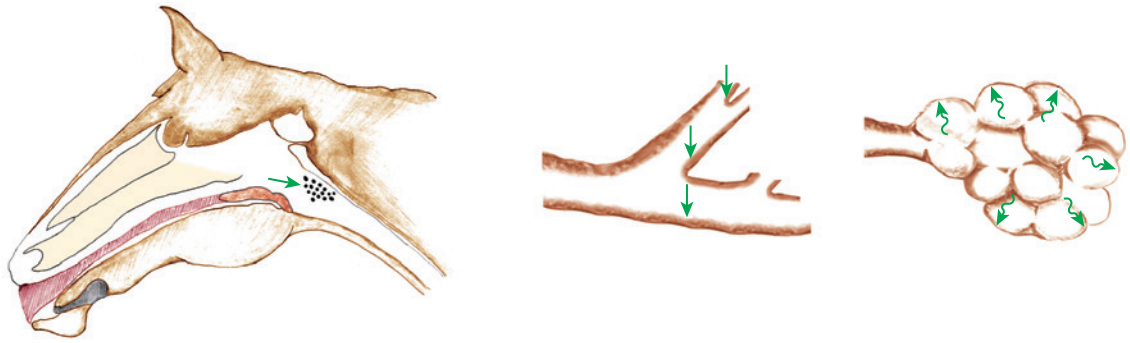
Nonspecific defenses are provided by the *innate immune system* (see Chapter 54), which immediately protects against many inhaled substances. Innate immunity includes *mucociliary clearance, cough, sneeze,* and *alveolar macrophages* as the resident phagocytic cells in the alveoli. In addition, *dendritic cells* continuously sample the airway environment and present any pathogens they find to cells that regulate further immune responses. Finally, *toll-like receptors* on the surface of many types of innate immune cells recognize molecules common to many bacteria and fungi. When activated, these receptors immediately initiate mechanisms that lead to expression of antimicrobial peptides and proinflammatory molecules.

Specific defenses involve the *adaptive immune system* (see Chapter 55) and are directed against specific pathogens following exposure. On initial exposure, adaptive immune responses need several days to become activated; they subsequently develop an immune memory that more rapidly protects against future attacks by the same organism. T and B lymphocytes are the primary effector cells of the adaptive immune response as they orchestrate the production and release of cytokines, chemokines, and antibodies, and the activation and recruitment of other cells.

When the immune response is overwhelmed by pathogens, animals can develop diseases such as pneumonia or pleuritis. Failure of the immune system to distinguish pathogens from innocuous materials results in allergy, whose symptoms are often respiratory in origin.

Particle Deposition in the Respiratory Tract Depends on Particle Size and Occurs by Impaction, Sedimentation, and Diffusion

Airborne material is inhaled as aerosols suspended in air or as toxic gases. The term *aerosol* refers to collections of *particles* or *liquid droplets* small enough to remain suspended in air for a period of time. As relates to the airways, particles can be described as inhalable or respirable. *Inhalable particulates* have a median diameter of 10 microns (micrometers, μm) or less (referred to as PM10). *Respirable particulates* have a diameter of 2.5 μm or less (PM2.5). Increases in the atmospheric concentrations of particulates are associated



• **Fig. 50.1** Mechanisms of particle deposition in the tracheobronchial tree. (A) Large particles are deposited by impaction in the bends in the larger airways. (B) Medium-sized particles are deposited in the smaller airways by sedimentation. (C) Small particles contact the walls of the alveoli by diffusion.

with adverse health effects, including worsening of respiratory and cardiovascular diseases. Exposure to PM2.5 poses greater risks than PM10 as the smaller particles penetrate farther into the lungs. Intensively-housed animals can be exposed to PM10 and PM2.5 concentrations as great as those known to cause respiratory disease in humans.

Particles and aerosols are removed from the air when they contact the moist epithelial surface of the upper airway or tracheobronchial tree (Fig. 50.1). The distance that particles and aerosols travel into the respiratory tract depends on their size. Larger particles, greater than 5 μm in diameter, contact the airway wall by *inertial impaction.* Inertial impaction occurs at the bends in the large airways because large particles traveling at high velocity have so much momentum that they fail to negotiate the turns. At sites of inertial impaction, there are accumulations of lymphoid tissue, such as tonsils and bronchus-associated lymphoid tissue (BALT), presumably to orchestrate an immune response to the material landing on the airway surface. As airflow rates diminish deeper in the lung, particles 1 to 5 μm in diameter settle onto the walls of the airways by *sedimentation.* The smallest particles reach the peripheral airways and alveoli, where they either contact the epithelial surface by *diffusion* through the thin layer of aqueous fluid lining the alveoli or are exhaled. Inhaled drugs must be delivered in a form with a particle size of 1 to 5 μm to be deposited onto the airway walls and remain in the lungs.

The deposition of particles and aerosols within the respiratory tract is also influenced by the pattern of breathing. Slow, deep breathing transports particles deep into the lung, whereas rapid, shallow breathing enhances inertial deposition in the larger airways. Bronchoconstriction enhances deposition of particles in more central airways, whereas bronchodilation favors more peripheral distribution.

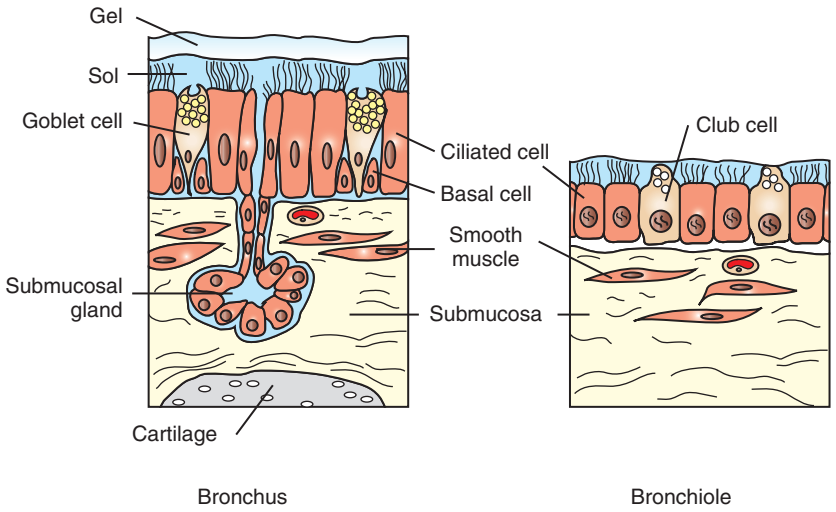
The deposition of *toxic gases* depends on their solubility and concentration. Highly soluble gases, such as sulfur dioxide (SO_2), in low concentrations are removed by the nose, but in higher concentrations they can penetrate deeper into the respiratory system. Less soluble gases can gain access to the alveoli. Toxic gases stimulate a variety of protective mechanisms, such as bronchoconstriction, mucus hypersecretion, coughing, and sneezing (see Chapter 49).

Mucociliary Clearance Removes Particles That Are Deposited on the Walls of the Upper Airway and Tracheobronchial Tree

Particles deposited on the epithelial surface of the upper airways and tracheobronchial tree are removed by *mucociliary clearance* as they are transported to the pharynx, where they are then swallowed or expectorated. The system of mucociliary clearance consists of *sol* and *gel* mucus layers overlying ciliated epithelial cells (Fig. 50.2). The low-viscosity sol layer, through which the beating cilia move, bathes the surface of the epithelial cells. On its forward stroke, the extended cilium catches the overlying viscous gel layer, in which inhaled particles are entrapped, and propels it up the tracheobronchial system or through the nasal cavity. The clearance rates and the beating frequency of cilia in the bronchi and trachea are faster than those in the bronchioles; this differential rate of mucus transport is required for the trachea and bronchi to be able to adequately move the mucus delivered to them from the vast surface area of the peripheral airways. In large mammals, gravity also plays an important role in speeding mucociliary clearance. If a horse is prevented from lowering its head, the rate of mucociliary clearance from the lungs is reduced. As a consequence, the number of bacteria in the trachea increases, and this can lead to pneumonia. Inability of horses to lower their heads in a horse trailer may explain why transportation over long distances is the greatest risk factor for development of pneumonia in horses.

Respiratory tract mucus originates from several sites (Fig. 50.2). In respiratory bronchioles, the nonciliated *club cells* (previously termed Clara cells) are a source of the fluid that lines the airways. In the larger airways, *goblet cells* produce mucous secretions. In the bronchi, *submucosal glands* produce both serous and mucous secretions. Secretion is under autonomic regulation. Throughout the respiratory tract, transepithelial movement of water and ions can change the composition of the mucus layer. Ion and fluid exchange is assisted by *microvilli* on the surface of epithelial cells, which function to increase the surface area available for exchange.

Changes in the amount, composition, and viscosity of mucus occur in response to many stimuli and can be the cause or the result of respiratory disease. Normal airway epithelia regulate the rates of sodium (Na^+) absorption and chloride (Cl^-) secretion to



• **Fig. 50.2** Diagram of the epithelium and submucosa of a bronchus and bronchiole. In the bronchus, the epithelium is pseudostratified columnar and includes goblet cells, ciliated cells, and basal cells, the latter of which do not reach the surface of the epithelium. A bronchial gland is shown in the submucosa with its duct passing through the smooth muscle. Cartilage underlies the submucosal layer. The cilia beat within a sol layer over which is a layer of gel-type mucus. In the bronchiole, the epithelium is cuboidal and is a mixture of ciliated cells and secretory club cells. Smooth muscle is shown in the submucosa. Bronchioles normally do not have submucosal glands or goblet cells, and there is no cartilage in their walls.

regulate the depth of the mucous layer for optimal ciliary function. A change in the depth or viscosity of the sol layer impairs ciliary function, and changes in the viscoelastic properties of the gel layer alter clearance rates. Primary ciliary dyskinesis is an inherited condition in which cilia function is impaired. This disease occurs in several dog breeds and results in chronic respiratory symptoms and pneumonia, along with clinical signs related to cilia dysfunction in other organs.

Coughing is part of the clearance mechanism of the respiratory tract and is initiated by stimulation of *subepithelial irritant receptors,* which are most numerous in the larger bronchi. Receptors can be stimulated by the mechanical deformation that results from foreign bodies or excessive amounts of material, such as mucus on the epithelial surface. The cough reflex becomes hyperresponsive when the air passages are inflamed and/or the respiratory tract epithelium is injured (e.g., by viral infections). Cough is effective in clearing mucoid secretions from the intrathoracic trachea and large bronchi, but it does not assist in removing mucus from the more peripheral bronchi and bronchioles. *Sneezing* is mediated by nociceptive C type neurons and functions to expel irritants, mucus, and the particles contained therein from the upper airways.

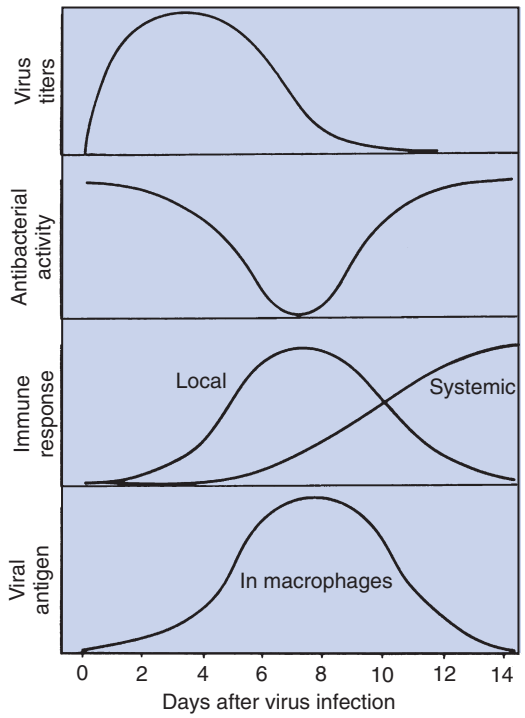
Dendritic Cells and Alveolar Macrophages Scavenge Particles Deposited on the Airway Surface and Bridge Innate and Adaptive Immune Responses

Lung residing *dendritic cells* are present primarily on the basolateral aspect of the airway epithelium. They serve as sentinel cells by sampling the alveolar and conducting airway surfaces for antigens and subsequently eliciting immune cell responses appropriate to the antigen encountered. Dendritic cells not only orchestrate the immune response to viral and bacterial pathogens, but they are also critically important for ensuring immune tolerance when presented with innocuous antigen so that inflammation is not induced unnecessarily; allergy is a result of immune tolerance failure.

Alveolar macrophages, which constitute the majority of cells in the alveolar lining fluids, are the principal resident phagocytes in the normal lung and an important part of the innate immune system. Surfactant proteins, complement, opsonins, and lysozymes in respiratory tract secretions assist macrophages in the killing and removal of pathogens, such as bacteria. When phagocytized, particles are destroyed or transported out of the lung by the macrophage. Some macrophages enter the mucociliary clearance system directly from the alveolus; others traverse the alveolar wall and enter the lymphoid tissues associated with the airways. In the lymphoid tissue, macrophages serve as *antigen presenting cells* (APCs), and in this role they are critical in orchestrating the lung's adaptive immune responses.

Alveolar macrophages have adapted to the high oxygen (O_2) levels of the alveolus, and their role as phagocytes is depressed by hypoxia. Macrophage function is also suppressed by endogenous glucocorticoids that are released from the adrenal glands at times of stress and by synthetic corticosteroids that are used to relieve inflammation (e.g., in arthritis or asthma). Stress-induced suppression of macrophage function contributes to respiratory disease in animals transported for long distances. In addition, excessive administration of synthetic corticosteroids can make animals more susceptible to viral and bacterial lung infections. Viral infections also suppress macrophage function; this occurs approximately 7 days after virus inoculation (Fig. 50.3) and contributes to the secondary bacterial infections that commonly follow viral respiratory disease.

Alveolar macrophages are a first line of defense. When large numbers of particles are inhaled, other phagocytes from the blood, particularly *polymorphonuclear neutrophils* (PMNs or neutrophils), assist the macrophage. Phagocytic cells release toxic *O_2 radicals* and *proteolytic enzymes* to break down invading bacteria, but they may also damage the lung tissue in the process. Endogenous protease inhibitors (e.g., α_1-antitrypsin) and antioxidants (e.g., glutathione



• **Fig. 50.3** Effects of viral infection on antibacterial activity of alveolar macrophages. Antibacterial activity is depressed 7 days after experimental viral infection. At this time, the viral antigen is located in the macrophages, which are damaged by the local immune response to the virus. (From Jakab GT. Viral-bacterial interactions in respiratory tract infections: A review of the mechanisms of virus-induced suppression of pulmonary antibacterial defenses. In: Loan RW, ed. *Bovine Respiratory Disease: A Symposium.* College Station, Tex: Texas A&M University Press; 1984.)

peroxidase, ascorbic acid) protect the lung from its own defense mechanisms.

Cytokines and Chemokines Communicate Adaptive Defense Responses

When infectious agents, an allergic response, or inhalation of toxic particles or gases injures the lung, an elaborate inflammatory process is elicited. *Cytokines* are chemical mediators produced and released by cells of the immune system that direct the actions of other cells through signaling pathways, providing a means for communication between the cells involved in the inflammatory process. *Chemokines* (*chemotactic cytokines*) are a subset of cytokines that promote chemotaxis, thereby attracting inflammatory cells to the site of injury. Cytokines and chemokines are produced and released by macrophages, lymphocytes, epithelial and endothelial cells, and various other cells involved in the inflammatory process. Along with establishing the initial inflammation associated with injury, they also are involved in downregulating the inflammatory response and orchestrating tissue remodeling to promote healing.

An example of the interplay of cytokines and chemokines in inflammation occurs when there is physical injury to the lung epithelium or the presence of bacteria in the lung; either of these stimuli cause the release of the cytokines' tumor necrosis factor (TNF) and interleukin (IL)-1 from macrophages. These cytokines act to draw neutrophils into the injured area of the lung. In addition, TNF and IL-1 initiate mechanisms that cause epithelial cells (e.g., alveolar type II) and endothelial cells to produce the chemokine IL-8, which prolongs the inflammatory response and is also a potent chemoattractant of inflammatory cells. Injured bronchial epithelial cells are also capable of producing IL-1, IL-6, granulocyte-macrophage colony-stimulating factor (GM-CSF), and IL-8, all of which have roles in the cascade of inflammation. Other cytokines including IL-4, IL-5, IL-9, IL-13, IL-17, and IL-22 (among others) are involved in allergic inflammation.

Pulmonary Fluid Exchange

The Lung Continuously Produces Lymph as a Result of the Net Fluid Movement From the Pulmonary Microvasculature Into the Pulmonary Interstitium

In the lung, as in other organs, there is continual movement of water and solutes from the capillary bed into the lung interstitium. The hydrostatic and osmotic forces governing fluid movement are the same as those in other organs, but because the lung vasculature operates at low pressure, the magnitude of the forces is different and the consequences are more severe. About 60% of fluid filtration takes place from alveolar capillaries, 15% from extraalveolar small arteries, and 20% from extraalveolar veins.

Fig. 50.4 represents a capillary in an alveolar septum. Fluid filtration normally occurs between the capillary and the interstitial tissue on the thick side of the alveolar septum where a layer of interstitium is interposed between the capillary endothelium and the epithelial basement membrane. Fluid filtering from pulmonary capillaries does not accumulate in, but moves through, this interstitium because the latter has a low compliance as a result of *proteoglycan* links between structures. On the thin side of the alveolar septum, interstitial tissue is absent because the capillary endothelium shares a basement membrane with the alveolar epithelium.

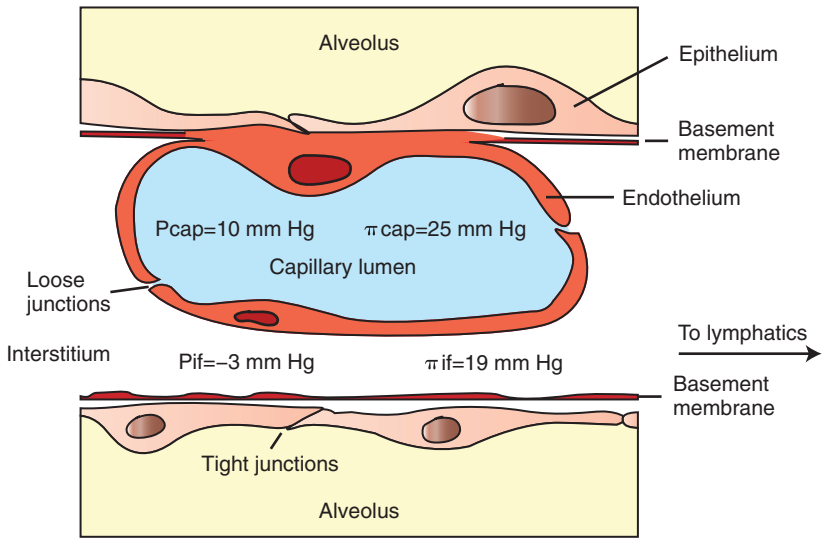
The cells of the pulmonary capillary endothelium are joined by loose junctions that provide a passage for movement of fluid, small molecules, and ions between the capillary lumen and the interstitial space. These junctions are loose enough that some small proteins, like albumin, can cross leading to pulmonary interstitial fluid albumin concentration approaching half that of blood. In pathological states such as inflammation, the endothelium becomes more *permeable;* this allows larger molecules to pass in between neighboring endothelial cells and thus move between the capillary lumen and the interstitial space. Increased permeability is indicated by a decreased reflection coefficient (σ) in the Starling equation (see below).

In contrast to the loose connections between pulmonary capillary endothelial cells, the alveolar epithelium contains tight junctions that largely restrict fluid and molecule passage under normal conditions. Thus the alveolar epithelium is less permeable than the pulmonary capillary endothelium and fluid does not leak from the interstitium into the alveolar air space unless the epithelium is damaged or there is considerable fluid accumulation in the interstitium.

Forces described in *Starling's equation* govern the movement of fluid across the pulmonary capillary endothelium as follows:

$$\text{Fluid flux} = K\{(Pcap - Pif) - \sigma\,(\Pi cap - \Pi if)\}$$

where K is the rate of flow (mL/h) per unit of pressure across the endothelium and is referred to as the *filtration coefficient; Pcap* is capillary hydrostatic pressure; *Pif* is interstitial fluid hydrostatic pressure; Πcap and Πif are capillary and interstitial colloid osmotic



• **Fig. 50.4** Diagrammatic representation of a pulmonary capillary in the alveolar septum. *Top,* on the "thin" side of the septum, capillary endothelium and alveolar epithelium share a basement membrane. *Bottom,* on the "thick" side of the septum, endothelium and epithelium are separated by a layer of interstitial tissue. Typical values for capillary and interstitial fluid hydrostatic pressures (Pcap and Pif) and oncotic pressures (Πcap and Πif) are shown.

(oncotic) pressures, respectively; and σ is the *reflection coefficient,* which indicates how effective the capillary endothelium is at preventing the movement of proteins and other solutes (also see Chapter 23). Values for the reflection coefficient (σ) range from 0, indicating no barrier (leaky) to passage of solutes, to 1.0, indicating no passage of solutes (tight barrier), with a typical range for pulmonary capillaries of 0.5 to 0.7 under normal conditions. The Pcap and Πif are both forces that promote fluid movement out of the capillaries, whereas Pif and Πcap serve to retain fluid within the capillary lumen.

When the normal values shown in Fig. 50.4 are inserted into Starling's equation, the net force is positive and favors fluid filtration out of the capillaries and into the interstitium of the lung. The fluid that is continuously filtered from capillaries does not accumulate in the low compliance interstitium of the alveolar septum but moves toward the perivascular and peribronchial tissues, where *lymphatic vessels* are located. Fluid transport out of the lung along lymphatic vessels, which is aided by lymphatic vasomotion and valves, maintains a low interstitial fluid pressure in the perivascular and peribronchial space.

Fluid filtration out of the pulmonary capillaries varies with changes in vascular permeability (σ) and with hydrostatic (P) and oncotic pressures (Π). Increases in capillary hydrostatic pressure occur during exercise and in animals with left-sided heart failure. These elevated pressures result in an increased fluid flux into the interstitium. During exercise, the lymphatic vessels remove this filtered fluid so rapidly that there is no net fluid accumulation. This is particularly notable in the intensely exercising horse in which pulmonary capillary pressure exceeds 60 mm Hg and fluid flux from the pulmonary capillaries has been reported to reach 8 L/min. Similarly, in heart failure, lymphatics accommodate large increases in fluid flux. As fluid transfer out of capillaries increases even further, excess fluid accumulates around the bronchi and large blood vessels in the compliant peribronchial and perivascular spaces. With large or sustained increases in fluid flux across the pulmonary capillary endothelium, the proteoglycan bridges in the alveolar septum break, which greatly increases the compliance of this region and allows even greater fluid movement into the interstitium. This fluid is termed *pulmonary interstitial edema* and is seen radiographically as peribronchial cuffing. When fluid pressure within the pulmonary interstitium exceeds the barrier capacity of the alveolar epithelium, then clinically evident *alveolar edema* occurs as fluid enters the air spaces across the alveolar epithelial cells or at the level of the bronchioles. This is seen radiographically as lung opacification and/or *air bronchograms,* in which air-filled bronchi (appearing black on a radiograph) are more visibly distinct when adjacent to fluid-filled alveoli (appearing gray or white on a radiograph). The foaming fluid typical of alveolar edema results from the mixing of air, edema fluid, and surfactant within the airways.

Increased fluid filtration and pulmonary edema can also result from a decrease in plasma oncotic pressure, which is a result of *hypoproteinemia.* Hypoproteinemia can be caused by disease, malnutrition, or by the overly vigorous administration of intravenous fluids that dilute the plasma proteins. Increased pulmonary vascular permeability occurs in many inflammatory lung diseases, such as pneumonia. This results from the effects of inflammatory cell products, such as O_2 radicals, on the endothelium. Protein-rich fluid leaks into the interstitium, elevating interstitial fluid oncotic pressure and causing osmotic attraction of water into the interstitium from the vasculature.

Pleural Fluid Originates by Filtration From Capillaries in the Parietal Pleura and Is Reabsorbed Through Stomata That Connect to Lymphatics

During respiration, the lungs glide along the chest wall with little frictional resistance. This is because the two pleural membranes, termed *visceral* (inner) and *parietal* (outer), are lined with negatively charged phospholipids that create slight repulsion between the membranes, allowing them to glide easily over one another. Additionally, the *pleural space* contains a small amount of fluid that further lubricates the pleural surfaces. The volume of pleural fluid is low

(approximately 0.26 mL/kg) and turns over about every hour. The continual uptake of *pleural fluid* by lymphatics helps to keep the visceral and parietal surfaces in close apposition. The protein content of pleural fluid is normally low (1.5 g/dL), but the net Starling forces favor filtration of fluid into the pleural space. Under normal conditions, fluid enters the pleural space by filtration from capillaries in the parietal (but not visceral) pleura and is removed by lymphatics that communicate directly with the pleural space through *stomata* (holes) in the surface of the parietal pleura. The density of stomata is particularly high near the tendinous part of the diaphragm and the mediastinal side of the pleural cavity. Fluid movement into the pleural space is governed by the same Starling forces that determine fluid flux across pulmonary capillary endothelium. Although the reserve capacity for clearance by lymphatics greatly exceeds the normal rate of pleural fluid production, fluid can accumulate in the pleural cavity when capillary pressures increase, when plasma oncotic pressures fall, or when vascular permeability is increased by inflammation of the pleura, termed *pleuritis.* Despite that, the fact that both pleural membranes are supplied by the systemic circulation (parietal pleura supplied by the intercostal arteries and visceral pleura supplied by the bronchial arteries), pleural fluid accumulation (plural effusion) is more closely tied to pulmonary vascular pressures than to systemic vascular pressures. This may be related to the fact that the visceral pleural veins drain into the pulmonary circulation (although the parietal pleural veins drain into the systemic circulation via the anterior vena cava). Furthermore, pleural effusion due to altered Starling forces occurs only after the development of pulmonary edema.

Several other pathological states are associated with pleural fluid accumulation. *Chylothorax* is lymph fluid accumulating in the pleural space as a result of disruption of intrathoracic lymph vessels, and similarly, *hemothorax* (accumulation of blood in the pleural space) is the result of disruption of intrathoracic blood vessels. Finally, ascites fluid in the peritoneal cavity can translocate to the pleural space through lymphatics or diaphragmatic rents (tears). In the presence of inflammatory conditions such as pleuritis, fibrin accumulates in the pleural space, lymphatic vessels become obstructed, and drainage of the pleural space is impaired. As a result, large volumes of fluid can accumulate between the lungs and chest wall, impeding ventilation and necessitating drainage by the use of chest tubes.

Metabolic-Related Functions of the Lung

The Lung Metabolizes Many Biochemical Substances, Removes Toxins From the Blood, and Its Actions Affect Thermoregulation

Because it receives the total cardiac output, the pulmonary capillary bed with its vast endothelial surface is ideally placed to cleanse the blood of substances produced in other parts of the body. The endothelial cell surface, which is enlarged by projections and depressions known as *caveolae,* is the site of many enzymes involved in the uptake and metabolism of vasoactive substances.

The lungs have an important role in regulating blood pressure and fluid volume through the renin-angiotensin-aldosterone system (see Chapters 34 and 41). In this system, *angiotensinogen* is produced in the liver and is metabolized to *angiotensin I* by *renin* from renal juxtaglomerular cells. Angiotensin I is then activated to *angiotensin II* by dipeptidyl carboxypeptidase, otherwise known as *angiotensin-converting enzyme* (ACE). ACE is present on the luminal surface of pulmonary endothelial cells and thus has ready access to the entire cardiac output. Angiotensin II is a potent vasoconstrictor and also stimulates the release of *aldosterone* and *antidiuretic hormone* (a.k.a., *vasopressin*), which further act to support blood pressure by retaining sodium and water, and promoting vasoconstriction. Pulmonary ACE also *inactivates* the potent vasodilator *bradykinin.* Thus the actions of ACE to increase angiotensin II and decrease bradykinin provide strong stimuli for fluid retention and increased blood pressure. ACE inhibitors constitute an important class of pharmaceuticals used to reduce blood pressure.

Several additional components of the renin-angiotensin-aldosterone system have been identified in recent years. One of the more important metabolites of this system is angiotensin-1-7 (ANG-1–7), which is produced when angiotensin-converting enzyme 2 (ACE2) acts on angiotensin II. ACE2, like ACE, is extensively expressed in the lungs. ANG-1-7 has vasodilatory properties and thus may have a therapeutic role in pulmonary hypertension.

The lungs also inactivate several neurotransmitters and inflammatory mediators. The neurotransmitter *serotonin* is almost totally removed by uptake into endothelial cells, where it is degraded by monoamine oxidase. *Norepinephrine* is also removed to some degree. *Leukotrienes* are broken down by neutrophils, which are numerous in the pulmonary circulation. Various cell types in the lung are major sources of both production and destruction of *prostaglandins* (PG), in particular PGE_2 and PGF_2.

Many exogenous toxic substances are also removed from the blood by the pulmonary endothelium. This process at times can cause severe lung injury. For example, the toxins from *Crotalaria* species of plants can cause smooth muscle hypertrophy in the pulmonary arterioles, which leads to pulmonary hypertension.

The body's metabolic processes can generate a great deal of heat. The respiratory system is important for thermoregulation as changes in breathing frequency (rate) and tidal volume (depth) affect heat loss. This topic is covered in Chapter 53.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Pleuritis in a Thoroughbred Horse

History

At your practice in California, you are asked to examine a 3-year-old thoroughbred stallion that arrived the previous day having traveled by horse van from a racetrack in New York. On arrival, the horse was lethargic. This morning, the stallion refused to eat or drink and was breathing rapidly. The owner reports that in New York the horse was at a racetrack where there was much through-traffic of young horses, many of which were coughing.

Clinical Examination

On arrival at the farm, the owner leads you to a stall where the horse is standing with its elbows slightly abducted, its head lowered, and its nostrils flared. The grain and hay from the morning's feed are untouched. The horse has a respiratory rate of 30 breaths/min (normal, 12–20). On further questioning, the owner reports that the horse looked much as it does now when it arrived from New York, but he thought it was just tired from traveling. The trucker reported that the horse drank little when it was offered

Continued

CLINICAL CORRELATIONS—cont'd

water on its 3-day trip across country and had only nibbled at its hay. It was in the van with four other horses. The condition of these horses is unknown.

You examine the horse and find that it is febrile, and its pulse rate and respiratory rate are elevated. The horse becomes anxious when approached and particularly when hands are laid and pressed on the thoracic cage. The horse's mucous membranes are a dull red. Auscultation of the abdomen reveals minimal gastrointestinal sounds, and there is no evidence of feces in the stall. You listen to the respiratory system and note louder, harsher sounds than normal in the trachea and in the dorsal part of the lung bilaterally. However, the ventral part of the lung is notably silent bilaterally.

You perform an ultrasound examination of the thorax and notice that there is fluid accumulation in the ventral half of the thorax, with areas of consolidated lung (in which the air spaces are filled with liquid and/or solid materials, such as exudate and cellular debris) and adhesions between the pleural membranes. With the assistance of ultrasound, a teat cannula is placed in the right pleural cavity to drain the pleural fluid, which is foul smelling, turbid, and contains bits of fibrinous tissue; 10 L is removed. This is repeated on the left side, with similar results.

Bacterial cultures of the pleural fluid grow an anaerobic organism (*Bacteroides fragilis*), which is probably responsible for the foul smell of the pleural fluid. A complete blood cell count reveals a decreased number of circulating neutrophils and a large number of immature forms of neutrophils (bands). This is an indication that the neutrophil resources of the body are being depleted and the bone marrow is putting out immature forms. Presumably, the neutrophils are being sequestered within the lung and pleural cavity.

Comment

The history and clinical findings in this horse are fairly typical of a case of pleuropneumonia. In New York the horse was exposed to other animals that were coughing, probably as a result of a viral infection, such as equine influenza or equine herpesvirus (also known as equine rhinopneumonitis virus). Respiratory viruses generally impair the defense mechanisms of the lung in two ways. First, they denude the tracheobronchial epithelium of cilia and therefore reduce mucociliary clearance of the airways. Second, they impair macrophage function. This combination of events results in the deposition of bacteria in the lung and failure of the lung to remove them by either the ciliary system or the macrophages. As a result, the bacteria multiply. The stress of shipping makes the situation even worse. Stress results in the release of corticosteroids from the adrenal gland, and this further suppresses the defense mechanisms of the lung. Failure to drink leads to dehydration, which can make mucus more difficult to clear. Keeping the horse's head elevated during transport also impairs mucociliary clearance. As a result of these events, the horse acquired a bacterial infection of the lung, which resulted in the migration of large numbers of neutrophils into the alveoli.

When the infection spread to the pleural cavity, neutrophils migrated to this region as well. The release of neutrophil products designed to kill bacteria causes extensive damage to the membranes of the alveolar epithelium, pulmonary capillaries, and pleural capillaries. The protein that leaks into the alveolar spaces, interstitium of the lung, and pleural cavity raises the osmotic pressure within these regions. This results in the movement of fluid from the vascular space into the alveolar spaces, interstitium, and pleural cavity. Within the pleural cavity, the fluid accumulates ventrally because of gravity and obscures lung sounds in the ventral part of the thorax.

Treatment

The horse is moved to the hospital ward so that intensive care can be provided. Indwelling chest tubes are placed in the pleural cavity bilaterally so that the fluid can be drained repeatedly. The horse is treated with broad-spectrum antibiotics pending culture and sensitivity results, nonsteroidal antiinflammatory drugs (flunixin meglumine), and supportive care. You counsel the owner that therapy will likely need to be continued for a number of weeks and that the chances of the horse returning to athletic performance are no better than 60%.

Mitral Insufficiency in a Dog

History

A 12-year-old spayed female Cocker Spaniel is brought to your veterinary hospital because of a recent deterioration in its condition. The dog has been a faithful pet and has always enjoyed exercising with its owner, but over the past few months the owner has noticed a growing reluctance to exercise. The dog has also coughed, especially when it gets up from resting. In the past few days the dog has refused to leave the house and has eaten little. The owner has noticed that the cough is much more frequent than previously and seems to be moist.

Clinical Examination

You have examined this dog on many occasions, and it has always been friendly, but when you walk into the examination room the dog greets you with only a modest tail wag. It stands with its head down and its tongue hanging out; it is panting. It walks reluctantly toward you when you call it. The dog was formerly fat but is now in about normal flesh, so over the past few months it has lost weight.

You lift the dog onto the examination table and begin by looking at the mucous membranes, which appear moist and normal in color, and capillary refill time. The dog's temperature is normal. The dog is panting, which makes auscultation of the chest difficult, but on the occasions when the dog breathes without panting you notice some increased sounds in the trachea and in all the lung fields, which sound like fluid bubbling within the air spaces of the lungs. The heart rate is increased and a loud murmur is audible over the mitral area during systole. You tell the owner that you suspect that the dog has a heart problem, which is leading to accumulation of fluid in the lungs.

You take chest radiographs and an arterial blood sample for measurement of blood gases. The chest radiograph shows an enlarged heart, particularly the left ventricle. The lungs are diffusely more dense than normal, and the densities have a fluffy appearance. An air bronchogram is noted. The air bronchogram and the appearance of the lung densities suggest alveolar pathology. There is also increased density along the walls of the major airways. Arterial O_2 partial pressure (Pa_{O_2}) is 70 mm Hg (normal 85–100 mm Hg) and arterial carbon dioxide partial pressure (Pa_{CO_2}) is 30 mm Hg (normal 35–45 mm Hg).

An echocardiogram is performed to further assess the heart and mitral valve. The echocardiogram demonstrates thickening of the mitral valve septal leaflet as well as the lateral leaflet. There is a flail of the mitral valve into the atria during ventricular systole. The left ventricle appears dilated and left ventricular contractility is decreased. Regurgitant blood flow through the mitral valve is noted.

Comment

Left-sided heart failure is accompanied by mitral valve insufficiency and thus blood leaking back into the left atrium during systole creates a murmur. The elevation in left atrial pressure as a result of mitral regurgitation is leading to increased pressure in the pulmonary veins and capillaries. The increased pulmonary capillary hydrostatic pressure causes fluid filtration into the interstitium and now into the alveolar air spaces. This condition has likely been progressing for some time and only when it became severe enough for fluid to accumulate in the air spaces of the lung did the owner notice the deterioration in the dog's condition.

The hypoxemia is a result of ventilation/perfusion mismatching because of accumulation of fluids within the alveolar spaces. These fluid-filled spaces are still perfused, but the blood passing through this region does not pick up a sufficient amount of O_2 as the fluid increases the barrier to diffusion. This results in hypoxemia. The hypoxemia stimulates ventilation, and the increase in total ventilation to the lung is sufficient to eliminate more carbon dioxide (CO_2) than normal, so Pa_{CO_2} is slightly reduced at 30 mm Hg.

Treatment

The dog is treated with a diuretic, an angiotensin-converting enzyme inhibitor, and a digitalis glycoside. The diuretic and angiotensin-converting enzyme inhibitor reduce vascular volume and intravascular pressures; they therefore reduce the amount of fluid being filtered into the lung as well as reduce the work of the heart. Over time this leads to resolution of the edema. The digitalis glycoside increases the contractility of the heart and thus the dog's cardiac output, which improves its ability to exercise. You counsel the owner that these treatments will improve the dog's quality of life, but they are not curative and the heart failure will continue to progress over time.

Practice Questions

1. Particles greater than 5 μm in diameter are deposited in the respiratory tract by:
- **a.** Inertial deposition in small airways.
- **b.** Sedimentation in airways.
- **c.** Diffusion in the alveoli.
- **d.** Inertial deposition in large airways.
- **e.** Sedimentation in the alveoli.

2. The mucociliary system:
- **a.** Consists of a gel layer in which cilia beat, overlaid by a sol layer that entraps particles.
- **b.** Is restricted to the nasal cavity and trachea and does not extend into the bronchi.
- **c.** Consists in part of mucus produced by goblet cells in the respiratory bronchioles and by club cells in the trachea.
- **d.** Has a more rapid transport rate in the trachea than in the bronchioles.
- **e.** Lacks ciliated cells in the bronchioles, so mucus must be pulled into the larger airways by viscous drag.

3. Phagocytosis of inhaled particles:
- **a.** Is generally by type II alveolar cells.
- **b.** Can always be accomplished by alveolar macrophages.
- **c.** Sometimes requires both macrophages and neutrophils.
- **d.** Is accentuated by alveolar hypoxia.
- **e.** None of the above.

4. Movement of fluid between the pulmonary capillaries and lung lymphatic vessels:
- **a.** Does not occur in a normal animal.
- **b.** Is accentuated by an increase in capillary hydrostatic pressure.
- **c.** Is accentuated by an increase in capillary oncotic pressure.
- **d.** Occurs by way of the alveolar surface.
- **e.** Is facilitated by a positive pressure in the peribronchial spaces of the bronchovascular bundle.

5. Which of the following occurs as a result of enzymes localized on the pulmonary endothelium?
- **a.** Conversion of angiotensin I to angiotensin II
- **b.** Conversion of angiotensinogen to angiotensin I
- **c.** Release of renin
- **d.** Conversion of renin to angiotensin II
- **e.** None of the above

Bibliography

Agostoni E, Zocchi L. Pleural liquid and its exchanges. *Respir Physiol Neurobiol*. 2007;159:311–323.

Chambers LA, Rollins BM, Tarran R. Liquid movement across the surface epithelium of large airways. *Respir Physiol Neurobiol*. 2007;159(3):256–270.

Dela Cruz CS, Koff JL. Innate and adaptive immunity in the lung. In: Grippi MA, Elias JA, Fishman JA, et al, eds. *Fishman's Pulmonary Diseases and Disorder*. 5th ed. New York: McGraw-Hill; 2015.

Kopf M, Schneider C, Nobs SP. The development and function of lung-resident macrophages and dendritic cells. *Nat Immunol*. 2015;16:36–44.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lekeux P, Art T, Hodgson DR. The respiratory system: anatomy, physiology and adaptations to exercise and training. In: Hodgson DR, McGowan CM, eds. *The Athletic Horse: Principles and Practice of Equine Sports Medicine*. St. Louis: Saunders; 2014.

Lumb AB. *Nunn's Applied Respiratory Physiology*. 8th ed. Elsevier; 2017.

Miserocchi G. Mechanisms controlling the volume of pleural fluid and extravascular lung water. *Eur Respir Rev*. 2009;18(114):244–252.

Siegford J, Powers W, Grimes-Casey HG. Environmental aspects of ethical animal production. *Poult Sci*. 2008;87(2):380–386.

51 Fetal and Neonatal Oxygen Transport

SUSAN L. EWART

KEY POINTS

1. The fetus depends on the placenta for the exchange of gas, nutrients, and metabolic byproducts.
2. The efficiency of gas exchange at the placenta depends on the species-variable arrangement of fetal and maternal blood vessels.
3. The fetal circulation mixes oxygenated and deoxygenated blood at several points, so the fetus exists in a state of relative hypoxemia.
4. Fetal oxygen transport is assisted by fetal hemoglobin, which has a high affinity for oxygen.
5. The lung develops in three stages, and pulmonary surfactant must be present at birth.
6. At or shortly after birth, umbilical vessels rupture, pulmonary vascular resistance decreases, and the foramen ovale, ductus arteriosus, and ductus venosus close.

The Fetus Depends on the Placenta for the Exchange of Gas, Nutrients, and Metabolic Byproducts

From conception until birth, the *fetus* depends on the mother for a supply of oxygen (O_2) and nutrients and for removal of carbon dioxide (CO_2) and other metabolic byproducts. Early in gestation, these substances are exchanged by diffusion through the uterine fluids. As the conceptus increases in size, the specialized exchange organ, known as the *placenta,* develops and becomes essential. The placenta brings maternal and fetal blood into close apposition over a large surface area provided by a network of capillaries. Nutrient and gas exchange occur in the placenta, and like the lung, deoxygenated blood is carried to the placenta in arteries and better oxygenated blood leaves the placenta in a vein.

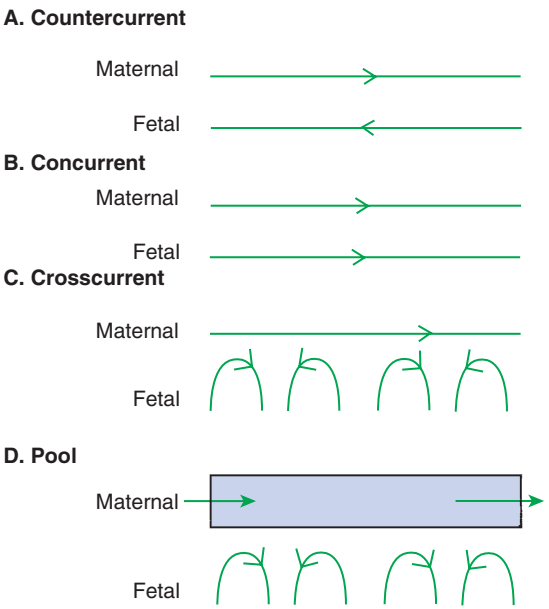
The gross classification of the placenta of different species varies widely (Table 51.1). In horses and pigs the placenta is *diffuse* and covers most of the uterine epithelium. In ruminants the placenta has rows of discrete circular-to-oval *cotyledons* that are attached to approximately 100 highly vascularized *caruncles* in the uterine epithelium. In dogs the placenta is *zonary,* forming a circular band around the *allantochorion* of each puppy.

In addition to differing in the amount of uterine surface to which they are attached, placentas also differ in the number of layers of cells that separate the maternal and fetal blood (see Table 51.1). In horses, pigs, sheep, and cows the fetal *chorion* is adjacent to the maternal uterine epithelium (*epitheliochorial* placentation), whereas in dogs and cats the chorion is adjacent to the endothelium of maternal vessels (*endotheliochorial* placentation). In rodents and most primates the chorion invades the uterine mucosa and erodes the maternal capillaries, so it becomes bathed by maternal blood (*hemochorial* placentation).

The Efficiency of Gas Exchange at the Placenta Depends on the Species-Variable Arrangement of Fetal and Maternal Blood Vessels

The exchange of gases and other substances across the placenta is determined by several factors, including the amount of surface apposition between fetal and maternal tissues, and the number of layers of cells separating fetal and maternal blood. Another factor determining exchange is the arrangement of fetal and maternal blood vessels within the small, interdigitating villi of the placenta (Fig. 51.1). *Countercurrent* flow of maternal and fetal blood provides the most efficient exchange and allows equilibration of fetal blood and maternal *arterial* gas partial pressures. In contrast, *concurrent* flow of fetal and maternal blood allows fetal vessels to equilibrate with the maternal *venous* gas partial pressures, thus providing a much lower partial pressure of O_2 (PO_2) to the fetus. In *crosscurrent* and *pool* types of equilibrators, fetal capillaries loop down to maternal vessels or into a pool of maternal blood. No simple model easily describes these types of exchangers. It is likely that several different arrangements of vessels are found in the placentas of all species, but some species seem to have more of the characteristics of countercurrent exchangers, and others have those of venous (concurrent) equilibrators. For example, the cotyledonary placenta of sheep functions as a venous equilibrator, whereas the hemochorial placenta of the rabbit seems to be a countercurrent exchanger. Fig. 51.2 shows the arrangement of vessels in the microcotyledon of the diffuse placenta of the horse, a species in which fetal and maternal blood flow is primarily countercurrent. Note the deoxygenated (blue) blood in the *umbilical arteries* of the fetus and the relatively more oxygenated (reddish) blood in the fetal *umbilical vein.*

Placental gas exchange has been best studied in sheep (Fig. 51.3), a species in which concurrent flow of fetal and maternal blood predominates. Typically, maternal blood enters the uterus

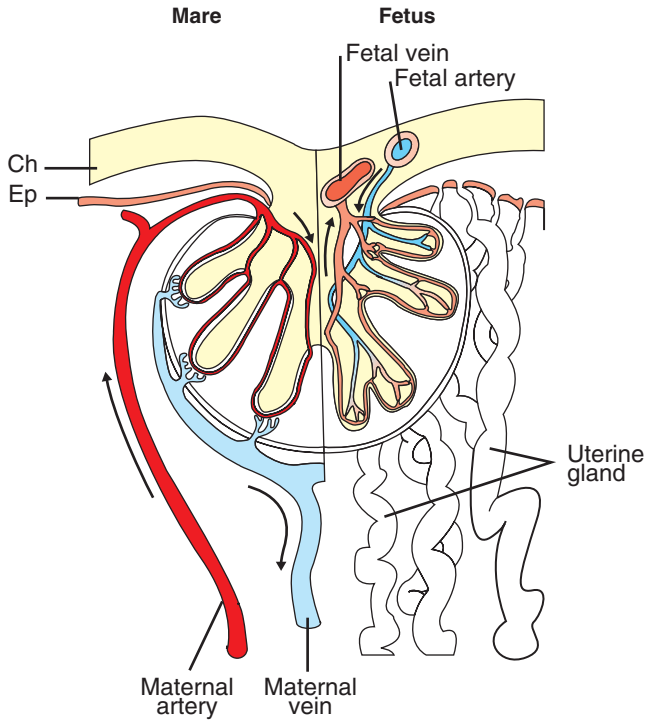


• **Fig. 51.1** Schematic representation of possible arrangements of fetal and maternal blood vessels. Arrows indicate the direction of blood flow. (From Dawes GS. *Fetal and Neonatal Physiology: A Comparative Study of the Changes at Birth.* Chicago, IL: Year Book Medical; 1968.)

TABLE 51.1 Placentation in Domestic Mammals

	CLASSIFICATION	
Species	Gross	Histological
Horse	Diffuse	Epitheliochorial
Pig	Diffuse	Epitheliochorial
Cow	Cotyledonary	Epitheliochorial
Sheep	Cotyledonary	Epitheliochorial
Goat	Cotyledonary	Epitheliochorial
Dog	Zonary	Endotheliochorial
Cat	Zonary	Endotheliochorial
Rabbit	Discoid	Hemochorial
Guinea pig	Discoid	Hemochorial

through the uterine artery with a P_{O_2} of 99 mm Hg and leaves through the uterine veins with a P_{O_2} of 53 mm Hg. Some of the maternal blood entering the uterus supplies the myometrium and endometrium, but most participates in gas exchange in the cotyledon. Deoxygenated fetal blood reaches the placenta through the umbilical artery and enters the cotyledon with a P_{O_2} of 23 mm Hg. Placental gas exchange occurs, and the blood leaving the placenta in the umbilical vein has a P_{O_2} that has been raised to 35 mm Hg. Because the sheep placenta is a venous equilibrator the maximal possible P_{O_2} would be 53 mm Hg, however, this maximum is not reached because fetal blood that has provided nutrient blood flow to the chorion dilutes the better-oxygenated blood draining from the cotyledon. The countercurrent exchanger of the mare is somewhat more efficient as umbilical artery P_{O_2} averages 33 mm Hg and umbilical venous P_{O_2} averages 48 mm Hg.

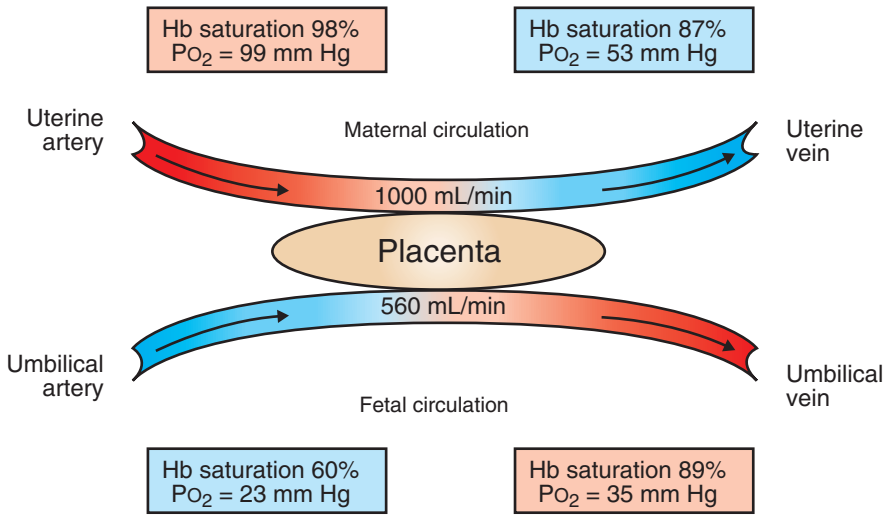


• **Fig. 51.2** Diagram showing the arrangement of maternal and fetal blood vessels in the microcotyledons of the diffuse equine placenta. *Upper middle arrows* demonstrate the countercurrent directions of maternal and fetal blood flow. For illustrative clarity, the maternal and fetal vessels are shown to be on the opposite sides of the central vertical line in this figure, whereas in situ the maternal and fetal vessels are in close apposition to each other. *Ch,* Chorioallantois; *Ep,* uterine epithelium. (Adapted from Tsutsumi Y. The vascular pattern of the placenta in farm animals (horse, pig, cow, sheep and rabbit). *J Facul Agr, Hokkaido Univ*. 1962; 52(3):372–482.)

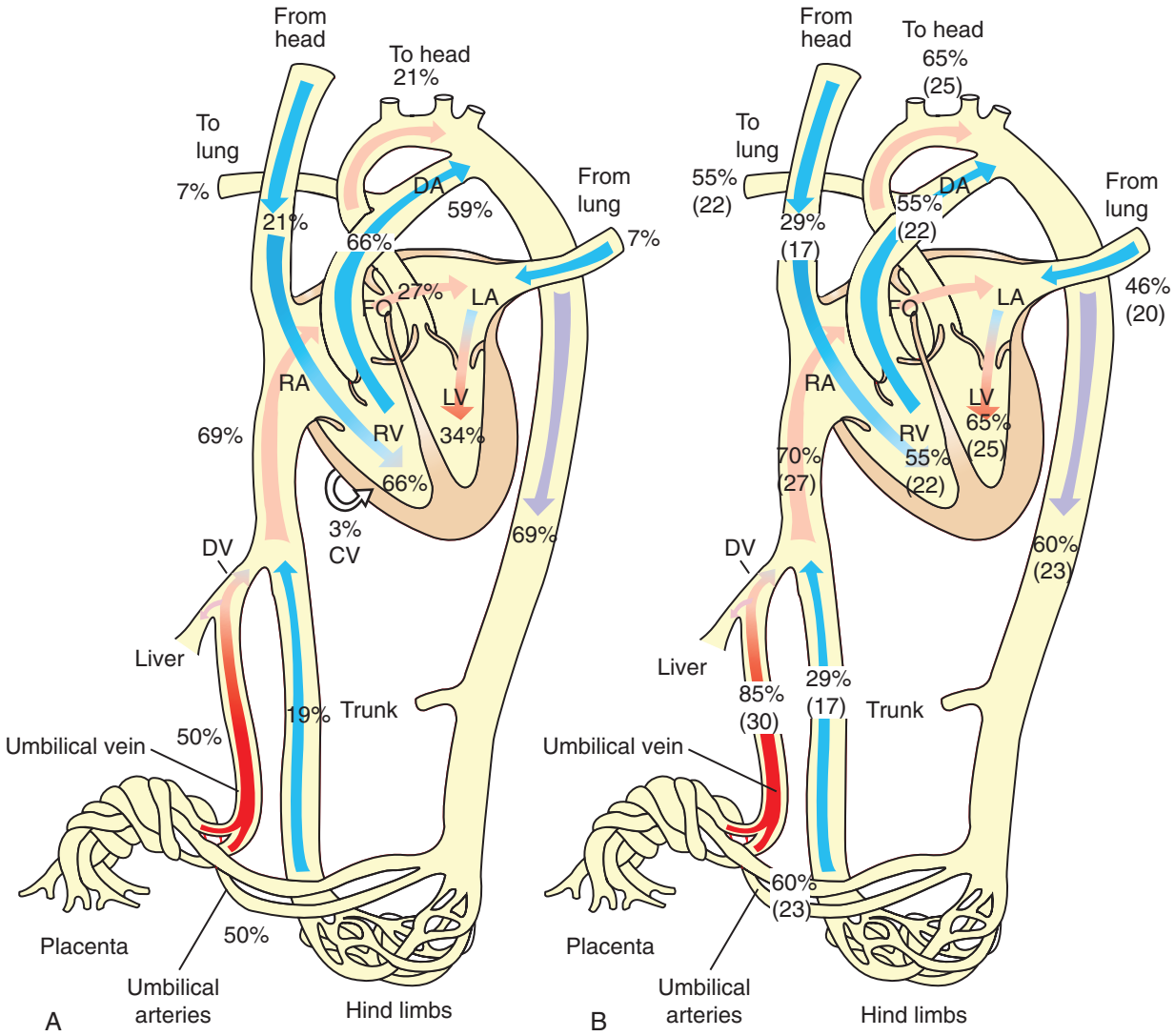
The amount of placenta available for exchange partly determines the ultimate size of the fetus. If uterine caruncles are surgically removed from sheep so that there are fewer sites for formation of fetal cotyledons, the full-term weight of lambs is reduced. The diffuse placenta of the horse can support only one full-size fetus. One foal in a set of twins usually dies in utero or is very small. It is rare for horse twins to survive to term and be of equal size.

The Fetal Circulation Mixes Oxygenated and Deoxygenated Blood at Several Points, so the Fetus Exists in a State of Relative Hypoxemia

In the adult the cardiac output of the right and left ventricles is separate and perfuses the pulmonary and systemic circulations, respectively. In the fetus the output of the two sides of the heart mixes at several points, so when referring to fetal circulation the term *combined cardiac output* is used to indicate the total (combined) output of the right and left ventricles. The combined cardiac output averages 500 mL/min/kg in fetal sheep. The output of the right ventricle exceeds that of the left (Fig. 51.4A). Because the right and left ventricles of the fetus pump into a common circulation, the two are of equal size and wall thickness.



• **Fig. 51.3** Placental blood flow, O_2 partial pressure (P_{O_2}), and hemoglobin (*Hb*) saturation in the uterine (maternal) and umbilical (fetal) circulation of sheep. (Adapted from Comline RS, Silver M. P_{O_2}, P_{CO_2} and pH levels in the umbilical and uterine blood of the mare and ewe. *J Physiol.* 1970; 209:587–608, and Battaglia FC, Meschia G. *An Introduction to Fetal Physiology.* Orlando, FL: Academic Press; 1986.)



• **Fig. 51.4** Diagrammatic representations of the fetal circulation showing (A) percentage of cardiac output in different parts of the circulation, and (B) percent O_2 saturation of hemoglobin and partial pressure of O_2 in mm Hg (in parentheses). *CV,* Coronary vessels; *DA,* ductus arteriosus; *DV,* ductus venosus; *FO,* foramen ovale; *LA,* left atrium; *LV,* left ventricle; *RA,* right atrium; *RV,* right ventricle.

The placenta, which has a *low vascular resistance,* receives approximately 50% of the combined cardiac output through the *umbilical arteries.* The *umbilical vein* carries blood that has been oxygenated in the placenta back to the fetus where it first encounters the liver. In species such as sheep, about half of the umbilical venous blood bypasses the liver via a low-resistance channel known as the *ductus venosus;* in other species, such as pigs and horses, the ductus venosus disappears early in gestation and umbilical venous blood flows through the liver capillaries. Within the liver, the oxygenated blood from the placenta is mixed with a small amount of more poorly oxygenated blood draining the liver sinusoids. The blood from hepatic veins or the ductus venosus enters the posterior vena cava, where it mixes with poorly oxygenated blood draining the hind end of the fetus. Thus the blood returning to the right atrium has a P_{O_2} of approximately 27 mm Hg. Fig. 51.4B shows the O_2 saturation of hemoglobin and P_{O_2} throughout the fetal circulation.

A low-resistance pathway, the *foramen ovale,* connects the right and left atria, and a structure known as the *crista dividens* directs the better-oxygenated blood from the posterior vena cava through the foramen ovale into the left atrium. The most poorly oxygenated blood returning to the right atrium in the cranial vena cava is directed into the right ventricle. In contrast to blood flow patterns postnatally, most of the output of the right ventricle does not go through the lungs; this is because fetal lungs have a high vascular resistance. Another low-resistance channel, the *ductus arteriosus,* connects the *pulmonary artery* with the *aorta* and allows the majority of blood leaving the right ventricle to bypass the lungs and instead enter the aorta. It is important to note that the arrangement of the fetal circulation allows the better-oxygenated blood (P_{O_2} 25 mm Hg) to travel from the left ventricle to the ascending aorta and then brachiocephalic vessels that supply the anterior portion of the fetus, including the brain and forelimbs. The less well-oxygenated blood from the ductus arteriosus (P_{O_2} 22 mm Hg) enters the aorta downstream from the brachiocephalic vessels; thus the tissues of the posterior portion of the animal and the placenta receive blood with slightly lower oxygenation.

Flow of blood from the right atrium to the left atrium through the foramen ovale and from the pulmonary artery to the aorta through the ductus arteriosus requires that the pressure in the right side of the fetal circulation is greater than that in the left side. This pressure difference is opposite of that in the postnatal state and occurs in the fetus because the left side of the circulation provides most of its output to the *low-resistance* placenta, whereas the right side of the fetal circulation is opposed by the *high-resistance* pulmonary circulation. At the end of gestation, systemic arterial pressure in the lamb is about 42 mm Hg.

The fetal circulation is not a passive system and is capable of considerable regulation, particularly as the fetus matures. Fetal hypoxia can stimulate vasodilation in the heart and brain and vasoconstriction in the gut, kidneys, and skeletal tissues. The fetal pulmonary circulation constricts vigorously when the fetus is hypoxic. This constriction diverts more blood through the ductus arteriosus to the systemic tissues.

Fetal Oxygen Transport Is Assisted by Fetal Hemoglobin, Which Has a High Affinity for Oxygen

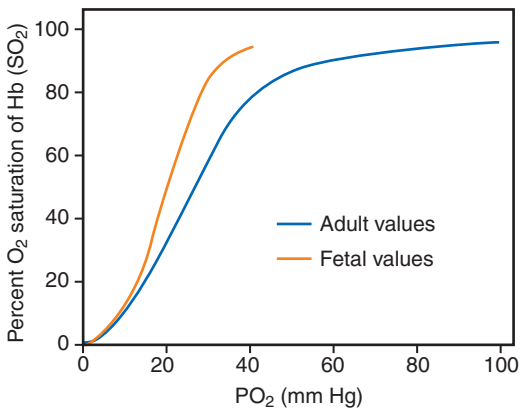
Fetal arterial blood has a lower P_{O_2} than in the postnatal state because the placenta is not a highly efficient gas exchanger and because oxygenated and deoxygenated blood mix at several points in the fetal circulation. The fetus is adapted to this state of chronic hypoxia in the following ways. Compared with the postnatal state, the fetus has a high cardiac output that delivers a large volume of blood per minute to the tissues. The fetus also has a higher hematocrit, which increases its O_2 carrying capacity. Furthermore, its erythrocytes contain hemoglobin with a *high affinity* for O_2. Finally, lower tissue P_{O_2} levels create a diffusion gradient for O_2 unloading from fetal hemoglobin.

Depending on the species, erythrocytes in the developing conceptus contain *embryonic, fetal,* or *adult hemoglobin.* Hemoglobin of adult mammals has four globin subunits together consisting of two alpha chains and two beta chains ($\alpha_2\beta_2$) of polypeptides. Most species have one or more forms of embryonic (ε) hemoglobin, and some species (such as sheep) have a distinct fetal hemoglobin that has gamma chains in place of the beta chains ($\alpha_2\gamma_2$). These embryonic and fetal globin chains bind O_2 with higher affinity. After birth there is gradual replacement of fetal hemoglobin by adult hemoglobin.

Horses and dogs do not have $\alpha_2\gamma_2$ fetal hemoglobin, instead their erythrocytes have reduced levels of *2,3-bisphosphoglycerate* (2,3-BPG, also called 2,3-DPG), which is a molecule that decreases the binding affinity of hemoglobin for O_2; its absence thus serves to increase the affinity of hemoglobin for O_2. Other species, including ruminants, have 2,3-BPG within their fetal erythrocytes, but their fetal hemoglobin is unresponsive to it.

These various mechanisms by which fetal hemoglobin achieves a higher affinity for O_2 result in a lower partial pressure at which hemoglobin is 50% saturated with O_2 (P_{50}) (see Chapter 48). This is reflected in the *left shift* of the *oxyhemoglobin dissociation curve* for fetal blood compared with maternal blood (Fig. 51.5). In some species, such as cats, the difference in the P_{50} between fetus and adult is small, whereas in ruminants the difference is 10 to 20 mm Hg.

The high affinity of fetal hemoglobin for O_2 allows the hemoglobin in the umbilical vein, with a P_{O_2} of 30 mm Hg, to be 85% saturated with O_2 and allows the hemoglobin in the aorta, with a P_{O_2} of 23 mm Hg, to be 60% saturated (Figs. 51.4B and 51.5). Thus although fetal O_2 *partial pressures* are considerably lower than postnatal values, pre- and postnatal O_2 *saturation* levels are less discrepant. The high affinity of fetal hemoglobin for O_2 not only



• **Fig. 51.5** Fetal and adult oxyhemoglobin dissociation curves. Note that the dissociation curve of the fetus is displaced to the left of the adult so that at a given P_{O_2}, the fetal hemoglobin has a higher percent saturation than does the adult. P_{O_2}, O_2 partial pressure.

allows O_2 transport at the low PO_2 in the fetal vasculature but also makes it necessary for fetal tissues to have an extremely low PO_2. The low tissue PO_2 provides an O_2 concentration gradient to unload O_2 from the fetal hemoglobin. Therefore the fetus exists in a state of tissue hypoxia compared with the adult.

The Lung Develops in Three Stages and Pulmonary Surfactant Must Be Present at Birth

By the time of birth, the lung must be ready to assume the gas-exchange functions of the placenta. The lung develops in three stages of equivalent duration. Beginning as an outgrowth of the foregut, the lung bud invades the *mesenchyme* of the thorax and divides into all the major airways during the first third of gestation. Concurrently, major blood vessel branches develop. Because the primordial airways are lined with a cuboidal epithelium and look like a gland in cross-section, this stage of development is known as the *pseudoglandular* stage. In the second or *canalicular* stage of development, the bronchioles develop, the airways develop lumen, and capillaries invade the developing bronchiolar region. In the final stage, or *alveolar sac* stage, alveolar sacs, and in some species alveoli, develop. The stage of maturity of the lung at birth, in general, matches the maturity of the fetus. Lambs and piglets have well-developed alveoli, but humans and especially rodents have thicker-walled alveolar sacs. In these latter species, alveoli develop as the animal grows postnatally.

Pulmonary surfactant is essential if the lung is to become and remain inflated after birth (see Chapter 45). Synthesis of surfactant begins during the pseudoglandular stage of development, and at about midgestation there is an increase in the synthesis of surfactant components, such as lecithin, within the lung. This increase in lecithin synthesis coincides with the appearance of *type II alveolar epithelial cells* (a primary source of surfactant) and with an increase in pulmonary blood flow. Some of this lecithin is secreted into the alveolar lumens and appears in the amniotic fluid, where it can be measured as an indicator of the state of lung maturity. Lung maturity coincides with an increase in the serum glucocorticoid (cortisol) level in the fetus. Indeed, glucocorticoids (two thirds fetal and one third maternal in origin) play an essential role in lung maturation.

Until the time of birth, the vascular resistance of the fetal pulmonary circulation is high for two reasons. First, the fetal lung is not inflated; therefore the pulmonary vasculature is relatively compressed at these reduced lung volumes. In addition, the hypoxia of the fetus maintains the pulmonary vascular smooth muscle in a state of contraction that narrows the arteries (see Chapter 46). The first few breaths after birth alleviate both these conditions.

The fetal lung continuously secretes fluid, which is rich in chloride and low in bicarbonate and protein, until about 2 days before birth. The fluid travels up the trachea and through the fetus's mouth into the amniotic cavity. This fluid serves to partially distend the lungs and is required for normal lung growth. The fluid in the alveolar spaces and airways is, in part, squeezed out of the lung as the thorax is compressed during birth. The majority is reabsorbed into lymphatic and blood vessels shortly after birth.

Beginning at about one third of gestation, the fetus makes *breathing movements,* although it moves little of the fluid to and fro in the airways. These movements prepare the respiratory muscles for their postnatal function.

At or Shortly After Birth, Umbilical Vessels Rupture, Pulmonary Vascular Resistance Decreases, and the Foramen Ovale, Ductus Arteriosus, and Ductus Venosus Close

Prenatally the fetus is dependent on the placenta and the mother for exchange with the environment, but the lung and other organs must be ready to assume their postnatal functions at birth. During a normal parturition, the newborn emerges from the birth canal at about the time the placenta is detaching from the uterine wall. Placental gas exchange probably continues well into third-stage labor. If labor is prolonged, the placenta may detach before the newborn is delivered. This is a medical emergency.

Normally, the newborn takes the first breath immediately after delivery. The stimuli for this include (1) hypoxia and hypercarbia, which result from the loss of the placental gas exchanger; (2) cooling of the fetus as it encounters ambient temperatures and the fetal fluids evaporate from the skin; and (3) a generalized increase in sensory input to the fetus as it is licked and nuzzled by its mother. To expand the lungs initially, the respiratory muscles must create an intrathoracic pressure 60 cm H_2O less than atmospheric pressure. This is vastly lower than the −2 to −5 cm H_2O generated by adults breathing at resting (tidal) volumes. The large effort during the first breath is necessary to move airway fluids down the tracheobronchial tree before air can enter the alveoli and to open the fluid-filled alveoli. Not all alveoli may inflate during the first breath, but subsequent inhalations inflate the entire lung and distribute surfactant over the alveolar surface. This surfactant makes the alveoli stable and prevents their collapse so that a stable end-expiratory lung volume, known as *functional residual capacity,* can be established (see Chapter 45). After the first few breaths, arterial O_2 partial pressure (P_aO_2) is much higher than it was before birth.

Even though the fetus exists in a state of hypoxia compared with the adult, the carotid bodies are not responsible for stimulating breathing movements in utero. In fact, exaggerated fetal hypoxia inhibits fetal breathing movements and reduces the metabolic rate to reduce O_2 demand. At birth, the carotid bodies are relatively insensitive to hypoxia, but in the presence of O_2 concentrations found in air, they develop the adult sensitivity to O_2 deficits within a few weeks.

Inflation and oxygenation of the lung at birth reduces the pulmonary vascular resistance, which leads to a decreased pressure in the pulmonary artery, right ventricle, and right atrium. At about the same time, the umbilical vessels rupture because the animal struggles to stand or the umbilical cord is torn by the mother. Umbilical blood flow is arrested by local vasoconstriction in the umbilical vessels. The loss of the low-resistance placental circulation increases systemic vascular resistance, which results in an increased pressure in the aorta, left ventricle, and left atrium. Additionally, the immediate postnatal increase in tissue oxygenation that follows breathing reverses the systemic hypoxia-induced vasodilation, which further increases systemic vascular pressures. As a result of these changes, aortic pressure exceeds pulmonary arterial pressure, and left atrial pressure exceeds right atrial pressure; therefore blood flow through the ductus arteriosus and foramen ovale reverses, respectively. Flow reversal in the foramen ovale causes a flap valve to close and occlude the foramen. Over succeeding days to weeks,

this valve becomes adherent to the wall of the atrium, thus permanently closing the foramen.

The ductus arteriosus contains extensive smooth muscle and during gestation is actively maintained in a vasodilated state by both tissue hypoxia and prostaglandins, in particular prostaglandin E_2 (PGE_2). Reversal of blood flow in the ductus arteriosus postnatally exposes the ductus wall to well-oxygenated blood; this removes the tissue hypoxia-driven vasodilation and causes constriction of smooth muscle in the wall of the ductus arteriosus, thus arresting blood flow. Additionally, PGE_2 levels fall postnatally, and the loss of this vasodilatory stimulus further contributes to constriction of the ductus arteriosus. When the ductus arteriosus has constricted and flow has been arrested, it is gradually converted into a nonpatent fibrous band of tissue. Drugs such as indomethacin, which inhibit prostaglandin synthesis, can be used to constrict the ductus arteriosus if it remains patent.

In a species where the ductus venosus remains patent throughout gestation, like in sheep, this vascular shunt closes via smooth muscle constriction within a few hours after birth, thus increasing blood flow through the liver. The functionally closed ductus venosus becomes a fibrous remnant known as the *ligamentum venosum.* Failure of the ductus venosus to close results in portosystemic shunting.

The changes just described convert the fetal circulation into the adult circulation, which is able to support the gas exchange function of the lung. Remarkably, these changes occur routinely and without medical assistance in almost all animal births.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Patent Ductus Arteriosus in a Pomeranian

History

A 7-week-old female Pomeranian puppy is presented to you because it is not growing as fast as its littermates. The breeder says it is lethargic and prefers to sleep, whereas the other puppies play.

Clinical Examination

Clinical examination reveals a small puppy with a rapid heart rate. The mucous membranes of its gums are pink, and its temperature is normal. While holding the puppy around the thorax, you notice a vibration in the region of the heart. When you listen with a stethoscope, you hear a loud murmur that is almost continuous through systole and diastole, and you recall that this is called a *machinery murmur.* It is difficult to listen to the breath sounds because the murmur is audible all over the thorax.

A radiograph reveals a slightly enlarged heart, but the lungs appear normal, although somewhat compressed by the heart. An echocardiogram is performed, which better characterizes the cardiac blood flow and confirms the suspected patent ductus arteriosus (PDA). The velocity of blood flow through the ductus and its mean diameter indicate that the dog is a candidate for surgical correction of the PDA.

Comment

The clinical and radiographic findings in a puppy of this age are characteristic of PDA. In some animals the ductus fails to close after birth and blood continues to flow through it, usually from the aorta to the pulmonary artery. This presents the animal with two problems. First, the left ventricle must increase its output to supply the systemic tissues because so much blood is passing through the ductus. Second, the pulmonary circulation has a volume overload that increases the pressure against which the right ventricle must work. Depending on the magnitude of the PDA, these extra loads result in dilation of the ventricles and sometimes in hypertrophy of the myocardium, which is seen on radiographs as an enlarged heart. The puppy is not growing and is listless because the tissues are not receiving a normal blood flow. Although surgical correction of the PDA allows the puppy to live a normal life, it would be unwise to breed this animal in the future because the condition may be inherited.

Treatment

The patent ductus is closed by surgical ligation, and the puppy is expected to have a normal life. Recently, use of a transarterial coil occlusion procedure has been successful in the treatment of animals with PDA. The treatment choice is somewhat dictated by the size of the shunt as well as the presence or absence of heart failure.

Pregnant Mare With Placentitis

History

A client calls you because their 16-year-old warmblood mare, who is pregnant and due to foal in 30 days, has some foul-smelling vaginal discharge. This is the mare's sixth pregnancy. She has not had pregnancy-related complications previously. The mare had been vaccinated against equine herpes virus (also termed rhinopneumonitis virus) at 5, 7, and 9 months of gestation. The clients recently learned that fescue grass, when it is infected with the fungus *Acremonium coenephialum*, can cause broodmares to abort in late gestation. Alarmed at that risk and the presence of fescue in their pasture, they switched the mare's feed to fescue-free hay and began turning her out in a dry lot instead of the pasture.

Clinical Examination

The mare's temperature, pulse rate, and respiratory rate are normal. The results of a general physical examination are within normal limits, other than some brown vaginal discharge that is mildly malodorous. You recommend to palpate the mare per rectum, perform a reproductive ultrasound of the mare and fetus, perform a vaginal examination, and possibly collect samples for culture and cytology. Your rectal examination reveals that the fetus is in normal presentation. On ultrasound, you note that the fetal heart rate is elevated and that the combined thickness of the uterus and placenta is increased in a few places. You collect samples in the vaginal vault for culture and cytology.

Comment

You are concerned the mare has placentitis, either due to a bacterial infection or fescue toxicosis. The lack of fever is more indicative of fescue toxicosis than infection, but you will put her on antibiotics to decrease the chance of ascending infection and losing the pregnancy. Horses have a diffuse epitheliochorial placenta and placentitis can affect the oxygen flow to the fetus, thus causing the increased fetal heart rate.

Treatment

In this case because either a microbial infection or fescue toxicosis could be the cause of the clinical signs, the mare is treated for both conditions. You prescribe the following medications: (1) oral progesterone as a daily supplementation to help support the placenta and pregnancy; (2) the nonsteroidal antiinflammatory medication flunixin meglumine to reduce inflammation, and (3) broad-spectrum antibiotics. You warn the client that the pregnancy has a guarded prognosis because placentitis can limit the delivery of oxygen and nutrients to the fetus. In this case, the fetus dies in utero and is aborted by the mare.

Continued

CLINICAL CORRELATIONS—cont'd

Dyspneic Foal

History

An 8-hour-old Arabian filly presents to the clinic for respiratory distress. The owners report that the mare foaled normally, but the foal has had an increased respiratory effort and is acting distressed; it has not nursed often and is spending a lot of time laying down. The mare is acting normally and appears to have a good supply of colostrum and milk.

Clinical Examination

The foal has increased temperature, respiratory rate, and heart rate, with a continuous murmur that is loudest over the left heart base. Its mucous membranes are injected (bright red in color), capillary refill time is prolonged, peripheral pulses are weak, and extremities are cool; these findings suggest poor perfusion and possibly sepsis. Gastrointestinal sounds are decreased. Blood work, consisting of a complete blood count and serum chemistry profile, reveals that the foal has a leukocytosis (increased white blood cell count), increased fibrinogen, low globulins and albumin, and normal electrolytes; these findings are indicative of an infection. Arterial blood gas values indicate a respiratory acidosis (see Chapter 52) with low oxygen partial pressure (P_aO_2), termed *hypoxemia* (see Chapter 47). Ultrasound of the foal's chest reveals pneumonia.

Comment

The foal has pneumonia, sepsis, and evidence that fetal circulatory shunts are likely to still be open. These conditions may be related to abnormalities of pulmonary vascular resistance (see Chapter 46). Normally, after birth the pulmonary vascular resistance decreases to approximately 10% of fetal levels, which allows pulmonary blood flow to increase as the pulmonary artery pressures fall. Vasoactive substances are thought to be responsible for regulating pulmonary arteriole tone, and thus pulmonary vascular resistance and blood pressure. Increasing the alveolar and arterial oxygen partial pressures, which occurs after birth as the foal begins to breathe, is required for resolution of the pulmonary hypertension that is characteristic of the fetal state. Although the exact mechanisms for the normal decrease in pulmonary vascular resistance after birth are unknown, nitric oxide and endothelin are thought to be involved. In some foals, pulmonary vascular resistance does not fall and the foal maintains the right-to-left shunting of blood through the patent foramen ovale and/or ductus arteriosus. Furthermore, the condition of sepsis, which is systemic inflammation in response to an infection, may act to sustain the prenatal levels of elevated pulmonary vascular resistance.

Treatment

The foal is administered broad-spectrum antibiotics and supportive care. In an attempt to treat the hypoxemia, the foal is supplemented with oxygen through a nasal cannula. Despite oxygen supplementation, the foal's oxygenation does not improve. The failure to respond to nasal oxygen suggests the foal has maintained (or reverted to) its fetal circulatory right-to-left shunting of blood (see Chapter 47) through the patent foramen ovale and ductus arteriosus due to persistence of pulmonary hypertension. Foals with evidence of pulmonary hypertension and fetal circulatory shunts are often put on ventilators to treat the hypoxemia and resolve the respiratory acidosis. Strategies to dilate the pulmonary vessels using inhaled nitric oxide have been tried.

Practice Questions

1. The vascular channel that allows fetal blood to pass from the pulmonary artery to the aorta is known as the:
- **a.** Foramen ovale
- **b.** Ductus arteriosus
- **c.** Ductus venosus
- **d.** Fetal cotyledon
- **e.** Allantois

2. Which of the following is the correct sequence of events that follow birth?
- **a.** Closure of foramen ovale, first breath, rupture of umbilical vessels
- **b.** Decrease in right atrial pressure, first breath, closure of the ductus arteriosus
- **c.** First breath, closure of the ductus arteriosus, decrease in pulmonary arterial pressure
- **d.** First breath, decrease in pulmonary arterial pressure, closure of the foramen ovale
- **e.** Closure of the foramen ovale, closure of the ductus arteriosus, first breath

3. Which of the following fetal blood structures contains blood with the highest P_{O_2}?
- **a.** Aorta
- **b.** Ductus arteriosus
- **c.** Pulmonary artery
- **d.** Left ventricle
- **e.** Umbilical artery

4. Which of the following statements about the fetal circulation is true?
- **a.** Right atrial pressure is higher than left atrial pressure
- **b.** Pulmonary vascular resistance is high
- **c.** The placenta receives about 45% of the combined output of both ventricles
- **d.** The output of the right ventricle is greater than that of the left ventricle
- **e.** All the above

5. Which of the following does not correctly describe the lung in utero?
- **a.** Type II cells, which produce surfactant, are present within the first few days of gestation in sheep.
- **b.** Chloride-rich fluid is secreted into the airways and flows into the amniotic cavity.
- **c.** Components of surfactant can be detected in the amniotic fluid when the lung approaches maturity.
- **d.** All the major branches of the tracheobronchial tree are present at birth, but alveoli continue to multiply after birth in some species.
- **e.** Breathing movements occur in utero, but the volume of fluid moved in and out of the lungs is small.

6. Fetal O_2 transport is assisted by:
- **a.** Fetal hemoglobin, which has a lower O_2 capacity than adult hemoglobin.
- **b.** Fetal hemoglobin, which has a lower P_{50} than adult hemoglobin.
- **c.** A cardiac output that is less per kilogram of body weight than in the adult.
- **d.** A cardiac output that preferentially delivers the blood with the highest P_{O_2} to the placenta.
- **e.** A fetal lung, which is an efficient gas exchanger.

7. Which of the following domestic mammals has a diffuse, epitheliochorial placenta in which fetal and maternal blood flow is countercurrent in the microcotyledons?
- **a.** Dog
- **b.** Cow
- **c.** Horse
- **d.** Rabbit
- **e.** Sheep

Bibliography

Battaglia FC, Meschia G. *An Introduction to Fetal Physiology*. Orlando, Fla: Academic Press; 1986.

Campbell FE, Thomas WP, Miller SJ, et al. Immediate and late outcomes of transarterial coil occlusion of patent ductus arteriosus in dogs. *J Vet Intern Med*. 2006;20(1):83–96.

Carroll JL, Agarwal A. Development of ventilatory control in infants. *Paediatr Respir Rev*. 2010;11(4):199–207.

Comline RS, Silver M. PO2, PCO2 and pH levels in the umbilical and uterine blood of the mare and ewe. *J Physiol*. 1970;209:587–608.

Dawes GS. *Foetal and Neonatal Physiology: A Comparative Study of the Changes at Birth*. Chicago: Year Book Medical; 1968.

Faber JJ, Thornburg KL. *Placental Physiology: Structure and Function of Fetomaternal Exchange*. New York: Raven Press; 1983.

Hafez S. Comparative placental anatomy: divergent structures serving a common purpose. *Prog Mol Biol Transl Sci*. 2017;145:1–28.

Kay HH, Nelson DM, Wang Y, eds. *The Placenta: From Development to Disease*. Oxford, UK: Wiley-Blackwell; 2011.

Kitchen H, Brett I. Embryonic and fetal hemoglobin in animals. *Ann N Y Acad Sci*. 1974;241(0):653–671.

Koterba A, Drummond W, Kosch P. *Equine Clinical Neonatology*. Baltimore, MD: Williams and Wilkins; 1990.

Leff AR, Schumacker PT. *Respiratory Physiology: Basics and Applications*. Philadelphia: Saunders; 1993.

Lister G, Jones EE. Fetal and neonatal physiology. In: Boron WF, Boulpaep EL, eds. *Medical Physiology*. 3rd ed. Philadelphia: Elsevier; 2017.

Morton S, Brodsky D. Fetal physiology and the transition to extrauterine life. *Clin Perinatol*. 2016;43(3):395–407.

Reed SM, Bayly WM, Sellon DC. *Equine Internal Medicine*. 3rd ed. St. Louis, MO: Saunders.; 2010.

Silver M, Steven DH, Comline KS. Placental exchange and morphology in ruminants and the mare. In: Comline KS, Cross KW, Dawes GS, et al, eds. *Foetal and Neonatal Physiology: Proceedings of the Sir Joseph Barcroft Centenary Symposium*. Cambridge, UK: Cambridge University Press; 1973.

Smith B. *Large Animal Internal Medicine*. 4th ed. St. Louis, MO: Mosby Elsevier.; 2009.

Smith LJ, McKay KO, van Asperen PP, et al. Normal development of the lung and premature birth. *Paediatr Respir Rev*. 2010;11(3):135–142.

52

Acid-Base Homeostasis

SUSAN L. EWART

KEY POINTS

Acid-Base Regulation

1. Relative constancy of the body's pH is essential because metabolism requires enzymes that operate at an optimal pH.
2. Hydrogen ion concentration is measured as pH.
3. An acid can donate a hydrogen ion and a base can accept a hydrogen ion.
4. Buffers reversibly bind hydrogen ions and prevent major changes in pH.
5. Carbonic anhydrases catalyze the reversible hydration of carbon dioxide.
6. Bicarbonate and hemoglobin are the most important blood buffers.
7. The blood buffers provide the first defense against a change in blood pH, but the lungs and kidneys must ultimately correct the hydrogen ion load.
8. Changes in ventilation can rapidly change carbon dioxide partial pressure and therefore alter pH.
9. Metabolic production of fixed acids requires that the kidneys eliminate hydrogen ions and conserve bicarbonate.
10. Buffers and ion exchangers regulate intracellular pH.

Acid-Base Disturbances

1. Primary acid-base disturbances are typically accompanied by compensatory mechanisms that move pH toward normal.
2. Base excess is a measure of the metabolic component of acid-base status and its assessment facilitates the evaluation of mixed acid-base disturbances.
3. Respiratory acidosis is caused by the accumulation of carbon dioxide, which decreases blood pH.
4. Respiratory alkalosis is caused by the loss of carbon dioxide, which increases blood pH.
5. Metabolic acidosis is caused by the accumulation of fixed acids or the loss of buffer base, which decreases blood pH.
6. Metabolic alkalosis is caused by the excessive elimination of hydrogen ions or the accumulation of buffer base, which increases blood pH.
7. Respiratory compensation for acid-base abnormalities occurs rapidly; renal compensation takes hours to days.
8. Appropriate compensation for acid-base disturbances can be predicted.
9. Potassium, chloride, and other ions affect acid-base homeostasis.
10. Arterial blood gases are used to diagnose acid-base disturbances.

Acid-Base Regulation

Relative Constancy of the Body's pH Is Essential Because Metabolism Requires Enzymes That Operate at an Optimal pH

For optimal functioning of cells, the ionic composition of body fluids is maintained within fairly narrow limits. *Hydrogen ion* (H^+, also termed *proton*) concentration is extremely important because it determines the *acidity* or *alkalinity* of the body fluids. The concentration of hydrogen ions ($[H^+]$) is reported as *pH*. The term pH derives from the phrase "potenz hydrogen," which means *power of hydrogen* and is calculated as the negative $\log_{10}$ of the $[H^+]$ in mol/L. Serious deviations of pH outside the normal range disrupt cell metabolism and therefore body function. For example, the activity of the sodium-potassium-ATPase (Na^+,K^+-ATPase) pump decreases by half when pH falls by one unit, and the activity of phosphofructokinase (a key regulatory enzyme in the glycolytic pathway) decreases by 90% when pH decreases by only 0.1 unit.

The pH of arterial blood under normal conditions is 7.4 (normal range 7.36–7.44). Acidemia is the state of increased acidity of the blood, whereas alkalemia is the opposite state of decreased blood acidity (that is, increased alkalinity). The terms *acidosis* and *alkalosis* describe the process by which these blood pH derangements occur. Because pH is reported as a negative $\log_{10}$, more acid results in a lower pH, thus blood pH below and above the normal range are referred to as acidosis and alkalosis, respectively (see the next section for further details of the pH measurement). The range of pH compatible with life is approximately 6.85 to 7.8, but such extreme changes are rare because buffers, the lungs, and kidneys work in concert to tightly regulate pH.

In a 70-kg animal, approximately 15,000 millimole (mM) of H^+ are produced daily when carbon dioxide (CO_2) is added to the blood for transport from the tissues to the lungs. If the lungs eliminate CO_2 as fast as it is produced in the tissues, there is no net H^+ gain by the body. However, the balance between CO_2 production and CO_2 elimination may be disturbed during exercise or in respiratory disease, thus threatening the acid-base homeostasis of the body.

Approximately 40 mM/day of H^+ are produced during metabolism of proteins (such as sulfuric and phosphoric acids), fats (such as keto acids), and the incomplete oxidation of glucose (lactic acid). Another approximately 30 mM/day of H^+ are absorbed from the intestines. Although the amount of H^+ from these sources is less than that produced in CO_2 transport, they must be eliminated continually. In a disease state, the H^+ load imposed on the body is frequently increased because of an increase in tissue

breakdown (catabolism) or because the kidneys fail to eliminate H^+. Loss of H^+, as occurs in vomiting, can create a relative deficit of H^+.

To understand how the body regulates pH and how acid-base disorders are diagnosed and managed, it is necessary to understand acids, bases, and buffering.

Hydrogen Ion Concentration Is Measured as pH

By convention, the concentration of H^+ is expressed in pH units, where pH is the negative $\log_{10}$ of the $[H^+]$ in mol/L. The scale of pH values ranges from 0 (most acidic) to 14 (most basic), with a value of 7 being neutral. Water contains 1×10^{-7} mol/L $[H^+]$ and an equal concentration of hydroxyl ions (OH^-) and thus has a pH of 7.0, that is, a neutral pH. An increase in $[H^+]$ decreases pH, resulting in a more acidic solution. For example, a tenfold increase in $[H^+]$ decreases pH by 1.0 unit, whereas doubling $[H^+]$ decreases pH by 0.3 unit. Conversely, a decrease in $[H^+]$ increases pH, resulting in a more alkaline solution.

The physiological range of blood pH, 6.85 to 7.80, represents $[H^+]$ of 1.4×10^{-7} to 1.6×10^{-8} mol/L (equivalent units of $[H^+]$ are reported as Eq/L). Thus although the $[H^+]$ is regulated, changes up to 10-fold in magnitude can occur, much greater than the fluctuations observed in the concentration of other ions, such as Na^+ or K^+. The very small ranges for normal and physiological pH indicate that pH is tightly regulated; however, small changes in pH reflect large changes in $[H^+]$.

An Acid Can Donate a Hydrogen Ion and a Base Can Accept a Hydrogen Ion

Acids are molecules that can *donate a proton* or *accept an electron pair. Strong acids* are those that completely dissociate in aqueous solutions and thus are powerful proton donors. Hydrochloric acid (HCl) is a strong acid because it dissociates completely in water into H^+ and Cl^-. In contrast, *weak acids* do not completely dissociate in solution. Carbonic acid (H_2CO_3) is a weak acid because it dissociates incompletely in solution into hydrogen and bicarbonate (HCO_3^-) ions with some remaining H_2CO_3, thus not all of its protons are released and available to alter the pH. Note that acid strength refers to the dissociation of the first proton; some strong acids, like H_2SO_4, only lose one of two protons ($H_2SO_4 \rightarrow H^+ + HSO_4^-$), just as some weak acids, like H_2CO_3, only lose a single proton ($H_2CO_3 \rightarrow H^+ + HCO_3^-$).

Bases are molecules that *accept protons* or *donate an electron pair.* Chloride ion (Cl^-) is a base because it has the potential to accept a H^+, but it is a *weak base* because HCl dissociates so completely in water. Bicarbonate, however, is a relatively *strong base* that can accept a H^+ and form undissociated carbonic acid. The latter reaction removes H^+ from solution and the $[H^+]$ decreases, causing the pH to rise. Bases do not have to be ions; for example, ammonia (NH_3) is a base because it can accept a proton and become ammonium ion (NH_4^+). This reaction is of little importance in the blood, but it is important in the renal collecting duct.

For physiological compounds, acids and bases have *conjugate pairs* in which the weak acid (e.g., H_2CO_3) dissociates into protons and its conjugate strong base (HCO_3^-). Similarly, a strong acid (e.g., HCl) dissociates into protons and its conjugate weak base (Cl^-). In addition, proteins act as buffers by virtue of the terminal carboxyl and amino groups, which can donate (R-COOH $\rightarrow$ R-COO^- + H^+) or accept (R-NH_2 + $H^+ \rightarrow$ R-NH_3^+) protons, respectively.

Buffers Reversibly Bind Hydrogen Ions and Prevent Major Changes in pH

The pH is a measure of H^+ that are *not bound to buffers. Buffers* are substances that *reversibly bind* H^+. In this manner, buffers minimize the pH change of a solution in the face of added acid (by binding the added protons) or base (by releasing protons to bind to the added base). The most important buffer system in the blood is based on the chemical interactions between CO_2 and HCO_3^-. This *buffer system* is variably referred to as the bicarbonate or CO_2/HCO_3^- or HCO_3^-/H_2CO_3 buffer system and is characterized by the following reactions:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

Note that these reactions are reversible. The primary direction of the reactions is dependent on the relative amounts of compounds on either side of the equations.

Physiological buffer systems contain a weak acid and its conjugate base or salt (a salt is an acid in which part of the H^+ is replaced by another cation). For example, carbonic acid (H_2CO_3) dissociates incompletely into hydrogen (H^+) and bicarbonate (HCO_3^-) ions, with some remaining undissociated carbonic acid; its conjugate salt, sodium bicarbonate ($NaHCO_3$), dissociates completely into sodium (Na^+) and bicarbonate (HCO_3^-) ions. Thus in a solution buffered with carbonic acid and sodium bicarbonate, there are sodium, hydrogen, and bicarbonate ions and undissociated carbonic acid. If a strong acid such as hydrochloric acid is added to the solution, the added H^+ upsets the dissociation equilibrium of carbonic acid; H^+ will then combine with bicarbonate ion to form carbonic acid, thus reducing the $[H^+]$ and preventing a major change in pH.

In contrast, if a base such as sodium hydroxide (NaOH) is added to a solution buffered with H_2CO_3 and $NaHCO_3$, the hydroxyl ions (OH^-) created by dissociation of NaOH combine with H^+ to form water. The decrease in H^+ causes dissociation of more carbonic acid and liberation of H^+, again preventing a large change in pH.

The extent to which a weak acid dissociates, and the resultant concentrations of H^+, base, and undissociated acid, is determined by the *dissociation constant* (K_a) and can be described by the *law of mass action.* For carbonic acid:

$$K_a = [H^+][HCO_3^-]/[H_2CO_3]$$

Taking logarithms of both sides of this equation results in:

$$\log K_a = \log([H^+][HCO_3^-]/[H_2CO_3])$$

Using the properties of logarithms yields:

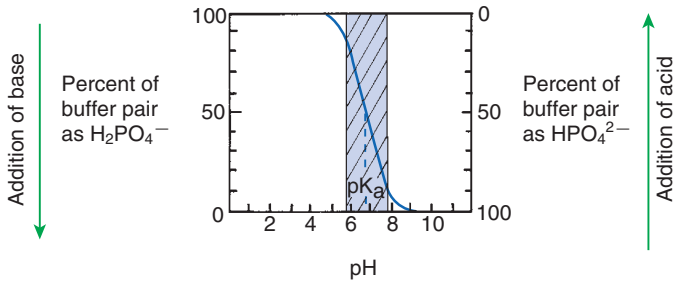
$$\log K_a = \log[H^+] + \log[HCO_3^-]/[H_2CO_3]$$

Rearrangement of this equation yields:

$$-\log[H^+] = -\log K_a + \log[HCO_3^-]/[H_2CO_3]$$

Because $-\log[H^+]$ is pH and $-\log K_a$ is called pK_a, therefore:

$$pH = pK_a + \log[HCO_3^-]/[H_2CO_3]$$



• **Fig. 52.1** Titration curve for the phosphate buffer system. The pK_a is 6.8. The *shaded area* represents the range of pH over which this buffer is effective.

This is the *Henderson-Hasselbalch equation* written for the bicarbonate buffer system. It can be written for any buffering system in the generic form:

$$pH = pK_a + \log[base]/[acid]$$

This equation shows that the pH of a solution is determined by the ratio of the concentration of base (the H^+ acceptor) to that of undissociated acid (the H^+ donor) and by the pK_a of the buffering system. The pK_a for a buffer system is the pH at which there are equal amounts of the weak acid and its conjugate base. Each buffer system has a specific pK_a value. Buffers that have a pK_a close to the normal body pH of 7.4 are the most effective physiological buffers.

Fig. 52.1 is a graphic presentation of the Henderson-Hasselbalch equation showing the change in pH that results when acid is added to a phosphate buffer system with a pK_a of 6.8. Initially, as acid is added, there is a large decrease in pH. As considerably more acid is added to the solution, the pH changes little. Hydrogen ions combine with hydrogen phosphate ions (HPO_4^{2-}) and form dihydrogen phosphate ions ($H_2PO_4^-$). Finally, the pH decreases considerably. The zone over which the pH changes little as acid is added (i.e., where buffering capacity is optimal) is within ±1 pH unit of the pK_a. Note that when the pH equals the pKa, 50% of the buffer has been consumed. From this buffer curve, it is obvious that an effective buffer must have a pKa within ±1 pH unit of the solution in which it operates. Thus the optimal blood buffers should have a pK_a between 6.4 and 8.4. In addition, buffers must be sufficiently plentiful to be effective.

Carbonic Anhydrases Catalyze the Reversible Hydration of Carbon Dioxide

The enzyme *carbonic anhydrase* (CA) is a catalyst for the reversible reaction between carbonic acid and carbon dioxide plus water, as follows:

$$CO_2 + H_2O \overset{CA}{\rightleftharpoons} H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

Although the interconversion of CO_2 and H_2O into H_2CO_3 (and the reverse reaction) will proceed in the absence of carbonic anhydrase, its rate is greatly accelerated by this catalyst. The dissociation/association between H_2CO_3 and the ions H^+ and HCO_3^- is rapid and does not require a catalyst.

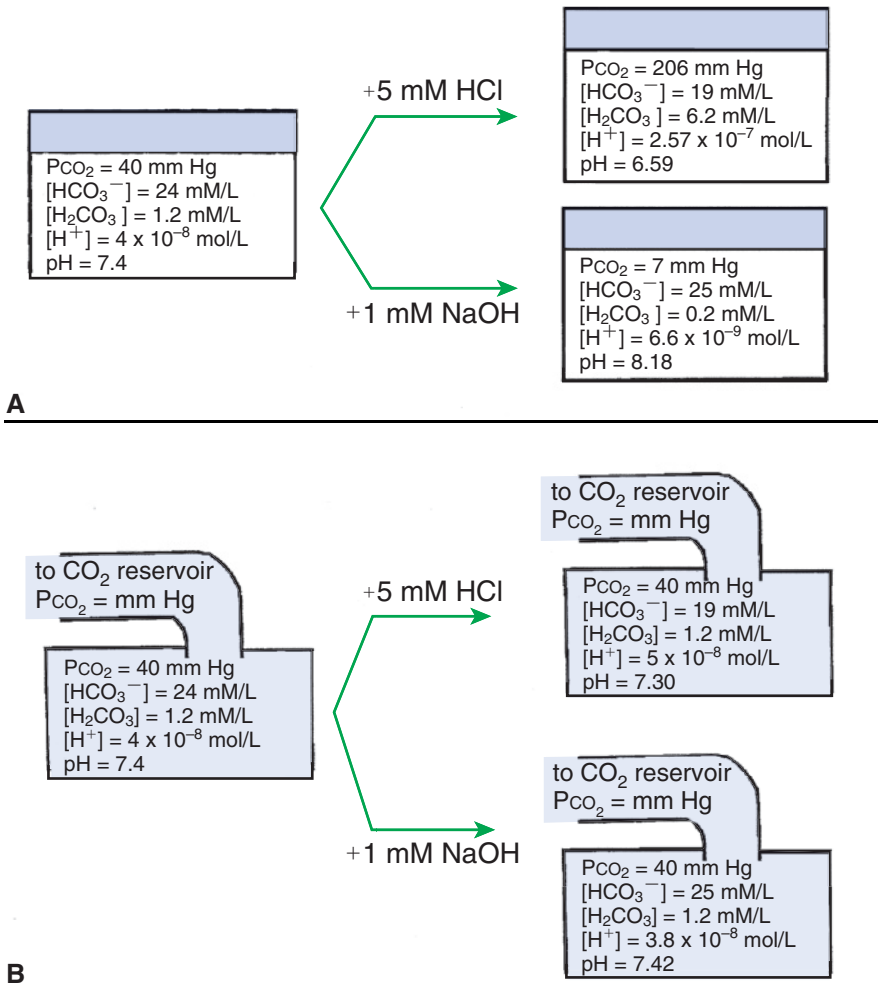
There are a number of different isoforms of carbonic anhydrases present in a variety of tissues. Carbonic anhydrases are present within erythrocytes and in most segments of the kidney, with important actions in the proximal tubules and collecting duct (see Chapter 42). Carbonic anhydrase keeps the hydration of CO_2 moving at physiological speeds; its presence in erythrocytes and renal cells is critically important for rapid response to the dynamic CO_2 levels that result from cellular respiration and gas transport, and thus maintenance of normal body pH.

Bicarbonate and Hemoglobin Are the Most Important Blood Buffers

The most important system for buffering the blood is based on the reversible reactions between HCO_3^-, H_2CO_3, and CO_2. This bicarbonate buffer system has a pK_a of 6.1; as this pK_a is more than 1 unit away from the normal physiological pH of 7.4, it may seem to make this buffer system ill-suited for buffering blood. However, for this system, the greater difference between pK_a and physiological pH is of lesser consequence for two reasons. First, there is a large amount of HCO_3^- in the blood, typically 24 millimole/liter (mM/L), making it readily available for buffering. Second, the kidneys can regulate the concentration of HCO_3^- and the lungs can regulate the concentration of CO_2 (and by extension, H_2CO_3). Because the base and acid concentrations can be regulated, the bicarbonate system is an *open system.*

Fig. 52.2 shows the value of this open system in maintaining body pH. It should be noted that the concentration of H_2CO_3 ($[H_2CO_3]$) in solution is directly proportional to the CO_2 partial pressure (P_{CO_2}); 1 molecule of H_2CO_3 is in equilibrium with 340 molecules of CO_2. Therefore $[H_2CO_3]$ can be calculated as 1/340 × P_{CO_2}, which simplifies to 0.03 × P_{CO_2}. In Fig. 52.2A–B, *top panels,* 5 mM of HCl is added to plasma. Fig. 52.2A shows what happens in a closed system where the total quantity of buffer remains constant. In this closed system, the 5 mM of HCl reacts with 5 mM of HCO_3^- to form 5 mM of H_2CO_3. As a consequence of this reaction, $[HCO_3^-]$ decreases from 24 to 19 mM/L, $[H_2CO_3]$ increases from 1.2 to 6.2 mM/L, and P_{CO_2} increases from 40 to 206 mm Hg. Using the law of mass action, one can calculate that $[H^+]$ increases from 4×10^{-8} mol/L to 2.57×10^{-7} mol/L (= 40–257 nEq/L) or, stated in other terms, pH decreases from 7.4 to 6.59. If the system is open (*top panel* of Fig. 52.2B), CO_2 is released to the environment as fast as it is produced so that P_{CO_2} and therefore $[H_2CO_3]$ remain constant, then $[H^+]$ increases to only 5×10^{-8} mol/L (= 50 nEq/L) and the pH decreases only to 7.3. The *lower panel* in Fig. 52.2B shows similar advantages to the open system when a strong base is added to the buffer system. Under most conditions, the body functions as an open system with regard to the HCO_3^-/H_2CO_3 buffer system so that pH changes are minimized. When tissues are ischemic, however, they have no connection to the lungs so CO_2 cannot be eliminated. The ischemic tissue then functions as a closed system and pH changes within the tissue can therefore be drastic.

The HCO_3^-/H_2CO_3 buffering system is of great value to clinicians because its components can be readily measured in the clinical laboratory and thereby used to diagnose acid-base disturbances. It is not necessary to measure the components of every buffering system to diagnose acid-base disturbances. If one system is known, changes in other systems can be predicted. For clinical estimation of acid-base balance, the Henderson-Hasselbalch equation for the bicarbonate system (pK_a = 6.1) is used as follows; under normal



• **Fig. 52.2** Buffer function of the bicarbonate buffer system under closed and open conditions. (A) Under closed conditions, the total quantity of the buffer (acid plus base components) remains constant. This condition results in greater changes in pH upon addition of acids or bases. (B) Under open conditions, the P_{CO_2} of the system, and thus the concentration of carbonic acid $[H_2CO_3]$, is maintained at a fixed level by continuous equilibration of the liquid phase with a gas reservoir of constant P_{CO_2}. This condition minimizes the effect of addition of acids or bases. The term *$[H_2CO_3]$* denotes the combined concentration of carbonic acid and dissolved CO_2. (From Madias NE, Cohen JJ. Acid-base chemistry and buffering. In: Cohen JJ, Kassirer JP, eds. *Acid-base*. Boston: Little, Brown; 1982.)

conditions, the pH of arterial blood is 7.4, $[HCO_3^-]$ is 24 mM/L, and the arterial CO_2 partial pressure (P_aCO_2) is 40 mm Hg (recall that $[H_2CO_3] = (0.03 \times P_{CO_2})$:

$$7.4 = 6.1 + \log[(24)/(0.03 \times 40)] = 6.1 + \log 20$$

This equation demonstrates that a normal blood pH requires a $[HCO_3^-]$ to $[0.03 \times P_{CO_2}]$ ratio of 20 : 1. An increase or decrease in this ratio increases or decreases pH, respectively.

Hemoglobin is also an important blood buffer because it is plentiful and because the imidazole residues of globin histidine have a pK_a closer to normal blood pH. In actuality, the pK_a of hemoglobin changes with its degree of oxygenation. Because the pK_a of deoxyhemoglobin ($pK_a = 7.93$) is closer to normal blood pH than the pK_a of oxyhemoglobin ($pK_a = 6.68$), deoxyhemoglobin provides more buffering capacity than does oxyhemoglobin. When arterial blood enters the tissue capillaries, oxygen (O_2) leaves hemoglobin so the resulting deoxyhemoglobin is an excellent buffer for the H^+ produced by the bicarbonate-carbonic acid reactions when CO_2 is added to the blood.

The other blood buffer with a physiological pK_a is the $HPO_4^{2-}/H_2PO_4^-$ system, with a pK_a of 6.8 (Fig. 52.1). The normally low phosphate concentration in the blood makes this buffering system quantitatively less important in this location; however, it is important in the renal tubules where phosphate becomes concentrated. Plasma proteins also provide a small amount of blood buffering.

The Blood Buffers Provide the First Defense Against a Change in Blood pH, but the Lungs and Kidneys Must Ultimately Correct the Hydrogen Ion Load

When normal body pH is threatened by a change in the production or elimination of H^+, the blood and tissue buffers provide the first line of defense. However, buffers only mitigate the immediate and large-scale changes in pH; they cannot correct the underlying

problem by increasing or decreasing the elimination of H^+ or by replacing lost buffering capacity. Ultimately, pH must be corrected by adjustments in ventilation or by changes in renal function. Because the lungs can alter P_aCO_2 (which reflects $[H_2CO_3]$) and the kidneys can regulate the concentration of HCO_3^-, the Henderson-Hasselbalch equation can be thought of as follows:

$$pH = pK_a + \log(\text{renal function}/\text{ventilation})$$

Changes in Ventilation Can Rapidly Change Carbon Dioxide Partial Pressure and Therefore Alter pH

As blood flows through the tissues, CO_2 diffuses into the plasma and the erythrocytes, where carbonic acid forms and then dissociates into hydrogen and bicarbonate ions:

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$

Because the initial $[HCO_3^-]$ in the blood is greater than $[H_2CO_3]$, the relative increase in $[H_2CO_3]$ is greater than the increase in $[HCO_3^-]$, and thus the $[HCO_3^-]/[H_2CO_3]$ ratio (i.e., $[HCO_3^-]/[0.03 \times PCO_2]$) decreases; consequently, pH decreases. In other words, CO_2 enters this system as PCO_2 making the denominator of the $[HCO_3^-]/[0.03 \times PCO_2]$ equation increase, and thus the pH falls. In the lungs, CO_2 leaves the blood and the pH increases again. For these reasons, venous blood is more acidic than arterial blood. Normally, the lungs eliminate CO_2 as fast as the tissues produce it so the P_aCO_2 and pH of arterial blood remain relatively constant.

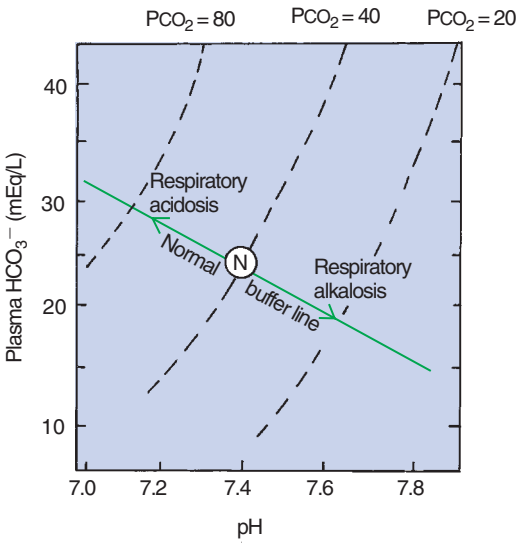
The lungs can cause rapid changes in blood pH by increasing or decreasing the elimination of CO_2. When ventilation increases in relation to CO_2 production (*hyperventilation*), P_aCO_2 decreases, the $[HCO_3^-]/[0.03 \times P_aCO_2]$ ratio increases, and pH increases. Conversely, when ventilation decreases in relation to CO_2 production (*hypoventilation*), P_aCO_2 increases, $[HCO_3^-]/[0.03 \times P_aCO_2]$ decreases, and pH decreases. Fig. 52.3 shows how $[HCO_3^-]$ and pH change as the PCO_2 of the blood increases or decreases.

Metabolic Production of Fixed Acids Requires That the Kidneys Eliminate Hydrogen Ions and Conserve Bicarbonate

Although the reaction of CO_2 with water to release H^+ is a major source of acid, the body produces other acids, known as *fixed acids, metabolic acids,* or *nonvolatile acids,* that cannot be eliminated by respiration. When fixed acids are added to the blood, for example, sulfuric acid from protein metabolism, the H^+ are buffered in part by HCO_3^-. Buffering results in the conversion of HCO_3^- to H_2CO_3 and then to CO_2, which is eliminated by the lungs. Fixed acids are produced continuously and would consume the body's HCO_3^- if the kidneys were not continually regenerating HCO_3^-.

The role of the kidneys in acid-base balance is described in Chapter 44. In brief, large amounts of HCO_3^- are filtered daily through the glomerulus and subsequently reabsorbed in the renal tubule. The amount of HCO_3^- reabsorbed depends on the amount filtered, which is determined by the plasma concentration of HCO_3^-, the *glomerular filtration rate,* and the rate of H^+ secretion by renal tubular cells. This last rate is controlled in part by the acid-base status of the body.

When PCO_2 is high, the reaction:



• **Fig. 52.3** A pH-bicarbonate diagram showing the effect of increasing and decreasing PCO_2 on pH and HCO_3^- concentration. As PCO_2 increases or decreases, the changes in pH and bicarbonate concentration are predicted by the normal buffer line. *N,* Normal arterial blood composition.

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$

within the *renal tubules* is driven to the right, producing more H^+ for secretion into the *tubular lumen* and more HCO_3^- for return to the blood. When PCO_2 is low, H^+ elimination and therefore HCO_3^- reabsorption decrease.

Ammonia, an important buffer in the distal renal tubule, is produced by the action of *glutaminase* on *glutamine.* In acidosis, the activity of glutaminase increases, resulting in increased ammonia production, an increased buffering capacity of the renal tubular fluid, and therefore increased ability to eliminate H^+.

Buffers and Ion Exchangers Regulate Intracellular pH

Whereas bicarbonate and hemoglobin provide the most immediately available source of buffers to prevent drastic changes in blood pH, *intracellular buffers* within the body tissues, other than blood, provide another large reserve of buffering capacity. To enter cells, H^+ must be exchanged with other cations, such as Na^+ or K^+. When inside the cell, H^+ is buffered by *amino acids, peptides, proteins,* and *organic phosphates.* These buffers provide approximately five times the buffering capacity of the extracellular fluid. As a general rule, intracellular pH follows extracellular pH but because of the large intracellular buffering capacity, the pH changes are less dramatic in the cytosol.

Two *ion exchangers* play a major role in regulation of intracellular pH. The Na^+/H^+ exchanger uses the energy derived from the extracellular to intracellular Na^+ gradient to move H^+ out of cells. The Cl^-/HCO_3^- exchangers use a similar mechanism to move HCO_3^- out of cells. When H^+ concentration within the cell increases, the activity of the Na^+/H^+ exchanger increases so that more H^+ is extruded. At the same time, the activity of the Cl^-/HCO_3^- exchanger is *inhibited* so that HCO_3^- accumulates within the cell. These two mechanisms can restore intracellular pH to normal as long as extracellular pH is also normal. However, if extracellular pH is low, it is more difficult for the Na^+/H^+ exchanger to extrude H^+, and thus intracellular pH tends to become acidic.

If the cytosol becomes alkaline, the activity of the Cl^-/HCO_3^- exchanger is facilitated and the Na^+/H^+ exchanger is inhibited to restore pH. This corrective action is inhibited if extracellular pH is also alkaline.

Acid-Base Disturbances

Primary Acid-Base Disturbances Are Typically Accompanied by Compensatory Mechanisms That Move pH Toward Normal

In most diseases, the buffering systems, lungs, and kidneys keep pH within tolerable limits, but in severe disease, these homeostatic mechanisms may be inadequate and life-threatening changes in pH can occur. In the diagnosis and treatment of acid-base abnormalities, it is important to realize that a *primary* abnormality causes the change in blood pH, which is followed by *compensatory* changes that try to restore pH toward normal. Because of the body's attempt to correct the abnormality, the clinician must differentiate the primary cause of the problem from the compensatory changes. The primary problems are excessive accumulation or elimination of CO_2 (respiratory abnormalities) or the excessive accumulation or elimination of fixed acids or buffer bases (metabolic abnormalities).

Base Excess Is a Measure of the Metabolic Component of Acid-Base Status and Its Assessment Facilitates the Evaluation of Mixed Acid-Base Disturbances

The *total buffer base* is the sum of the concentrations of available blood buffers. To facilitate the clinical interpretation of acid-base status, clinicians use the parameters of *base excess* and/or the inverse term *base deficit* to indicate the metabolic component of acid-base disturbances. By convention, total buffer base is typically reported as base excess, with a normal range of +3 to −3 mM/L. Base excess values greater than +3 mM/L indicate too much base relative to the amount of *nonvolatile acids* (acids in the body other than CO_2, such as lactic acid, acetoacetic acid, and β-hydroxybutyric acid), which is consistent with metabolic alkalosis. In contrast, base excess values less than −3 mM/L indicate insufficient base relative to the nonvolatile acid load, which is indicative of a metabolic acidosis. Although the value of HCO_3^- is fundamental to assessing the acid-base balance of a patient, its interpretation can be complicated by the fact that HCO_3^- can be affected by changes in P_{CO_2} due to normal compensatory responses. Thus the base excess is used in conjunction with HCO_3^- to more accurately assess acid-base status. Furthermore, assessment of base excess is particularly valuable in the face of mixed acid-base disorders, where multiple metabolic and respiratory and/or acidotic and alkalotic disturbances coexist. Base excess values remain in the normal range in acid-base disturbance of solely respiratory origin; this is because respiratory acid-base disturbances are due to changes in P_{CO_2} values, rather than the gain or loss of nonvolatile acids or buffer base.

Respiratory Acidosis Is Caused by the Accumulation of Carbon Dioxide, Which Decreases Blood pH

Respiratory acidosis is characterized by an acidic pH and increased P_aCO_2. It is caused by *alveolar hypoventilation,* which can result from damage to or depression of the respiratory control centers, injury to the respiratory pump (e.g., fractured ribs, bloated abdomen), or severe respiratory disease that either obstructs the airways or excessively stiffens the lungs. When alveolar hypoventilation occurs, P_{CO_2} increases because the lungs incompletely eliminate CO_2. The reaction:

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^-$$

is driven to the right by the accumulating CO_2; H^+ accumulates, and pH decreases. Bicarbonate accumulates simultaneously, but the amount is too small to keep the $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio at a normal value of 20 : 1.

In the blood, other nonbicarbonate buffers not only take up H^+ produced by the accumulation of CO_2 but also assist in the accumulation of HCO_3^- as follows:

$$\begin{array}{l} CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^- \\ \qquad\qquad\qquad\qquad\qquad\quad \downarrow \\ \qquad\qquad\qquad\qquad\qquad\quad H^+ + Hb^- \rightarrow HHb \end{array}$$

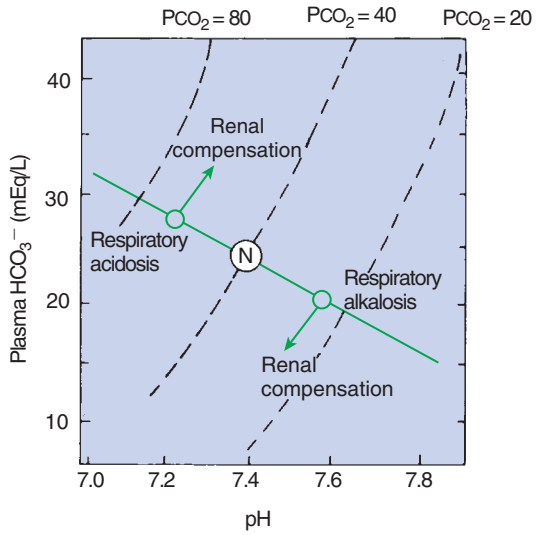
By buffering H^+, hemoglobin (Hb^-) pulls the first reaction to the right and produces HCO_3^-. This accumulation of bicarbonate, shown on the *normal buffer line* in the pH-HCO_3^- diagram (Fig. 52.3), is still insufficient to maintain a normal $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio, and therefore the pH decreases. As a result of these various reactions, the characteristic findings in acute respiratory acidosis are an elevated P_aCO_2, a decreased pH, and a minor compensatory increase in $[HCO_3^-]$. The base excess values are in the normal range as there has been no change in amount of nonvolatile acids or total buffer base.

The ideal way for the clinician to correct respiratory acidosis is to restore alveolar ventilation. However, because disease processes that impede ventilation cause the respiratory acidosis, this option is not available to the untreated animal, and other means to correct pH, primarily renal mechanisms, must be used. The elevated P_{CO_2} and decreased pH lead to increased H^+ and NH_3 production by the kidney. This increases the elimination of H^+ in the urine and generates new HCO_3^-, and as the plasma $[HCO_3^-]$ increases, the $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio and pH are adjusted toward normal. The newly generated HCO_3^- adds to the total buffer base and therefore causes a base excess. Fig. 52.4 shows how this accumulating HCO_3^- adjusts the pH toward normal during respiratory acidosis, even though P_{CO_2} remains constant.

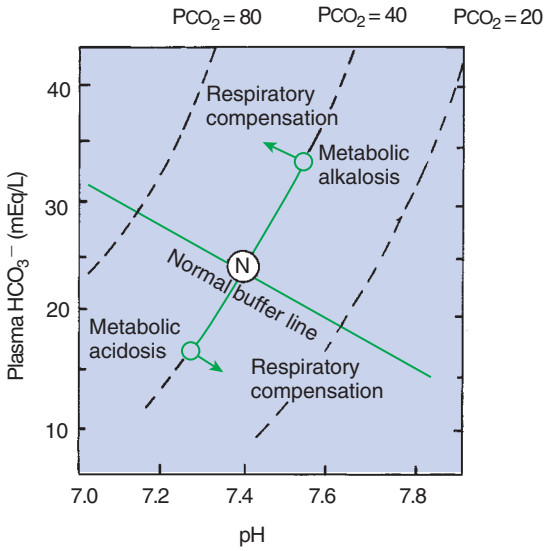
Respiratory Alkalosis Is Caused by the Loss of Carbon Dioxide, Which Increases Blood pH

Respiratory alkalosis is characterized by an alkaline pH and decreased P_aCO_2. It is caused by *alveolar hyperventilation,* which results from stimulation of the chemoreceptors by hypoxemia or from stimulation of intrapulmonary receptors by lung injury or inflammation. Overly vigorous use of a ventilator can also cause hyperventilation in an anesthetized animal. Carbon dioxide is eliminated faster than the tissues produce it and so blood P_{CO_2} decreases. The changes in blood chemistry are the inverse of those in respiratory acidosis as the relative CO_2 deficit drives the reactions of the bicarbonate buffer system to the left:

$$\begin{array}{l} CO_2 + H_2O \leftarrow H_2CO_3 \leftarrow H^+ + HCO_3^- \\ \qquad\qquad\qquad\qquad\qquad\quad \uparrow \\ \qquad\qquad\qquad\qquad\qquad\quad H^+ + Hb^- \leftarrow HHb \end{array}$$



• **Fig. 52.4** A pH-bicarbonate diagram showing the effects of respiratory acidosis and alkalosis on pH, HCO_3^- concentration, and P_{CO_2} of arterial blood. In acute respiratory acidosis, as P_{CO_2} increases, the changes in pH and bicarbonate concentration are predicted by the normal buffer line. Renal compensation leads to an accumulation of HCO_3^-, which increases the pH, whereas P_{CO_2} remains constant (upward green arrow). In respiratory alkalosis, P_aCO_2 and HCO_3^- decrease and pH increases. The kidneys compensate by eliminating HCO_3^-, which decreases pH (downward green arrow). *N,* Normal arterial blood composition.



• **Fig. 52.5** A pH-bicarbonate diagram showing the effects of metabolic acidosis and alkalosis on pH, HCO_3^- concentration, and P_{CO_2} of arterial blood. In uncompensated metabolic acidosis, there is a decrease in the bicarbonate concentration, which leads to a decrease in pH, whereas P_{CO_2} remains constant (lower open circle). Respiratory compensation results in a decrease in P_{CO_2}, with a subsequent increase in pH and movement of data points parallel to the normal buffer line (lower arrow). In metabolic alkalosis, HCO_3^- increases and causes an increase in pH (upper open circle). Respiratory compensation is by alveolar hypoventilation, which leads to an increase in P_aCO_2 and a return of pH toward normal (upper arrow). *N,* Normal arterial blood composition.

As CO_2 is eliminated, H_2CO_3 is formed from H^+ and HCO_3^-, and thus the pH increases and $[HCO_3^-]$ decreases. Hydrogen ion is supplied by release from nonbicarbonate buffers, such as hemoglobin. As a result of these various reactions, the characteristic findings in acute respiratory alkalosis are decreased P_aCO_2, increased pH, and a compensatory $[HCO_3^-]$ decrease. Other buffers then replace the reduced bicarbonate. There is no change in total buffer base so base excess values remain in the normal range. The increase in the $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio increases pH.

Fig. 52.4 shows the increase in pH and decrease in $[HCO_3^-]$ as P_{CO_2} decreases. To adjust the pH toward normal, hyperventilation must be stopped, or the kidneys must eliminate HCO_3^-. The latter occurs because the low P_{CO_2} and alkalosis reduce H^+ and NH_3 production by the kidney. When H^+ is not produced in sufficient amounts to capture all the filtered HCO_3^-, the latter spills into the urine and is lost from the body, moving the pH toward a more normal value.

Metabolic Acidosis Is Caused by the Accumulation of Fixed Acids or the Loss of Buffer Base, Which Decreases Blood pH

Metabolic acidosis is characterized by an acidic pH and decreased HCO_3^-. It is the most common acid-base abnormality. During metabolism, there is a continuous production of fixed acids; an increase in their production or a failure of H^+ elimination by the kidneys is one cause of metabolic acidosis. Increased production of fixed acids results from protein catabolism or ketone production during negative energy balance or from anaerobic metabolism that leads to lactic acidosis. In ruminants, excessive feeding of carbohydrates can lead to increased H^+ production in the rumen (*rumen acidosis*). The H^+ ions that are absorbed cause metabolic acidosis.

Because it is the relative amount of H^+ in relation to buffering capacity that determines pH, a loss of buffer base also causes metabolic acidosis. For example, diarrhea is a common cause of metabolic acidosis because excessive amounts of HCO_3^- are lost in the feces.

In metabolic acidosis, accumulating H^+ in the blood combines with HCO_3^- and other buffers, which causes a base deficit (reported in blood gas results as base excess more negative than −3 mM/L). The increased H^+ also drives the bicarbonate buffer system reactions to the left, and the CO_2, resulting from the combination of H^+ and HCO_3^-, is eliminated by ventilation. The HCO_3^- depletion decreases the $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio; therefore the pH decreases (Fig. 52.5).

$$CO_2 + H_2O \leftarrow H_2CO_3 \leftarrow H^+ + HCO_3^-$$

The characteristic findings in acute metabolic acidosis are decreased HCO_3^-, decreased pH, abnormal base excess, and a compensatory P_aCO_2 decrease, as follows. The decrease in pH accompanying metabolic acidosis stimulates ventilation. This increase in alveolar ventilation has a compensatory effect on metabolic acidosis by eliminating CO_2, leading to a decreased P_{CO_2}, which ultimately adjusts the $[HCO_3^-]/[0.03 \times P_{CO_2}]$ ratio toward normal and pH increases toward normal (Fig. 52.5). Complete restoration of normal acid-base balance requires the restoration of the depleted base by the kidney or by therapy with intravenous fluid that contains buffers, such as bicarbonate or lactate.

Metabolic Alkalosis Is Caused by the Excessive Elimination of Hydrogen Ions or the Accumulation of Buffer Base, Which Increases Blood pH

Metabolic alkalosis is characterized by blood gas parameters of an alkaline pH and increased HCO_3^-. This bicarbonate increase is relative to the $[H^+]$, and as such, metabolic alkalosis is more commonly due to a loss of H^+ than an absolute gain of HCO_3^-. The most common cause of metabolic alkalosis is *vomiting,* during which H^+-rich gastric contents are lost from the body. In ruminants, right-sided displacement or volvulus of the abomasum causes metabolic alkalosis because H^+ ions are trapped in the abomasum. Metabolic alkalosis can also result from K^+ depletion because low concentrations of K^+ in the blood (hypokalemia) result in increased H^+ excretion by the kidney (see later discussion).

The loss of H^+ from the body frees buffer, and thus the plasma $[HCO_3^-]$ and total buffer base increase. The $[HCO_3^-]/[0.03 \times PCO_2]$ ratio, pH, and base excess all increase (Fig. 52.5). The increase in pH reduces the drive to ventilate; alveolar ventilation decreases, and so PCO_2 increases. This adjusts the $[HCO_3^-]/[0.03 \times PCO_2]$ ratio toward normal, and thus the pH decreases toward normal (Fig. 52.5). Thus the characteristic blood gas findings in metabolic alkalosis are increased HCO_3^-, increased pH, and a compensatory P_aCO_2 increase.

Respiratory Compensation for Acid-Base Abnormalities Occurs Rapidly; Renal Compensation Takes Hours to Days

The lungs compensate for disturbances of metabolic origin and the kidneys compensate for disturbances of respiratory origin. Because the *chemoreceptors* respond almost immediately to changes in blood pH, and because changes in ventilation rapidly change PCO_2, respiratory compensation for primary metabolic acid-base disturbances occurs almost immediately. For this reason, it is uncommon to observe primary metabolic acidosis or alkalosis without a respiratory compensation, unless the patient has a concurrent respiratory disorder. The response of the kidneys to a primary respiratory acid-base disturbance is less rapid, and changes in NH_3 and HCO_3^- production occur over hours to several days.

As compensatory mechanisms adjust the pH toward normal, there is less "error signal" to drive the compensation; thus compensatory mechanisms alone rarely return the pH to normal. In metabolic acidosis, for example, the low pH drives ventilation to decrease PCO_2. As the pH returns to normal, however, the respiratory drive to compensate is reduced; therefore restoration of normal pH is rarely complete. Full restoration of normal acid-base balance requires correction of the primary cause of the acidosis or alkalosis.

Appropriate Compensation for Acid-Base Disturbances Can Be Predicted

Many disease processes result in acid-base disturbances and it is not uncommon for critically ill patients to have more than one condition that alters the acid-base balance. For example, a cat that has diabetes and asthma may have both a metabolic acidosis and a respiratory acidosis. A cat in renal failure that is vomiting may have both metabolic acidosis and metabolic alkalosis. When more than one primary acid-base abnormality is present, this is referred to as a *mixed acid-base disturbance.* In *simple* (or *uncomplicated*) *acid-base disturbances,* where a single disorder is underlying the altered blood gases, appropriate compensation can be predicted.

As noted earlier, respiratory compensation for metabolic disorders is rapid as changes in ventilation respond almost immediately to chemoreceptors signals. For metabolic acidosis, the compensatory respiratory response is hyperventilation to decrease P_aCO_2. In uncomplicated metabolic acidosis, the expected P_aCO_2 measured in mm Hg $= 35 - [(22 - HCO_3^-$ mM/L$) \times 0.7]$. If the patient has a P_aCO_2 that has fallen to the expected value based on this formula, the patient is said to be *fully compensated.* If the P_aCO_2 is notably higher than the expected value, the patient is not fully compensating and thus may have additional primary disease preventing a complete response.

For metabolic alkalosis, the compensatory respiratory response is hypoventilation to retain CO_2. In uncomplicated disorders, the expected P_aCO_2 in mm Hg $= 35 + [(22 - HCO_3^-$ mM/L$) \times 0.7]$. In this case, if the P_aCO_2 has risen to the expected level, the patient is fully compensating. However, the compensatory response of hypoventilation has been found to be quite variable as there are many drivers of ventilation, including pain and fear, so a patient may not decrease ventilation to the extent needed to correct its acid-base disturbance. For this reason, P_aCO_2 values within ±3 mm Hg of the expected value based on the formula are considered to be indicative of appropriate compensation.

The metabolic compensatory response to primary respiratory disturbance is biphasic, with modest compensation in acute conditions and greater compensation expected in chronic conditions where the kidney has had time to adjust the retention or secretion of appropriate ions. In respiratory acidosis, the appropriate compensation is a retention of HCO_3^- and thus an increase in arterial blood HCO_3^- values. In acute respiratory acidosis the expected compensatory HCO_3^- in mM/L $= 22 + [(PCO_2 - 35) \times 0.15]$. In chronic respiratory acidosis, the expected compensatory HCO_3^- in mM/L $= 22 + [(PCO_2 - 35) \times 0.35]$, reflecting the larger change in HCO_3^- the kidneys can produce over time. The range for both these compensation formulas is +/− 2 mM/L.

The metabolic compensation for respiratory alkalosis is a decrease in HCO_3^-, based on the following formulae. In acute respiratory alkalosis, the expected HCO_3^- in mM/L $= 22 - [(35 - PCO_2) \times 0.25]$. In chronic respiratory alkalosis, the expected HCO_3^- in mM/L $= 22 - [(35 - PCO_2) \times 0.55]$. The range for both these compensation formulas is +/− 2 mM/L.

It is important to note that appropriate compensation for acid-base disturbances moves the pH toward normal, but typically does not return it to normal.

Potassium, Chloride, and Other Ions Affect Acid-Base Homeostasis

Acid-base disturbances affect the distribution of K^+ within the body. A decrease in blood pH leads to an increase in the plasma K^+ concentration $[K^+]$ (hyperkalemia), whereas a blood alkalosis leads to a decrease in plasma $[K^+]$ (hypokalemia). The hyperkalemia accompanying blood acidosis is a consequence of reduced activity of the Na^+/K^+-ATPase pump and the $Na^+/K^+/Cl^-$ cotransporter, both of which normally transport K^+ into the cell. In addition, intracellular acidosis liberates K^+ from nondiffusible intracellular anions so that more K^+ is free to diffuse out of the cell. In the case of blood alkalosis, high $[HCO_3^-]$ can stimulate K^+ uptake into cells.

Just as changes in blood pH affect $[K^+]$, the converse is also true. *Hypokalemia* is frequently associated with induction of *metabolic alkalosis* through the actions of K^+ in the renal tubule and collecting duct. Specifically, hypokalemia leads to increased H^+ elimination by (1) increasing tubular Na^+/H^+ exchange and basolateral Na^+/HCO_3^- cotransport; (2) increasing NH_3 synthesis and NH_4^+ excretion; and (3) shifting H^+ into cells in exchange for K^+ by H^+/K^+-ATPase in the collecting duct. Similarly, *hyperkalemia* is associated with the induction of *metabolic acidosis* as increased K^+ inhibits NH_3 synthesis and NH_4^+ excretion in the renal tubule and collecting duct.

Chloride depletion can maintain a metabolic alkalosis after the initiating cause of alkalosis (such as vomiting) is resolved because in the absence of Cl^-, the kidney must reabsorb HCO_3^- along with Na^+ to maintain electroneutrality. Conversely, when metabolic acidosis is due to a loss of HCO_3^- (instead of a gain of H^+), the decreased HCO_3^- will be countered by a retention of Cl^-, leading to hyperchloremic metabolic acidosis.

In the blood the total cation concentration (concentration of $Na^+ + K^+ + Mg^{2+} + Ca^{2+}$) should approximately equal the total anion concentration (concentration of $HCO_3^- + Cl^-$). Usually, the total measured cations exceed the total measured anions; the difference is called the *anion gap*. This gap results from the presence of unaccounted for (unmeasured) anions from fixed acids, such as lactate. When metabolic acidosis is due to increased production of fixed acids the anion gap increases.

Arterial Blood Gases Are Used to Diagnose Acid-Base Disturbances

Arterial blood samples are ideal to determine acid-base abnormality, especially if a respiratory component is suspected. Although venous blood samples are comparable to arterial samples for the assessment of pH and HCO_3^-, the venous and arterial values for P_{CO_2} are only comparable when in the normal range (not comparable in hypercapnia). Venous and arterial values of P_{O_2} are not comparable as they differ by up to 60 mm Hg in normal individuals. Blood gas samples must be obtained and handled anaerobically.

It is recommended that arterial blood gas data be analyzed in a step wise manner as follows:

1. Evaluate P_aO_2 to determine whether oxygenation is adequate.
2. Evaluate pH to determine whether there is an acid-base disturbance; is pH normal (7.36–7.44), acidotic (<7.36), or alkalotic (pH >7.44)?
3. If the pH is acidotic, is it due to an increased P_aCO_2 (which would indicate a respiratory origin) or a decreased HCO_3^- (which would indicate a metabolic origin)?
4. If the pH is alkalotic, is it due to a decreased P_aCO_2 (which would indicate a respiratory origin) or an increased HCO_3^- (which would indicate a metabolic origin)?
5. Determine whether there is compensation, and if it is to the expected extent.

Examples are provided in Table 52.1.

TABLE 52.1 Examples of Blood Gas Abnormalities

pH	P_aCO_2	HCO_3^-	Diagnosis
7.4	40	24	Normal
7.26	60	27	Respiratory acidosis; fully compensated if acute, partially compensated if chronic
7.35	60	32	Respiratory acidosis; fully compensated (chronic)
7.2	40	15	Metabolic acidosis; uncompensated
7.35	24	11	Metabolic acidosis; fully compensated
7.47	20	14	Respiratory alkalosis; fully compensated (chronic)
7.55	42	34	Metabolic alkalosis; partially compensated
7.2	50	19	Mixed metabolic and respiratory acidosis
7.6	20	20	Respiratory alkalosis; fully compensated if acute, partially compensated if chronic
7.3	28	9	Metabolic acidosis; partially compensated

P_aCO_2, Arterial CO_2 partial pressure (mm Hg); *HCO_3^-*, bicarbonate (mM/L).

CLINICAL CORRELATIONS

Upper Airway Obstruction in a Boston Terrier

History

A 1-year-old female spayed Boston Terrier exhibits signs of severe respiratory distress. It has difficulty inhaling and makes a snoring sound during inhalation. The effort of walking magnifies the distress. An arterial blood gas sample reveals that P_aCO_2 is 80 mm Hg (normal, 35–45 mm Hg), pH is 7.3, $[HCO_3^-]$ is 39 mM/L (normal, 22–26), and base excess is 10 mM/L (normal, ±2) (Fig. 52.6).

Clinical Examination

Examination reveals excessively narrowed (stenotic) nares and excessive folds of tissue in the soft palate, the latter occluding the glottis. The larynx and trachea appear normal.

Treatment

Reconstructive surgery is performed on the dog to enlarge the nares and remove the excessive tissues from the palate. Two weeks after surgery, respiratory distress is greatly reduced. Arterial blood gas analysis reveals that P_aCO_2 is 45 mm Hg, pH is 7.39, $[HCO_3^-]$ is 27 mM/L, and base excess is 2 mM/L.

Comment

Before surgery, the animal is acidotic with an elevated P_aCO_2 and base excess. Only the high P_aCO_2 explains the acidosis, therefore the dog has respiratory acidosis. The increase in $[HCO_3^-]$ is caused primarily by creation of new HCO_3^- (a base excess) by the kidneys and indicates the condition is of at least several days' duration. Respiratory acidosis was caused by

CLINICAL CORRELATIONS—cont'd

alveolar hypoventilation resulting from the upper airway obstruction. Surgery corrects the obstruction and alleviates the hypoventilation. This returns the pH to a normal value. Two weeks after surgery, the base excess has been virtually eliminated.

Torsion of the Abomasum in a Cow

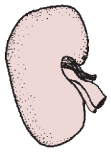
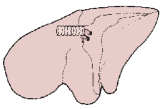
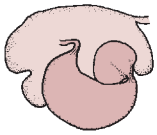
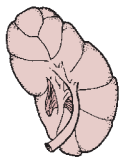
History

A Holstein cow gave birth 2 weeks ago and became inappetent 2 days ago. Over the past 12 hours, she has become lethargic, and her right flank is distended. Examination shows she is dehydrated and her extremities are cold. Rectal examination as well as percussion reveals a large, fluid-filled organ between the rumen and the right abdominal wall. A fluid sample obtained percutaneously from the distended organ is Cl^- rich and very acidic.

An arterial blood gas sample shows that P_aCO_2 is 50 mm Hg (normal, 35–45), pH is 7.6, $[HCO_3^-]$ is 50 mM/L (normal 22–26), and base excess is 24 mM/L (normal, ±2) (Fig. 52.6).

Comment

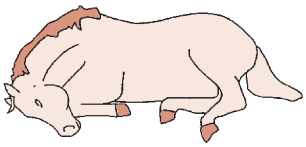
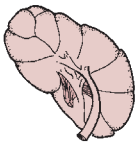
The history and physical findings are typical of a right displacement of the abomasum. This condition occasionally occurs shortly after parturition in dairy cows fed high levels of concentrates and chopped feeds. The abomasum distends and may displace, and thus its inlet and outlet are obstructed. Fluid rich in Cl^- and H^+ continues to be secreted into, and is trapped within, the abomasum. The loss of H^+ from the blood results in a base excess and causes the metabolic alkalosis. The alkalosis depresses ventilation, which elevates P_aCO_2. This is a compensation to restore pH toward normal.

Primary cause	Initial blood chemistry	Compensations	Blood chemistry following compensations
Upper airway obstruction  Upper airway obstructed Too little ventilation CO_2 retained	1) Elevated P_aCO_2 2) Decreased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 3) Decreased pH	 Increased H^+ elimination Increased HCO_3^- retention  Increased drive to ventilate Animal cannot respond because airway is obstructed	1) Increased $[HCO_3^-]$ 2) Base excess 3) Increased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 4) pH approaches normal
Abomasal displacement  H^+ accumulates in the distended abomasum	1) Increased $[HCO_3^-]$ as less H^+ to buffer 2) Base excess 3) Increased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 4) Increased pH	 Decreased H^+ production Decreased HCO_3^- retention Increased HCO_3^- elimination  Decreased ventilatory drive Decreased CO_2 elimination	Restoration of $[HCO_3^-]$ 1) Increased P_aCO_2 2) Decreased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 3) pH approaches normal

A

• **Fig. 52.6** Diagrammatic representation of the acid-base changes initiated by upper airway obstruction, abomasal displacement, and neonatal foal diarrhea.

Continued

Primary cause	Initial blood chemistry	Compensations	Blood chemistry following compensations
Neonatal diarrhea			
	1) Decreased $[HCO_3^-]$ 2) Base deficit 3) Increased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 4) Decreased pH		
Fluid and electrolyte, including HCO_3^- loss in feces		Increased H^+ elimination Complete $[HCO_3^-]$ reabsorption Increased HCO_3^- production	Restoration of $[HCO_3^-]$
			1) Decreased P_{CO_2} 2) Increased $\frac{[HCO_3^-]}{[0.03 \times P_aCO_2]}$ 3) pH approaches normal
B		Increased ventilatory drive Increased CO_2 elimination	

• **Fig. 52.6, cont'd**

CLINICAL CORRELATIONS—cont'd

Treatment

The displacement of the abomasum must be surgically corrected. However, the metabolic alkalosis and any fluid deficits should be treated concurrently to provide the best chance of recovery. The alkalosis is enhanced by loss of Cl^- into the abomasum along with H^+. Repletion of Cl^- allows the kidneys to eliminate the excess HCO_3^- and restores normal pH. In practical terms, this is accomplished by treating the cow intravenously with large volumes of 0.9% NaCl solution.

Neonatal Diarrhea in a Foal

History

A 2-week-old foal has profuse diarrhea. It is lethargic and cold to the touch, its eyes are sunken and dull, and it lies in a pool of feces. The foal's hematocrit is 65% (normal, 30–40), pH is 7.2 (normal, 7.4), P_aCO_2 is 30 mm Hg (normal, 35–45), $[HCO_3^-]$ is 12 mM/L (normal, 22–26), and base deficit is 15 mM/L (normal, ±2) (Fig. 52.6).

Comment

The foal shows typical clinical signs of severe dehydration as a result of excessive fluid loss in the feces. Fluid loss from the intravascular compartment reduces blood volume and cardiac output. To maintain blood pressure, vasoconstriction occurs in the extremities, which therefore have less blood flow and become cold. The loss of fluid from the interstitial space causes the dryness of the eyes and muzzle, the sunken appearance of the eyes, and inelastic skin. The increased hematocrit of 65% confirms the dehydration.

Feces contain HCO_3^- and its excessive loss causes a base deficit and a decrease in pH. In addition, poor tissue perfusion results in lactic acidosis. The acidosis results from loss of buffer base and accumulation of lactic acid; it is a metabolic acidosis. The acidosis stimulates ventilation, which reduces P_aCO_2 in an attempt to correct pH. The foal has a metabolic acidosis that is partially corrected by the reduction in P_aCO_2.

Treatment

This foal needs fluid replacement to increase plasma volume, raise cardiac output, and restore circulatory perfusion. The fluid should contain electrolytes, to replace those lost in diarrhea, and a source of buffer, such as lactate or bicarbonate. A good choice would be lactated Ringer's solution supplemented with additional bicarbonate. If the foal's respiratory and acid-base homeostasis can be maintained until the diarrhea ceases, the foal has a moderate chance of recovery. A serious concern is that this 2-week-old foal could have an infectious cause of diarrhea. The foal could become septic from the infection or the foal could have an initial septicemia, which is now manifested as diarrhea. In either situation, antibiotics are also warranted in most cases to treat the infection.

Practice Questions

1. Elevated P_aCO_2, low pH, and no base excess or deficit are characteristic of:
- **a.** Acute respiratory acidosis
- **b.** Acute respiratory alkalosis
- **c.** Metabolic acidosis
- **d.** Metabolic alkalosis
- **e.** None of the above

2. Elevated P_aCO_2, alkaline pH, and base excess are characteristic of:
- **a.** Chronic respiratory acidosis
- **b.** Chronic respiratory alkalosis
- **c.** Metabolic acidosis
- **d.** Metabolic alkalosis
- **e.** None of the above

3. Low P_aCO_2, acidic pH, and base deficit are characteristic of:
- **a.** Chronic respiratory acidosis
- **b.** Acute respiratory alkalosis
- **c.** Metabolic acidosis
- **d.** Metabolic alkalosis
- **e.** None of the above

4. The most likely acid-base disturbance to be found in a dog at the top of Mount McKinley (Denali) in Alaska (height, 20,320 ft [6353 m]) is:
- **a.** Respiratory acidosis
- **b.** Respiratory alkalosis
- **c.** Metabolic acidosis
- **d.** Metabolic alkalosis
- **e.** None of the above

5. The distal tubule of the kidney affects acid-base balance by:
- **a.** Altering the pK_a of the HCO_3^-/H_2CO_3 buffer
- **b.** Concentrating CO_2 in the renal tubular cell
- **c.** Generating new HCO_3^-
- **d.** Using ammonia to buffer H^+
- **e.** Both c and d

6. Which of the following buffers will be most effective in blood (pH = 7.4)?
- **a.** HX/X^-, $pK_a = 4.2$, plentiful
- **b.** HY/Y^-, $pK_a = 7.2$, scarce
- **c.** HZ/Z^-, $pK_a = 9.6$, scarce
- **d.** HW/W^-, $pK_a = 7.6$, plentiful
- **e.** HA/A^-, $pK_a = 10.2$, plentiful

Bibliography

Boron WF. Acid-base physiology. In: Boron WF, Boulpaep EL, eds. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Cohen JP, Kassirer JJ, eds. *Acid-Base*. Boston: Little, Brown; 1982.

DiBartola SP. Introduction to acid-base disorders. In: DiBartola SP, eds. *Fluid, Electrolyte, and Acid-Base Disorders in Small Animal Practice*. 4th ed. St Louis: Saunders; 2012.

Gamble JL. *Acid-Base Physiology: A Direct Approach*. Baltimore: Johns Hopkins University Press; 2000.

Hall JE. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Saunders; 2016.

Hamm LL, Nakhoul N, Hering-Smith KS. Acid-base homeostasis. *Clin J Am Soc Nephrol*. 2015;10:2232–2242.

Hlastala MP, Berger AJ. *Physiology of Respiration*. 2nd ed. New York: Oxford University Press; 2001.

Purkerson JM, Schwartz GJ. The role of carbonic anhydrases in renal physiology. *Kidney Int*. 2007;71:103–115.

West JB, Luks AM. *West's Respiratory Physiology: The Essentials*. 10th ed. Philadelphia: Wolters Kluwer; 2016.

53 Thermoregulation

SUSAN L. EWART

KEY POINTS

1. Temperature is a major factor affecting tissue function.
2. Ectotherms and endotherms use different strategies to regulate body temperature.
3. Body temperature depends on the balance between heat input and heat output.

Heat Production

1. Heat is a byproduct of all metabolic processes.
2. Shivering produces heat by muscle contraction.
3. Nonshivering thermogenesis is an increase in basal metabolic rate caused by the oxidation of brown adipose tissue to produce heat.

Heat Transfer in the Body

1. Heat is transferred in the blood.
2. Countercurrent heat exchange mechanisms are used to both conserve and lose heat.

Heat Exchange With the Environment

1. Heat exchange by convection occurs as the body warms air or water on its surface.
2. Heat exchange by conduction occurs when the body is in contact with a warmer or cooler surface.
3. Heat exchange by radiation occurs as heat is transferred between objects by infrared radiation.
4. Heat loss by evaporation occurs when the water in sweat, saliva, and respiratory secretions is converted into water vapor.

Temperature Regulation

1. Mammals and birds regulate the input and output of heat to maintain body temperature within a narrow limit.
2. Peripheral and central thermoreceptors sense environmental and core temperatures, respectively.
3. Information from peripheral and central thermoreceptors is integrated in the hypothalamus to regulate heat loss and conservation.

Integrated Responses

1. The responses to heat stress include peripheral vasodilation, increased evaporative cooling, and behavioral modifications.
2. The responses to cold stress include peripheral vasoconstriction, piloerection, increased metabolic heat production by shivering and nonshivering thermogenesis, and behavioral modifications.
3. Fever is an elevation of body temperature that results from an increase in the thermoregulatory set point.

Disorders of Heat and Cold Stress

1. Heat stroke occurs when heat production or input exceeds heat output, so body temperature rises to dangerous levels.
2. Thermoregulatory mechanisms of heat dissipation can cause hypovolemia and hypotension resulting in heat exhaustion.
3. Hypothermia occurs when heat output exceeds heat production, so body temperature decreases to dangerous levels.
4. Frostbite occurs when ice crystals form in the tissues of the extremities.

Temperature Is a Major Factor Affecting Tissue Function

Normal body function depends on chemical and physical processes that are sensitive to changes in temperature, so animals use a variety of strategies to regulate the temperature of their tissues. Each species has a fairly tight range for normal body temperature that it strives to maintain (Table 53.1). If *body temperature* is allowed to decrease too much below its normal value, metabolic processes are slowed and body functions become impaired. Below about 34° C (93.2° F), animals become unable to regulate their own temperature. At the other extreme, an increase in temperature above about 42° C (107° F) can result in heat stroke and cause fatal brain lesions.

Ectotherms and Endotherms Use Different Strategies to Regulate Body Temperature

Fish, reptiles, and amphibians are called *ectotherms* because they depend on external sources to maintain body temperature. The body temperature range of ectotherms can vary with the temperature of the environment; animals that experience a considerable range of body temperatures are termed *poikilotherms.* Ectotherms use behavioral methods to prevent major changes in their temperature. For example, the lizard basks on a sunbaked rock to increase its temperature early in the morning and hides beneath the rock later in the day to prevent overheating. Veterinarians are sometimes asked to advise on the management of captive ectotherms; it is important to remind owners to provide supplemental heat if they

want their ectothermic animal to be active at the cooler times of the year.

Mammals and birds are classified as *endotherms* because they rely more on internal sources for heat and as *homeotherms* because they maintain a fairly constant body temperature in the presence of considerable changes in environmental temperature. Although the maintenance of a constant temperature allows mammals to live in a wide variety of environments and to remain active during the cold times of the year, it is not without cost. Endotherms must maintain a high metabolic rate just to provide the heat necessary to maintain body temperature. This requires a high energy intake and therefore considerable time foraging for food. Ectotherms require much less energy and are better able to survive times of food shortage. Because most veterinarians are primarily concerned with mammals and birds, this chapter focuses on the maintenance of a normal body temperature by endotherms.

TABLE 53.1 Rectal Temperature (°C) of Domestic Mammals

Species	Average	Range
Cat	38.6	38.1–39.2
Cattle (beef)	38.3	36.7–39.1
Cattle (dairy)	38.6	38.0–39.3
Dog	38.9	37.9–39.9
Donkey	37.4	36.4–38.4
Goat	39.1	38.5–39.7
Horse	37.7	37.2–38.2
Pig	39.2	38.7–39.8
Sheep	39.1	38.5–39.9

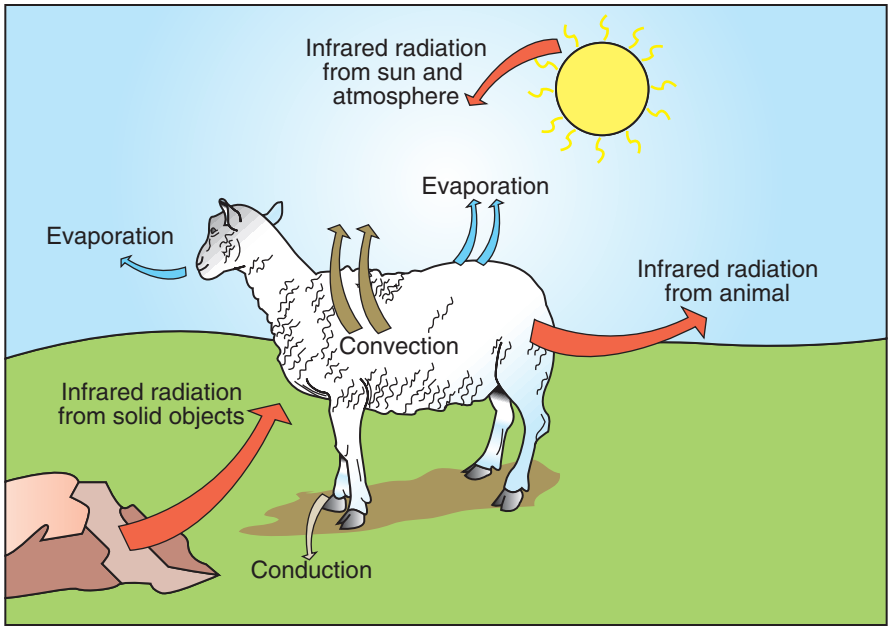
Body Temperature Depends on the Balance Between Heat Input and Heat Output

Heat inputs to the body are from metabolism and from external sources (Fig. 53.1). When food energy is ingested, heat is produced at all stages of the metabolic process. Eventually, all food energy is converted into heat, which is dissipated into the environment. Heat production by the body is related to metabolic rate. A basal metabolic rate (BMR) is necessary to maintain the function of cells. During exercise, metabolic heat production can increase more than 10-fold. If this heat is not dissipated to the environment, body temperature can increase to dangerous levels. Furthermore, this increase in body temperature increases metabolic rate, which further increases heat production.

Animals gain heat from the environment by *convection* (bulk movement of molecules in gas or liquid) when ambient temperature exceeds body temperature, by *conduction* when in contact with a warmer object, and when they are exposed to *radiant heat* sources. The latter occurs when an animal is exposed to sunlight or near solid objects that are warmer than its body temperature.

Heat is lost to the environment by *radiation* from the body surface to a cooler object; by *convection* as the surrounding air or water is warmed by the body; by *evaporation* of respiratory secretions, sweat, or saliva; and by *conduction* to cooler surfaces with which the animal is in contact. A small amount of heat is also lost with urine and feces.

Many of the metabolic heat sources, such as the liver, heart, and limb muscles, are remote from the skin, which is the site of heat loss. Therefore it is necessary to transfer heat among these sites. Body tissues are poor conductors, so heat is transferred mainly by convection in the circulation.



• **Fig. 53.1** Representation of the heat inputs and heat outputs between a mammal and the environment.

Heat Production

Heat Is a Byproduct of All Metabolic Processes

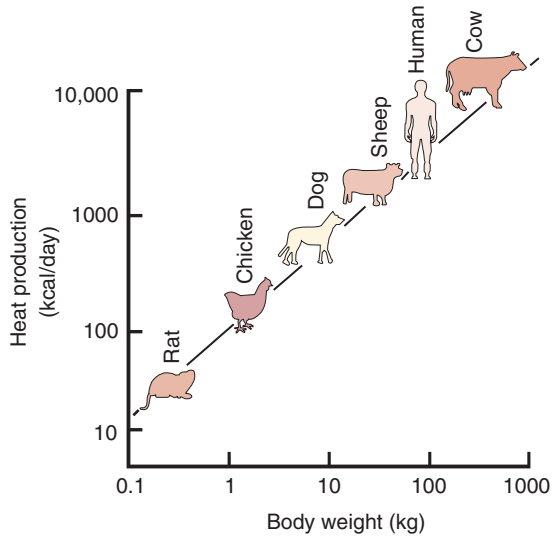
Table 53.2 shows the amount of heat produced by the metabolism of carbohydrates, fats, and proteins. The *basal metabolic rate* (BMR) is the rate of energy metabolism measured under minimal stress while the animal is fasting. BMR is greater in endotherms than in ectotherms because endotherms need to generate heat to maintain body temperature. The BMR per kilogram of body weight is greater in smaller than in larger animals. This is necessitated partly by the greater surface area to volume ratio of smaller animals, which provides a larger area for heat loss. Fig. 53.2 illustrates the relationship between body weight and daily heat production.

Shivering Produces Heat by Muscle Contraction

Shivering is one method of increasing the metabolic production of heat. Antagonistic groups of skeletal muscles (flexors and extensors) are activated so that they produce no useful work. The chemical energy used in shivering is transferred to the body core as heat. If necessary, shivering can go on for many hours and can double heat production. Short bursts of shivering can increase heat production up to fourfold.

TABLE 53.2 Amount of Heat Produced by Metabolism of Major Food Types

Food Type	HEAT PRODUCTION (KCAL/G) Per Gram of Food	Per Liter of O_2 Consumed	Per Liter of CO_2 Produced
Carbohydrates	4.1	5.05	5.05
Fats	9.6	4.75	6.67
Proteins (to urea)	4.2	4.46	5.57



• **Fig. 53.2** Relationship between body weight and heat production.

Nonshivering Thermogenesis Is an Increase in Basal Metabolic Rate Caused by the Oxidation of Brown Adipose Tissues to Produce Heat

When animals are chronically exposed to cold, they develop the ability to increase metabolic heat production without shivering, termed *nonshivering thermogenesis.* This increase in metabolism is mediated through an increase in secretion of thyroid hormones and the calorigenic effects of catecholamines on lipids. Table 53.2 shows that fat metabolism is an effective way to produce heat. *Brown adipose tissue* (*brown fat*) is a specialized vascular, mitochondria-rich fat used to generate heat by uncoupling the electron transport chain within its mitochondria. This uncoupling mechanism is mediated by the protein thermogenin and results in the generation of heat instead of ATP when protons moving down the electron transport chain cross the mitochondrial membrane. Brown adipose tissue is widely distributed within the body of all mammals and is especially prevalent in small mammals and neonates where it forms both subcutaneous and trunk deposits. Cold stress-induced neural release of norepinephrine primarily from the adrenergic nerve supply activates β_3-adrenergic receptors on brown adipocytes; this increases fat metabolism to produce heat distributed around the body through the bloodstream. Furthermore, under cold stress, anatomical sites that classically contain primarily white adipose tissue may undergo transformation to "beige" or "brite" adipose tissue that has characteristics of both brown (uncoupled mitochondria) and white (lipid) adipocytes and contributes to heat generation.

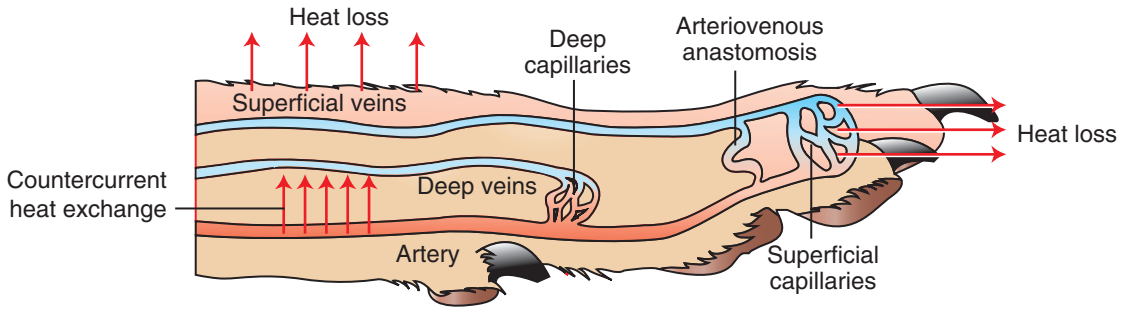
Heat Transfer in the Body

Heat Is Transferred in the Blood

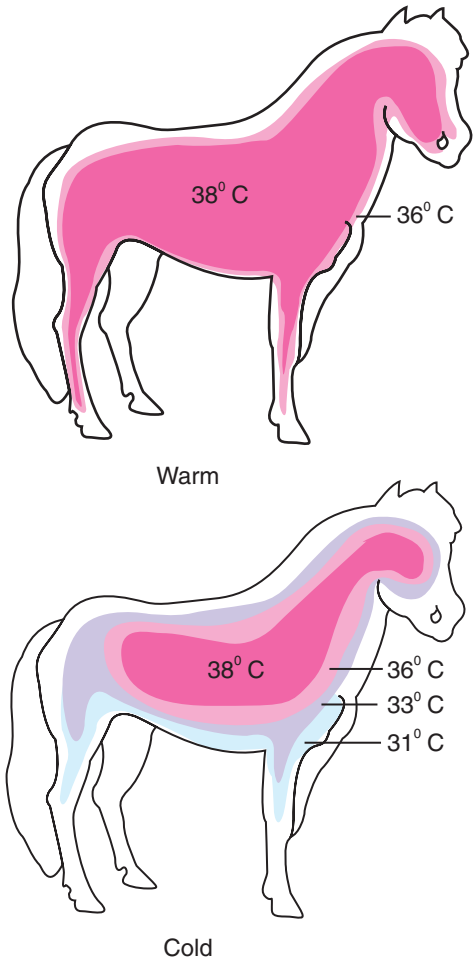
Because heat is produced by metabolically active tissues, most heat is generated by muscles of the limbs and in the liver. Heat is eliminated through the skin and the respiratory tract; therefore it is necessary to transfer heat around the body. Tissues have poor *thermal conductivity;* thus conduction is not an efficient means of heat *redistribution.*

The blood perfusing a metabolically active organ collects heat and transfers it to cooler parts of the body by *circulatory convection.* Redistribution of blood flow can deliver heat preferentially to certain body regions, or it can allow regions to cool, such as when the maintenance of the temperature of the brain and major viscera (*core temperature*) is threatened.

Under conditions of heat stress, circulatory transfer of heat to the skin can be increased dramatically by two mechanisms (Fig. 53.3). First, the *arterioles* of skin vascular beds dilate, which results in increased capillary blood flow. Second, *arteriovenous anastomoses* open in the limbs, ears, and muzzle, which channels blood flow to more superficial veins. These two actions greatly increase the total blood flow to the periphery, and the increased heat delivery increases the temperature of the skin, which facilitates heat loss. Conversely, under cold stress, skin vascular beds vasoconstrict and arteriovenous anastomoses close; thus skin and limb temperatures decrease. This results in reduced heat loss from the skin and in a gradient of temperatures along the limbs (Fig. 53.4). Under severe cold stress, the skin temperature of the extremities can approach ambient temperature. The lipids in the limb extremities have a lower melting point than those in the core, so fats in these locations do not readily solidify in extreme cold stress.



• **Fig. 53.3** Representation of a limb showing the arterial supply and venous drainage by deep and superficial veins. Under warm conditions, blood perfuses the more superficial capillary beds and heat is lost to the environment through the skin. The opening of arteriovenous anastomoses facilitates the distribution of blood to superficial vessels. Blood returns from these superficial vascular beds through the superficial veins and this provides an additional source of heat loss. Under cold conditions, peripheral vasoconstriction occurs, arteriovenous anastomoses close, and the blood flow to the limb is directed to the deeper vascular beds and returns to the trunk through the deep veins; countercurrent heat exchange between the arteries and deep veins conserves body heat.



• **Fig. 53.4** Representation of the distribution of temperatures in a pony under warm and cold environmental conditions. Under warm conditions, the core body temperature extends into the limbs and close to the skin surface of the animal. Under cold conditions, vasoconstriction in the peripheral blood vessels results in a gradient of temperatures between the core and the extremities. The core temperature is maintained only in the abdomen, thorax, and brain of the animal. The more peripheral tissues are allowed to cool considerably.

Countercurrent Heat Exchange Mechanisms Are Used to Both Conserve and Lose Heat

When the environmental temperature is high, the blood perfusing the skin vascular beds returns to the body core through superficial veins from which heat is lost to the skin and air. Under cold conditions, limb blood flow returns to the core through deep veins that accompany arteries (see Fig. 53.3). In this manner, heat is transferred by *countercurrent exchange* from the warm arterial blood to the cooler venous blood and thereby returned to the core of the body.

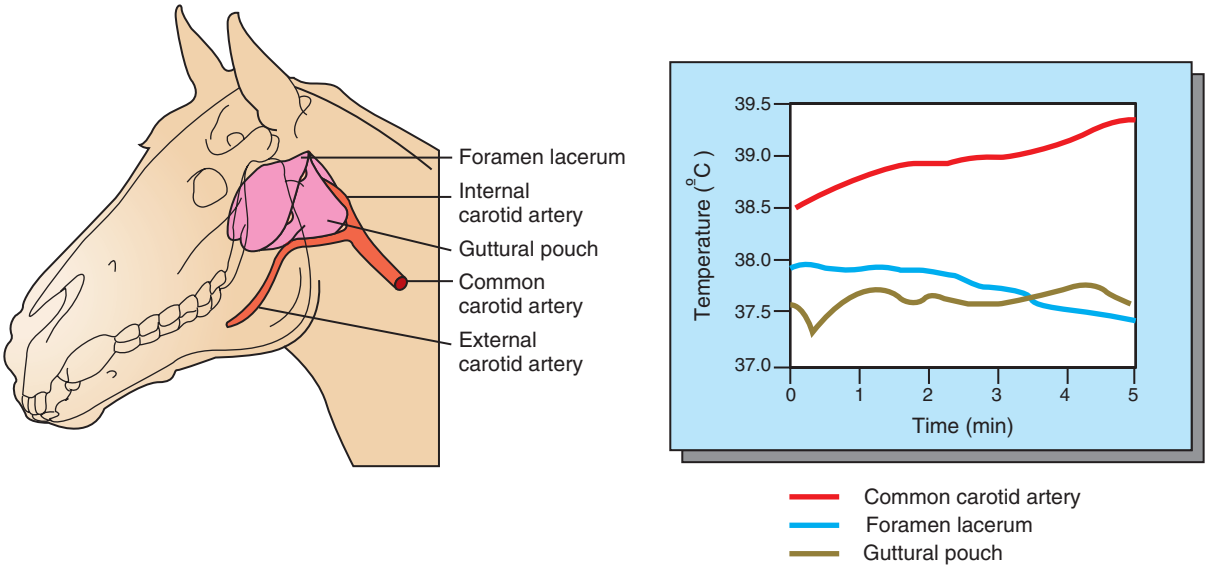
A similar countercurrent exchange of heat occurs in the *carotid rete* in sheep and some other ungulates. In this system, the carotid artery forms a rete (network of anastomosing vessels) bathed in a sinus of venous blood that has drained from the nasal cavity. The colder venous blood from the nose cools the arterial blood supplying the brain and protects the brain from elevated temperatures. This mechanism becomes important during exercise, when the increase in ventilation aids in cooling the blood that drains from the nose. As a result, the arterial blood carrying heat from the exercising muscles is cooled before it enters the brain.

Many animals, including humans and horses, do not possess a carotid rete and must rely on other thermoregulatory mechanisms to cool their brains during exercise. In horses the *guttural pouches* may serve that purpose. The guttural pouches are air-filled structures that surround the internal carotid arteries. This anatomical arrangement allows heat to transfer from the blood in the internal carotid arteries to the air in the guttural pouches, thus protecting the brain from hyperthermia (Fig. 53.5). In addition, the intracranial cavernous venous sinuses may assist in cooling the horse's brain during exercise. This mechanism is thought to function in the same manner as the carotid rete, but less efficiently.

Heat Exchange With the Environment

Heat Exchange by Convection Occurs as the Body Warms Air or Water on Its Surface

When fluids such as air or water in contact with the skin are heated by the body, they flow away, thereby exposing the skin to cooler



• **Fig. 53.5** Guttural pouches cool the blood passing through the internal carotid artery on its way to the brain. *Left,* Anatomical arrangement of the guttural pouches and carotid arteries in the skull. *Right,* Graph of blood and guttural pouch temperatures during a period of cantering. Note that even though the temperature of blood entering the guttural pouch in the common carotid artery increases with the duration of exercise, temperature at the foramen lacerum, where the internal carotid artery enters the cranium, decreases slightly.

fluids; this is cooling by *convection.* Because it takes more heat to warm water than it does to warm an equivalent mass of air, water-dwelling animals lose more heat in cooling by convection than do terrestrial mammals. The amount of heat lost by convection depends on the *thermal gradient* (temperature difference) between the skin of the animal and the fluid overlying the skin; a greater thermal gradient results in more heat loss. In *natural convection* the warmed air or water rises from the surface of the animal because it is less dense than the cooler fluid. In *forced convection,* cooler fluid is moved over the skin surface by a breeze or current or simply because the limbs and animal are moving. Forced convection is more effective than natural convection as a mode of heat loss because the thermal gradient is maintained by the constant renewal of the cooler air or water that blankets the surface of the skin. Young or small animals left in a cool draft can quickly lose much body heat by convection and should be protected from such situations.

The thermal gradient for heat loss can be altered by changes in skin blood flow and the amount of *insulation* separating the animal's core from the environment. Increasing blood flow to the skin raises skin temperature and therefore heat loss, whereas a reduction in skin blood flow reduces heat loss. Fur and hair trap air and impair convection. The thickness of the layer of hair can be altered by *piloerection* (making the hair stand up) and by growing a thicker hair coat in preparation for winter. The thick layer of blubber in sea mammals also provides a layer of insulation. Reducing the area of the exposed body surface also reduces convective heat loss. Animals accomplish this by curling up in a ball or by huddling with littermates. Conversely, animals can increase convective heat loss by assuming a splayed posture.

Convection is a major form of thermoregulation. Although heat can be both dissipated and gained by convection, convective heat gain is not common.

Heat Exchange by Conduction Occurs When the Body Is in Contact With a Warmer or Cooler Surface

When an animal is in direct contact with an object that has good conductive properties, heat can be readily lost or gained. For example, metal is a good heat conductor so heat is transferred rapidly and efficiently between the body and the metal in contact. A cold, stainless steel surgery table can form a heat sink, especially for a small, anesthetized animal. Insulation or a heat source should be provided for such animals. Similarly, newborn piglets can lose much heat by lying on a cold concrete floor. Adult pigs cool themselves by conduction when they wallow in cool mud puddles. Alternatively, heat can be gained by conduction when an animal is in contact with a warmer object, such as a hot water bottle. Thus heat can be both gained and dissipated by conduction.

Heat Exchange by Radiation Occurs as Heat Is Transferred Between Objects by Infrared Radiation

All solid objects emit invisible *electromagnetic radiation* in the *infrared* range. Warm objects emit a shorter wavelength and emit more emissions per unit time than do cool objects. When these emissions strike another object, some are absorbed and thus transfer heat. This radiant heat exchange occurs without contact between objects and can result in heat dissipation or gain. Although all objects at a temperature above absolute zero emit radiant heat, the net heat transfer is from warm to cool objects. Radiant heat loss can occur even when a thermally neutral or warm environment surrounds the animal. For example, heat can be lost from an animal to the uninsulated walls of a building even though the intervening air is warm.

Heat Loss by Evaporation Occurs When the Water in Sweat, Saliva, and Respiratory Secretions Is Converted Into Water Vapor

A calorie is a unit of heat energy; 1 calorie is the energy required to raise the temperature of 1 gram of water 1° C. The *evaporation* of 1 L of water into water vapor requires 580 kilocalories (kcal). If the body provides this heat, evaporation can be a major form of heat loss. Some *evaporative heat loss* occurs continuously by the diffusion of water through the skin and by loss of water vapor from the respiratory tract. This water loss is *obligatory,* but under thermal stress, evaporative cooling can be increased greatly either because *sweat glands* are activated, the animal begins to *pant,* or the animal smears itself with *saliva.*

Evaporation is the one form of thermoregulation limited to heat loss; heat cannot be gained by evaporation. Evaporative heat loss becomes increasingly important as the ambient temperature approaches body temperature; it is the only form of heat loss available when ambient temperature exceeds body temperature. The effectiveness of evaporative heat loss decreases as the ambient humidity increases, such that evaporative heat loss is not functional when the air is saturated with water vapor.

All placental mammals except rodents and lagomorphs have sweat glands, but in dogs and pigs these glands are poorly developed and of little use in thermoregulation. Sweating occurs from two types of coiled, tubular sweat glands located in the dermis. *Apocrine glands* produce a protein-containing secretion, whereas *eccrine glands* produce an aqueous secretion. Apocrine glands produce the thermoregulatory sweat of hoofed animals, whereas in primates, sweat is produced by the eccrine glands. Compared with plasma, secreted sweat has similar ions but is hypotonic; horses are the exception as their sweat is hypertonic to plasma. As sweat passes to the skin surface along the duct, its composition is altered by the reabsorption of ions. If secretion rates are low, almost all the sodium, chloride, and water are absorbed. Therefore the sweat reaching the skin is a concentrated solution of urea, lactic acid, potassium ions, and, in the case of hoofed mammals, protein. When secretion rates are high, less sodium and chloride are absorbed, more water is lost, and the other constituents are consequently diluted. In hot environments, acclimatization increases the sweating rates, and because of increased secretion of *aldosterone,* most of the sodium and chloride is reabsorbed before the sweat reaches the skin. In most species, sweating is under the control of *sympathetic cholinergic nerve fibers* (sweat glands are one of the few locations were sympathetic neurons utilize acetylcholine instead of norepinephrine), but in horses, control is through β_2 adrenoceptors activated by catecholamines originating from sympathetic nerves or the adrenal medulla.

Panting is one mode of increasing evaporation from the respiratory tract. Small tidal volumes are moved at a rapid frequency (200 breaths/min) over the respiratory dead space. The rate of panting is close to the resonant frequency of the respiratory system, and thus the work of breathing is minimized and does not substantially add to the heat load in normal animals. By ventilating primarily dead space, severe hyperventilation and respiratory alkalosis are avoided. In the panting animal, two mechanisms further elevate heat loss through evaporation. First, vascular engorgement of the respiratory and oral mucosa delivers maximal amount of heat to the surface of the respiratory tract and thus increases heat loss through evaporation. Additionally, increased salivation elevates evaporative heat loss as there is more fluid to be evaporated.

Animals with any form of airway obstruction, including brachycephalic breeds of dogs, are particularly susceptible to heat stress as their work of breathing is increased in the absence of thermal stress, thus their heat production is elevated at rest. Because of their respiratory compromise, these animals may not be able to effectively utilize panting to dissipate heat.

In birds, *gular flutter* (vibration of the upper throat) is another method of increasing airflow over the respiratory dead space. Even in mammals that do not pant, such as horses, evaporative heat loss from the respiratory tract probably increases during prolonged exercise because dead space ventilation increases.

Mammals vary in the relative importance of different modes of evaporative heat loss. In horses and cattle, sweating is the major form of evaporative heat loss. Sheep sweat, but panting is also of considerable importance. Dogs rely almost totally on panting. Even small rodents, which neither pant nor sweat, increase evaporative heat loss by smearing saliva or water on their fur.

Temperature Regulation

Mammals and Birds Regulate the Input and Output of Heat to Maintain Body Temperature Within a Narrow Limit

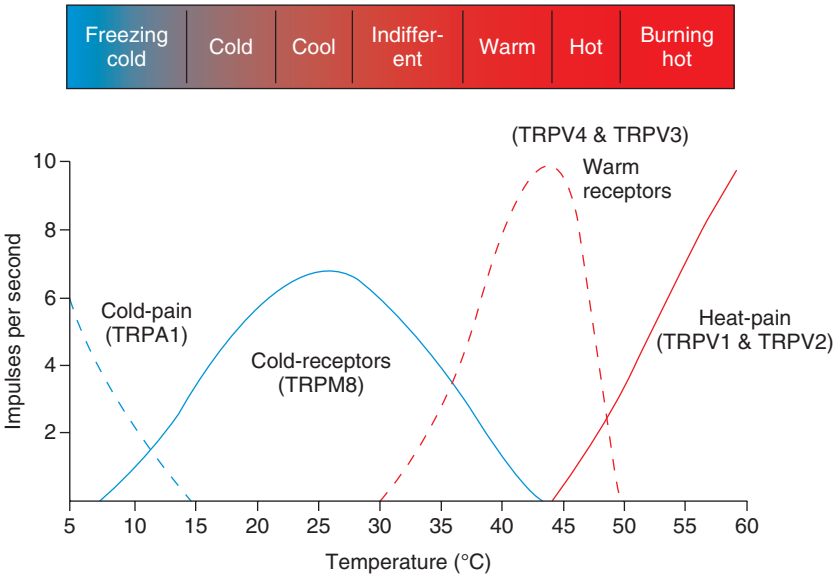
It is customary to measure body temperature as a part of the clinical examination of mammals. This is because body temperature is maintained within fairly narrow limits despite large variations in ambient conditions. In diseased animals, the ability to regulate temperature can be impaired, for example, by dehydration. In addition, pathogens and other agents produce *pyrogens* that can cause an increase in body temperature. The *rectal temperature* (see Table 53.1) is somewhat lower than the *core temperature* of the animal, and changes in rectal temperature lag behind changes in core temperature. However, rectal temperature is a convenient measure in domestic mammals and provides a useful indication of core temperature.

In well-hydrated homeothermic animals living in temperate climates, the range of normal body temperature is quite narrow. There are some physiologically interesting exceptions to this, however. For example, camels in desert conditions of high heat and water restriction experience several degrees increase in body temperature during the day and comparable decrease in body temperature at night. This circadian rhythm of heat gain and loss is a physiological tradeoff between heat loss and water loss; by reducing evaporative heat loss during the day, the animal is protected from dehydration with the adaptive consequence of increased body temperature.

To maintain temperature within narrow limits, the animal must regulate its heat inputs and outputs. The inputs and outputs clearly cannot be equal at all times. During exercise, for example, heat production exceeds heat loss. Heat is stored in the body and then dissipates when exercise ceases. The *specific heat capacity* (ability of mass to absorb heat without changing temperature) of body tissues is similar to that of water, which is relatively high compared with substances like metals. Therefore large amounts of heat can be stored without a potentially lethal increase in temperature.

Peripheral and Central Thermoreceptors Sense Environmental and Core Temperatures, Respectively

To regulate body temperature, animals have a variety of *thermoreceptors* that sense temperatures at various locations within the



• **Fig. 53.6** Skin thermoreceptors sense cold and warm temperatures, as well as cold pain and heat pain at temperature extremes. Shown are temperature-specific discharge frequencies for the different classes of thermoreceptors, as well as the various transient receptor potential (TRP) ion channels that are the substrates for those thermoreceptor responses. (From Hall JE. *Guyton and Hall Textbook of Medical Physiology.* 13th ed. Philadelphia, PA: Elsevier; 2016.)

body. These sensors relay information to the brain, which then initiates mechanisms either to increase or decrease heat loss or production.

When an animal is exposed to a change in ambient temperature, considerable heat loss or heat gain can occur before a change in core temperature occurs. Environmental temperatures are detected by *peripheral thermoreceptors* located in the skin. These peripheral thermoreceptors possess transient receptor potential (TRP) ion channels that are activated at specific temperatures. Different types of TRP channels respond to cold and warm stimuli (Fig. 53.6). There are some chemical agonists for TRP channels useful for understanding their function. For example, menthol is an agonist for the cold receptor TRPM8, and thus the application of menthol-containing compounds to the skin or mucous membranes evokes a sensation of cool. Capsaicin, found in hot peppers, is an agonist for the heat-pain receptor TRPV1.

The most numerous thermoreceptors in the skin respond to cold. If these thermoreceptors are activated, the body can initiate mechanisms for heat conservation and heat production before the core temperature decreases. Skin cold receptors are particularly sensitive to the *rate* of decrease in temperature. For this reason, shivering can occur after exercise as the skin is rapidly cooled by sweat evaporation, despite the core temperature being normal or slightly elevated. Skin receptors sensitive to heat also exist and can initiate heat loss when the skin temperature rises.

Temperature-sensitive neurons are also present in the hypothalamus, spinal cord, great vessels, and abdominal viscera. The deep location of these *central thermoreceptors* is ideal for detecting core body temperatures. They provide feedback to the thermoregulatory center in the hypothalamus, which integrates information from the various thermoreceptors and directs needed heat generation or dissipation.

Information From Peripheral and Central Thermoreceptors Is Integrated in the Hypothalamus to Regulate Heat Loss and Conservation

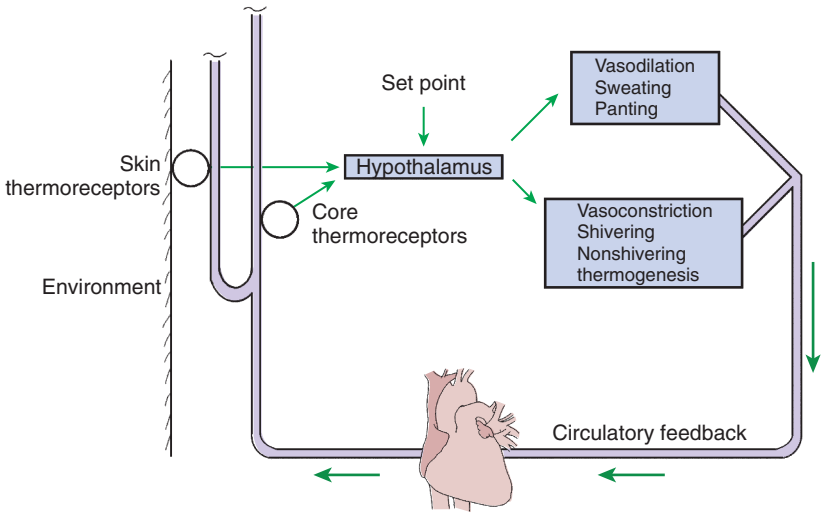
Fig. 53.7 shows the *feedback control mechanisms* for the regulation of body temperature. Central integration of the information from various thermoreceptors occurs in the anterior hypothalamus. Information from central thermoreceptors seems to predominate over information from peripheral thermoreceptors in the skin. For this reason, a rise in core temperature of only 0.5° C causes a sevenfold increase in the amount of skin blood flow; similarly, a modest decrease in core temperature initiates vasoconstriction and shivering. The effect of central thermoreceptors is about 20-fold greater than the effect of peripheral thermoreceptors.

The hypothalamus has a thermoregulatory *set point.* When the core temperature rises above the set point, heat-losing mechanisms are initiated; when temperature decreases below the set point, heat conservation or production begins. Information from peripheral thermoreceptors can slightly modify the set point, and thus shivering begins at a higher core temperature when the skin is cool than when it is warm. Similarly, sweating is initiated at a higher core temperature when the skin is cool than when it is warm.

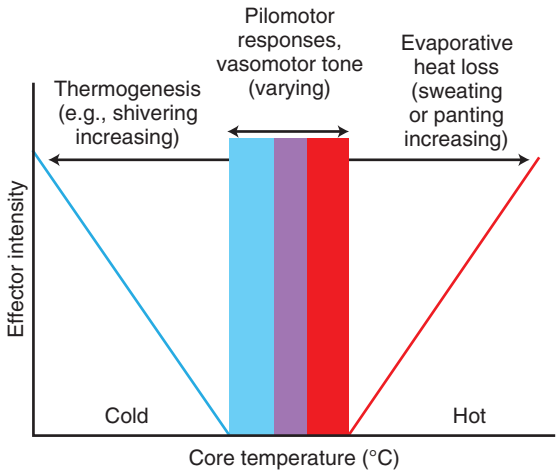
Integrated Responses

The Responses to Heat Stress Include Peripheral Vasodilation, Increased Evaporative Cooling, and Behavioral Modifications

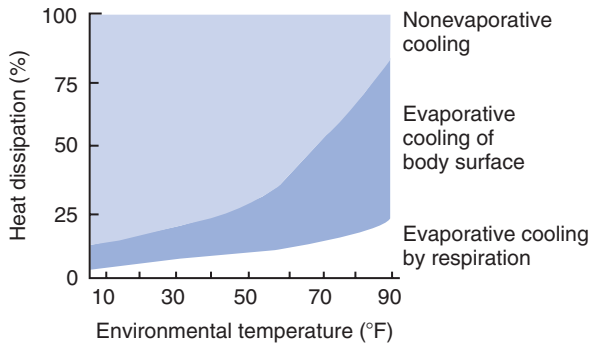
For all mammals and birds, an environmental temperature range exists at which body temperature can be maintained in a normal



• **Fig. 53.7** Feedback control mechanisms for the regulation of body temperature. Thermoreceptors in the skin and the core relay information to the hypothalamus, which adjusts the responses either to conserve and produce or to lose heat. The results of these responses are relayed to the receptors through circulatory feedback.



• **Fig. 53.8** Relationship between the intensity of thermoregulatory responses (blue and red lines) and core temperature. The set point for temperature regulation is indicated by the purple bar zone. On either side of this set point is a zone where temperature can be maintained by vasomotor responses (*blue* and *red bar zones*). As the core temperature deviates more dramatically from the set point, there is a need either to increase thermogenesis during cold stress or to increase heat loss during heat stress by means other than vasomotor action. (Modified from Bligh J. Temperature regulation in environmental physiology of animals. In: Bligh J, Cloudsley-Thompson JL, MacDonald AG, eds. *Environmental Physiology of Animals*. Oxford, UK: Blackwell Scientific; 1976.)



• **Fig. 53.9** Methods of heat loss used by a cow as the environmental temperature increases. At low temperatures, most heat loss is by nonevaporative cooling (*light purple shading*), but as the environmental temperature increases, the cow becomes increasingly dependent on evaporation (*darker purple shading and white*).

range, primarily by vasomotor mechanisms (Fig. 53.8). This *thermoneutral zone* varies with the metabolic rate and the amount of insulation. Pigs, which lack fur, clearly have a higher thermoneutral zone than do sheep, which have thick wool. Dairy cattle that have high milk production produce so much metabolic heat that their thermoneutral zone is 4° to 15° C (39° to 59° F), which is considerably lower than that of other domestic animal species. In the thermoneutral zone, body temperature can be regulated by vasomotor mechanisms that increase or decrease skin blood flow and therefore change the amount of heat loss by convection and radiation.

When an endotherm is exposed to heat stress, the initial response is vasodilation, which increases skin and limb blood flow. The resulting increase in skin temperature and the extension of core temperature down the limbs increase the temperature gradient between the skin and the environment, resulting in more heat loss by convection and radiation (see Fig. 53.4).

If vasodilation alone is ineffective in maintaining a normal temperature, evaporative cooling is increased by sweating, panting, or both. As noted earlier, evaporative cooling is the only method of heat loss available when the environmental temperature exceeds the skin temperature and is most effective when relative humidity is low. Fig. 53.9 shows that, at −10° C (14° F), cows lose less than 15% of their heat by evaporation, but as the ambient temperature rises to about 32° C (90° F), they lose 80% by evaporation. As relative humidity rises, animals have increasing difficulty losing heat; therefore exercise in hot, humid conditions may result in heat stroke. Alternatively, thermoregulatory mechanisms can lead to fluid balance disturbances that result in heat exhaustion. Heat stroke and heat exhaustion can be major concerns for horses competing in the summer when high temperatures are combined

with high humidity, such as occurred in the Olympics in Atlanta (1996), Hong Kong (2008), and Rio de Janeiro (2016). Heat stroke and heat exhaustion are described in more detail later in the section *Disorders of Heat and Cold Stress*.

Animals also use behavioral methods to resist heat stress. These methods, which include seeking shade, standing in water, and wallowing in mud, are not available to intensively managed livestock and access may be limited to companion animals as well. In hot environments, owners and producers must assume responsibility for their animals' comfort and survival. Because high-producing dairy cattle have such a low thermoneutral temperature, their primary requirement is shade in hot climates, which is a greater concern than a source of heat or insulation in cold climates.

The Responses to Cold Stress Include Peripheral Vasoconstriction, Piloerection, Increased Metabolic Heat Production by Shivering and Nonshivering Thermogenesis, and Behavioral Modifications

As the ambient temperature decreases, endotherms initially conserve heat by peripheral vasoconstriction. This sets up a temperature gradient along the limbs and reduces skin temperature, so there is only a limited temperature gradient for radiation and convective heat loss (see Fig. 53.4). Piloerection provides insulation and also decreases heat loss. Further cold stress initiates increases in metabolic heat production by shivering or nonshivering thermogenesis. Adult mammals can shiver, and neonates born in an advanced stage of development, such as lambs and foals, can also shiver. Puppies and other less-developed neonates cannot shiver; they rely on the warmth of their mother and the nest to protect them from cooling. Brown adipose tissue, which is present in many neonates and small mammals, provides a source of energy for nonshivering thermogenesis.

Chronic exposure of animals to cold results in increased secretion of *thyroxine* and an increase in basal metabolism, which increases basal heat production. When animals are housed in conditions where they receive natural light, the thickness of the hair coat increases at cold times of the year. Hair growth is the result of decreasing daylight as cold weather approaches.

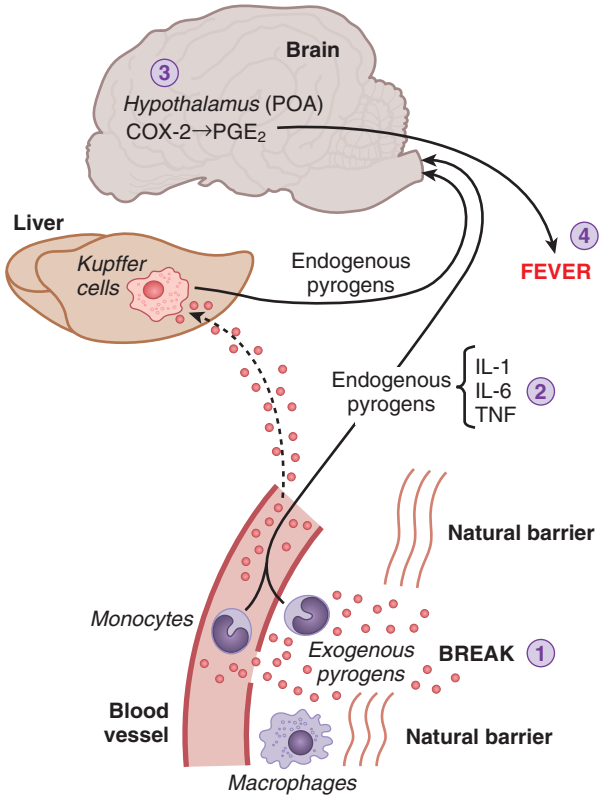
Behavioral adaptations to cold stress include curling up, piling or huddling with littermates, and seeking sunlight and/or contact with warm objects. These behaviors serve to decrease heat loss and increase heat gain.

Fever Is an Elevation of Body Temperature That Results From an Increase in the Thermoregulatory Set Point

Fever, also known as *pyrexia,* occurs in response to the elevation of an animal's thermoregulatory set point and often accompanies infectious diseases. Fever is believed to be an evolutionary adaptation to fight off infection. Studies indicate that during infection, an increase in body temperature enhances leukocyte activity (see Chapter 54). This results in a decrease in animal morbidity and mortality from infections.

The molecular mechanisms that underlie the development of fever are complex, involving cytokines, prostaglandins, complement factor 5a (C5a), and norepinephrine, among other factors that, in concert, trigger a febrile response. Both humoral (factors carried in fluids such as blood or lymph) and neural mechanisms for fever have been proposed. Knowledge about these intricate cascades of events is continually evolving, and a review of the basic principles is provided here.

Molecules that induce fever are termed *pyrogens. Exogenous pyrogens* are pyrogens that originate from outside the body; they may be microbial or nonmicrobial in origin and include various bacteria, viruses, fungi, toxins, antigen-antibody complexes, and pharmaceuticals, among other substances. The induction of fever begins when exogenous pyrogens enter the body and activate the innate immune system (Fig. 53.10). More specifically, the pyrogens activate complement (a system of plasma-borne proteins assisting with antibody/phagocyte attack or lysis of pathogens), as well as interact with Toll-like receptors (TLR; see Chapter 54) on monocytes and macrophages, including specialized liver macrophages, termed Kupffer cells, that are exposed to circulating exogenous pyrogen as blood is filtered through the hepatic sinusoids. The various mediators induced, as exogenous pyrogens stimulate monocytes and macrophages, are transported in the blood to the brain and thus this system of signaling is referred to as the *humoral mechanism of fever induction* (Fig. 53.10). Important components of the humoral mechanism of fever are generated by the initiation of the TLR-signaling pathway as it stimulates the production of cytokines (immune cell signaling molecules), which themselves can serve as



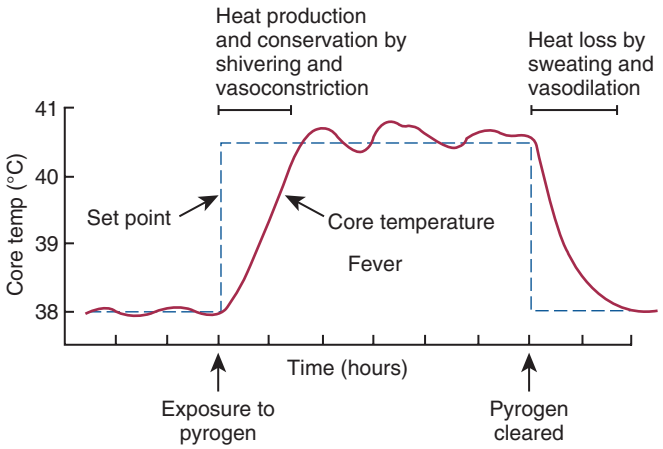
• **Fig. 53.10** Humoral mechanisms of fever induction involve (1) a break in the body's natural barrier to allow entry of exogenous pyrogens, (2) an immune signaling cascade that results in monocyte/macrophage/Kupffer cell production of endogenous pyrogens, which are transported to the POA where (3) COX-2 enzyme induces the production of PGE_2, and (4) PGE_2 acts to raise the thermoregulatory set point, resulting in fever. The role of the complement system is not represented. *COX-2,* Cyclooxygenase-2; *IL-1,* interleukin-1; *IL-6,* interleukin-6; *PGE_2,* prostaglandin E_2; *POA,* preoptic-anterior hypothalamic area; *TNF,* tumor necrosis factor.

endogenous pyrogens, so termed as they are produced within the body. The cytokines involved in the pyrogenic signaling cascade include interleukin (IL)-1, as well as tumor necrosis factor-alpha (TNF-α), IL-6, and interferons (IFN).

When released into the blood, endogenous pyrogens travel to the vascular network supplying the thermoregulatory region of the anterior hypothalamus called the *organa vasculosum of the lamina terminalis* (OVLT). This region of the hypothalamus is highly vascularized, and the blood-brain barrier there is almost nonexistent, so endogenous pyrogens and prostaglandins enter the brain easily from the bloodstream. In the nearby preoptic-anterior hypothalamic area (POA), the endogenous pyrogens act on the vascular endothelial cells to produce COX-induced prostaglandin E_2 (PGE_2), which is the central mediator of the febrile response. This locally produced PGE_2 causes the thermoregulatory set point to rise.

Along with the humoral-mediated pyrogen signaling described earlier, there is mounting evidence of a *neural mechanism for fever induction.* Like the humoral pathway for pyrogen signaling, the neural-mediated pathway starts with exogenous pyrogen circulating through hepatic sinusoids. In the neural mechanism, pyrogens activate complement on Kupffer cells, which leads to subsequent PGE_2 production; this PGE_2 locally binds to hepatic vagal afferent (sensory) nerve fibers that relay the pyrogenic message to the POA. In this manner, the presence of pyrogen is rapidly detected and the febrile response can be initiated within minutes.

When the thermoregulatory set point increases, the animal initiates responses to conserve and produce heat until the body temperature reaches the new set point (Fig. 53.11). Shivering, peripheral vasoconstriction, piloerection, and huddling behavior are all characteristics of the onset of fever. When the new set point is reached, the animal maintains its body at the new temperature until the pyrogen production ceases and it is cleared from the body. When this occurs, the set point decreases back to normal, and the animal initiates heat-losing mechanisms such as vasodilation and sweating to decrease body temperature. Because the local production of PGE_2 in the hypothalamus is involved in increasing the set point, nonsteroidal anti-inflammatory drugs (NSAIDs; e.g., aspirin, flunixin, ibuprofen) are used to treat fever. These antipyretic drugs block the enzyme cyclooxygenase, an integral enzyme in the arachidonic acid cascade, thus blocking prostaglandin production.



• **Fig. 53.11** Events involved in fever. Exposure to a pyrogen increases the set point for the temperature-regulating system. This results in heat production and conservation to elevate body temperature, which in turn results in fever. When pyrogen is cleared from the body the fever ends, the set point decreases, and heat must be lost from the body through sweating and vasodilation.

Disorders of Heat and Cold Stress

Heat Stroke Occurs When Heat Production or Input Exceeds Heat Output, so Body Temperature Rises to Dangerous Levels

In hot, humid weather it is difficult for animals to lose heat because evaporative cooling cannot occur effectively. Strenuous exercise under these conditions exacerbates the risk of ineffective cooling. Similarly, when dogs are closed in cars in warm weather, their panting saturates the air with water vapor, so further heat loss by evaporation is impossible. These environmental conditions deter or even prevent the loss of body heat, which can result in *heat stroke.* Heat stroke is the rise in body temperature due to failure of thermoregulatory mechanisms in conditions of high temperature, which is exacerbated by high humidity and/or exercise. As the body temperature rises, the metabolic rate increases, and more heat is produced. In addition, panting and sweating lead to dehydration, which decreases circulatory volume and makes it more difficult to transfer heat to the skin. When the body temperature exceeds 41.5° to 42.5° C (~107° to 109° F), cellular function is seriously impaired and consciousness is lost. As noted earlier, the respiratory system is important for heat loss through delivery of heat to the mucus membranes, as well as through panting. Thus animals with a compromised respiratory system, due to disease (such as pneumonia) or breed conformation (associated with the distorted nasal cavity and laryngopharynx in brachycephalic breeds of dogs and cats), are at increased risk of heat dissipation failure.

Thermoregulatory Mechanisms of Heat Dissipation Can Cause Hypovolemia and Hypotension Resulting in Heat Exhaustion

Heat exhaustion is another condition that occurs in hot ambient environments. Heat exhaustion is a product of dehydration that results from profuse sweating or panting, which leads to hypovolemia and subsequent hypotension. The body's attempt to dissipate heat by peripheral vasodilation further contributes to hypotension. It is this resultant profound hypotension that causes the signs of weakness or collapse characteristic of heat exhaustion. It should be noted that heat exhaustion is not a failure of heat loss mechanisms, and in many cases the body temperature of affected animals is in the normal range.

Hypothermia Occurs When Heat Output Exceeds Heat Production, so Body Temperature Decreases to Dangerous Levels

Small or sick animals exposed to a cold environment may lose more heat than they can generate, and body temperature may decrease to a point at which heat-regulating mechanisms no longer work. The ability of the thermoregulatory system to regulate core body temperature is impaired below a body temperature of about 34° C (94° F). Neonates seem to be able to withstand cooling more than adults, and in some cases lambs, piglets, and puppies that are comatose due to hypothermia can be revived through warming, leading to the adage "a hypothermic patient is not dead until it's warm and dead."

Frostbite Occurs When Ice Crystals Form in the Tissues of the Extremities

In extremely cold conditions, when the extremities are vasoconstricted to conserve heat, the tissues may cool below the freezing point of tissue water. *Frostbite* is the condition of ice crystals forming in the tissues, with resultant cell lysis and tissue necrosis. One mechanism used to prevent such cold injury is cold-induced vasodilation; this is a cyclic fluctuation in blood flow to peripheral tissues caused by intermittent dilation of vascular smooth muscle allowing increased blood flow to extremities. This provides sufficient warm blood to avoid cold injury, alternating with vasoconstriction to limit peripheral heat loss. This mechanism is important in animals that winter outdoors in cold climates.

Acknowledgement

The author thanks Dr. N. Edward Robinson for permission to build on his original rendition of this chapter.

CLINICAL CORRELATIONS

Influenza in Pigs

History

You are called to examine a group of 3-month-old pigs in an intensively managed grower and finishing facility. The group of 20 pigs is in a pen, and there are multiple similar pens within the barn. In the last 2 days in this particular pen of pigs, the animals have been reluctant to eat and have started huddling together. The owner has observed that the outer pigs in the huddle continually try to burrow toward the center of the pile of pigs and that they appear to be trembling. At this time the remaining pigs in the barn are not affected. When you enter the barn and the pigs are disturbed, they begin sneezing and coughing, and some are reluctant to move.

Clinical Examination

Three pigs are caught, and their rectal temperatures are found to be near 41° C (normal, 38.7°–39.8° C). They have nasal discharge and their conjunctiva and nasal mucosa are congested. You treat the pigs with antibiotics, and over several days, the pigs recover; however, the disease spreads progressively through the remaining pens in the barn. All pigs show the same clinical signs and no pigs die of the disease. Oral fluid samples are taken from the acutely affected pigs to test for influenza A virus and porcine reproductive and respiratory syndrome (PRRS) virus by reverse transcriptase-polymerase chain reaction (rt-PCR).

The samples are positive for influenza A virus, which has a high morbidity rate but a low mortality rate.

Comment

The clinical signs produced by this disease are caused largely by the development of fever. The pigs that were examined had an elevated body temperature because the infection had raised the set point of their thermoregulatory centers to a higher value. To raise body temperature to this new value, the pigs huddled together, and the pigs on the outside shivered to generate metabolic body heat. When the infection is overcome and pyrogens are removed by metabolism, the behavior of the pigs changes; they need to lose heat, so they separate and move around the pens more freely.

Heat Stroke in a Boxer Dog

History

At 3 pm on a hot (95° F, 35° C), humid day in August, you receive a frantic phone call from a client. The client let his 10-year-old Boxer out in the backyard midafternoon. There was no water and minimal shade in the yard. When he checked on the dog about an hour later it was lying prostrate with its tongue hanging out of its mouth and was minimally responsive when called. You instruct the owner to bring the dog to your practice immediately and to drive with the windows open for the half-mile trip.

Clinical Examination

On arrival at the clinic, the dog is weak and fails to respond to its name. Its mouth is open, its tongue is distended, and its mucous membranes are dry. Body temperature is 42.2° C (normal, 37.9°–39.9° C).

From the history, the animal's body temperature, and its lack of response, you diagnose heat stroke. The dog is placed in a bath of cool water and fluids are administered intravenously. Within 5 to 10 minutes, the dog begins to look around and recognizes its owner. The water bath treatments are continued for 2 hours, at which time the body temperature is close to normal. The dog remains in the hospital overnight and is then discharged to the relieved owner the next day.

Comment

The high ambient temperature and lack of access to shade led to a situation where the environmental temperature exceeded body temperature. At this time, the only mechanism available for losing heat is evaporation of water from the respiratory tract, which the dog attempts by panting and salivating. For a short time, this is an effective means of losing heat. But the panting results in dehydration, which reduces the circulatory volume and thus lessens the dog's ability to transfer heat to the body surface by convection. Curtailed convection results from peripheral blood vessels vasoconstricting to maintain blood pressure in the presence of reduced blood volume, which results in less blood (and thus less heat) flowing to the extremities from the core of the body. Eventually the animal cannot dissipate sufficient heat, and the body temperature begins to rise. Brachycephalic dogs, such as Boxers, have an added disadvantage in temperature regulation; the short nose and convolutions in the wall of the pharynx increase the work of breathing, especially when the dogs pant. This increased work is an additional source of body heat. The anatomical configuration of the brachycephalic upper airway, which is characterized by a foreshortened facial skeleton with soft tissue structures of the oral cavity that are not proportionally reduced, and may also include stenotic nares, elongated soft palate, everted laryngeal saccules, laryngeal edema, and hypoplastic trachea, further reduces the effectiveness of evaporative cooling in brachycephalic animals.

Therapy for this condition is to reduce body temperature and to restore circulatory function as rapidly as possible. For this reason, the dog is placed in a cool water bath to reduce body temperature and also receives intravenous fluids to rehydrate it by expanding its circulatory volume and restoring the ability of the circulation to redistribute heat within the body.

Practice Questions

1. Sweating is an effective cooling mechanism because:
- **a.** Sweat secretion produces heat, which is carried to the skin surface in the sweat.
- **b.** Conversion of sweat into water vapor requires heat, which is supplied to the skin by blood flow.
- **c.** Sweat dripping from the body carries away large amounts of heat.
- **d.** The ions in sweat carry large amounts of heat from the body.

2. In the cold, animals both conserve and produce heat. Which of the following is a method of heat conservation?
- **a.** Shivering
- **b.** Brown fat metabolism
- **c.** Increased thyroxine secretion
- **d.** Countercurrent heat exchange in the limbs
- **e.** All the above

3. Which of the following methods of heat loss can occur in an animal (body temperature = 38° C) standing in a room (temperature = 40° C) with relative humidity of zero? The walls of the room have a temperature of 30° C.
- **a.** Convection and evaporation
- **b.** Convection and radiation
- **c.** Evaporation and radiation
- **d.** Radiation alone
- **e.** Convection, evaporation, and radiation

4. Which of the following describes thermoregulation?
- **a.** Temperature receptors in both the brain and the skin can initiate thermoregulatory responses.
- **b.** The brain temperature receptors have a greater influence on thermoregulation than do skin receptors.
- **c.** The core temperature at which shivering begins is higher if the skin is cold than if it is warm.
- **d.** Skin cooling can initiate shivering even if core temperature is normal.
- **e.** All the above.

5. Which of the following correctly describes fever?
- **a.** It results when the set point for body temperature decreases.
- **b.** It is accompanied by sweating to lose heat as body temperature rises.
- **c.** It is accompanied by shivering to gain heat as body temperature decreases when pyrogens are metabolized.
- **d.** It can be initiated by pyrogens from bacteria or leukocytes.
- **e.** All the above.

Bibliography

Boron WF, Boulpaep EL, eds. *Medical Physiology*. 3rd ed. Philadelphia: Saunders; 2017.

Cinti S. Between brown and white: novel aspects of adipocyte differentiation. *Ann Med*. 2011;43(2):104–115.

Gordon CJ. The mouse thermoregulatory system: its impact on translating biomedical data to humans. *Physiol Behav*. 2017;179:55–66.

Hales JR. The partition of respiratory ventilation of the panting ox. *J Physiol*. 1967;188(2):45P–46P.

Li Y, Lasar D, Fromme T, Klingenspor M. White, brite, and brown adipocytes: the evolution and function of a heater organ in mammals. *Can J Zool*. 2014;92:615–626.

Maughan RJ, Lindinger MI. Preparing for and competing in the heat: the human perspective. *Equine Vet J Suppl*. 1995;20:8–15.

Randall D, Burggren W, French K. *Eckert Animal Physiology: Mechanisms and Adaptations*. 5th ed. New York: Freeman; 2001.

Roth J, Blatteis CM. Mechanisms of fever production and lysis: lessons from experimental LPS fever. *Compr Physiol*. 2014;4:1563–1604.

Schmidt-Nielsen K. *Animal Physiology: Adaptation and Environment*. 5th ed. Cambridge, UK: Cambridge University Press; 1997.

Sherwood L, Klandorf H, Yancey PH. *Animal Physiology, From Genes to Organisms*. 2nd ed. Belmont, CA: Cengage; 2013.

Tansey EA, Johnson CD. Recent advances in thermoregulation. *Adv Physiol Educ*. 2015;39:139–148.

54 Antigens and Innate Immunity

S. ANSAR AHMED AND GERHARDT G. SCHURIG

KEY POINTS

Antigens

1. Antigens (or immunogens) stimulate immune cells to induce an immune response.
2. The degree of immune response depends on several characteristics of the antigen.

Body's Defense Against Invading Antigens

1. Both nonimmune and immune mechanisms defend against invading antigens.
2. A first line of defense includes physical and chemical barriers such as the skin and internal body fluids.
3. A second line of defense consists of phagocytic cells of the myeloid and macrophage-monocyte lineages.
4. Macrophage-derived cytokines can induce a variety of physiological processes to help combat infectious antigens.

The immune system performs two vital functions critical for the maintenance of homeostasis and survival: (1) inducing an effective and safe response against foreign antigens (infectious and noninfectious) and (2) avoiding a response to components of "self" antigens by enforcing stringent regulatory controls over dangerous self-reactive cells capable of mounting devastating immune attacks on "self" tissues. Because the induction of immune responses depends on antigens, this chapter first discusses the nature and characteristics of antigens.

Antigens

Antigens (or Immunogens) Stimulate Immune Cells to Induce an Immune Response

An *antigen,* or *immunogen,* is defined as any substance that is capable of stimulating immune cells (T and B cells) to induce an immune response. Antigens can be broadly divided into two large categories: (1) infectious (microbial) and (2) noninfectious (Fig. 54.1). *Infectious antigens* include components derived from bacteria, viruses, protozoa, and helminths. *Noninfectious antigens* include those derived from "self" (autoantigens), food, plants, dust, or insect and animal venoms, as well as synthetic and cell surface proteins.

An antigen is composed of many molecular units to which an antibody binds. These small units on an antigen are called antigenic *epitopes,* or *antigenic determinants.* Thus a single antigen may be composed of many antigenic epitopes. In the strictest sense, antibodies bind to an antigenic epitope of an antigen. Some of these antigenic epitopes are shared among different bacteria (e.g., epitopes on *Brucella* and *Yersinia*) or between a bacteria and host cells (e.g., *Mycobacterium* heat shock proteins and synovial tissue; *Mycoplasma* and lung tissue). These types of antigenic epitopes are called *cross-reactive epitopes.*

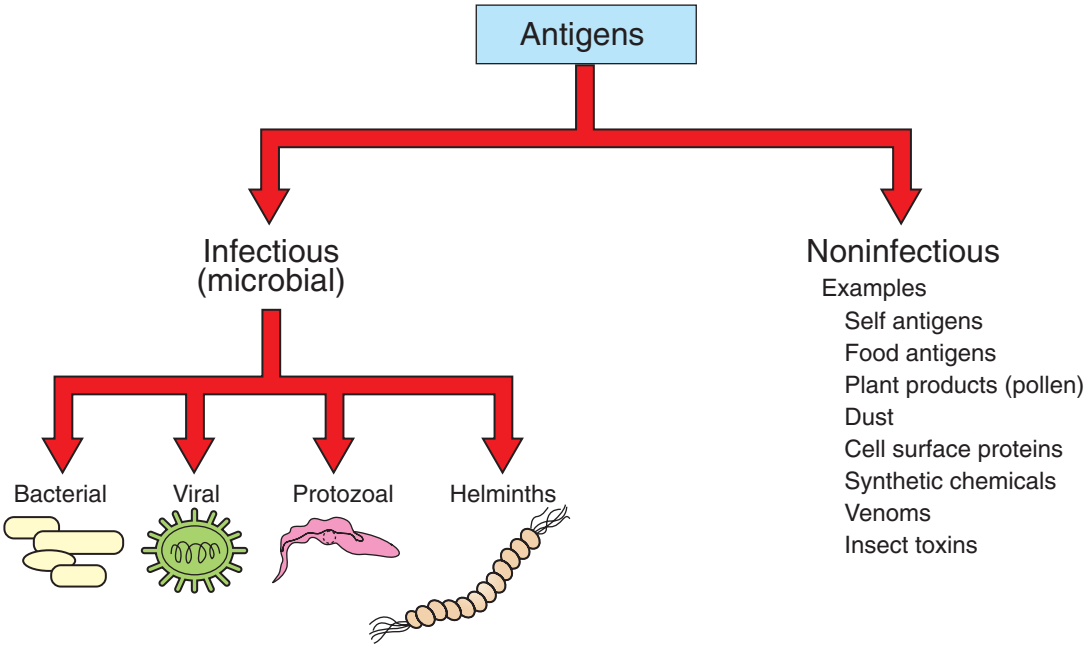
Fig. 54.2 shows the following antigenic structures of bacteria:

- *Bacterial cell wall.* Cell walls of gram-positive bacteria differ from those of gram-negative bacteria. Gram-positive bacteria are composed of a thick layer of short chains of amino acids or peptides and carbohydrates (*peptidoglycans*). The cell wall of gram-negative bacteria has a thin layer of peptidoglycan and is largely composed of *lipopolysaccharides,* which are potent endotoxins.
- *Capsule.* Certain bacteria produce a protective outer covering called a capsule, which is composed of polysaccharides.
- *Pili.* These small, hairlike protein structures on some bacteria enable the bacteria to adhere to target host cells and transfer genetic information from one bacterium to another.
- *Flagella.* Some bacteria possess flagella for mobility. Flagella contain a protein called *flagellin,* which can be antigenic.
- *Nucleic acids.* Nucleic acids, such as bacterial deoxyribonucleic acid (DNA), tend to be antigenic because of differences in methylation compared with mammalian DNA. The antibodies against bacterial DNA tend to cross-react with the host's DNA.

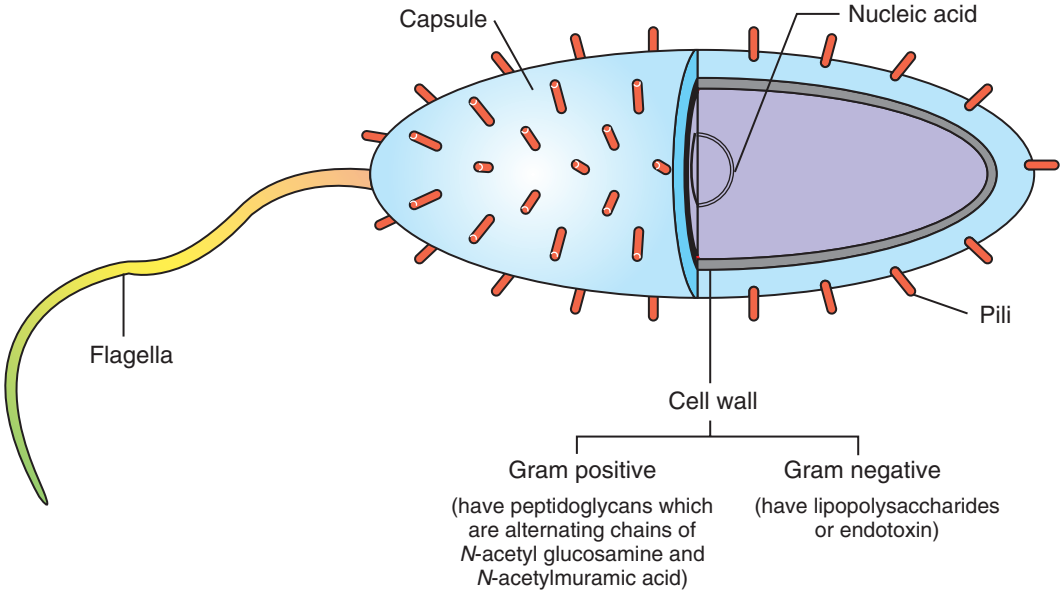
Viruses have nucleic acids (ribonucleic acid [RNA] or DNA), surrounded by a protein coat called a capsid. Some viruses have an *envelope,* a lipid membrane–like structure covering the capsid. On the envelope are glycoprotein projections that the viruses use to attach to the host target cells. All these components may be antigenic.

External structures of protozoa and helminths tend to be antigenic. Similarly, fungal spores are antigenic. Pollens, glycoproteins of certain foods, the unique biochemical structure of synthetic chemicals, insect saliva, and venoms are all good antigens. It is beyond the scope of this chapter to discuss each of these antigens in detail.

The immune system is exposed to and tolerates "self" antigens found on all of its own tissues. These antigens can be cell surface antigens (e.g., thyroglobulin, myelin peptides) or internal antigens



• **Fig. 54.1** Classification of antigens.



• **Fig. 54.2** Antigenic structures of bacteria (not all bacteria have all these structures).

(e.g., cardiolipin, nucleic acids, histones). In certain individuals who are allergic, antigens derived from food (e.g., peanuts, strawberries, fish) or plants (e.g., pollen, spores) induce an immediate and potent immune reaction. Many synthetic chemicals and drugs are minute in size and tend to be adsorbed onto cell surface antigens to create a new antigenic epitope. With ever-increasing synthesis of chemicals (pesticides, agricultural chemicals, drugs, and consumer products, to name a few), it is likely that synthetic chemicals may become an important class of antigens in the future.

The Degree of Immune Response Depends on Several Characteristics of the Antigen

The degree of immune response induced by an antigen is called *antigenicity* or *immunogenicity.* Understanding the characteristics of antigens that provoke a strong or weak immune response provides important insight into the body's ability to combat invading antigens successfully. Furthermore, this understanding is useful in designing a vaccine preparation with potent antigenicity. Characteristics that contribute to potent antigenicity include the following:

- *Foreign versus self antigens.* Antigens considered to be foreign to the host tend to be highly antigenic. For example, if a horse is injected separately with antigens derived from a dog or from its own tissues, the horse will mount a strong immune response to the dog antigens but not to its own (self) tissues.
- *Size.* The size of an antigen also influences the level of immune response. Large antigens enable better processing by antigen-presenting cells (e.g., macrophages, dendritic cells) and subsequent presentation of antigenic peptides to lymphocytes for induction of an immune response. Examples of large antigens

include bacterial and insect toxins, viral capsids, surface proteins on protozoa and helminths, and venoms. At the other extreme, very small antigens (e.g., small synthetic antigens, endogenous hormones, pesticides) tend to be ineffective in provoking an immune response. Very small antigens are inherently incapable of inducing immune responses; however, when bound to a larger protein, they can be potent antigens. Such small compounds are referred to as *haptens.* A good example of a hapten is a poison ivy–derived chemical, urishiol, which readily combines with many proteins (e.g., skin proteins) to induce a vigorous immune response.

- *Biochemical structure and complexity.* In general, proteins tend to be more antigenic than lipids or carbohydrates. Large size alone is insufficient to provoke a good immune response. For example, many sugars and lipids, even though large in size, are ineffective in inducing an immune response because they consist of simple repeating units (e.g., repeating sugars in starch), which lack complexity. Complex carbohydrates and lipids, on the other hand, as found in many microbes, are strong immunogens. Carbohydrates and lipids, when combined with protein to form glycoproteins and lipoproteins, respectively, have increased complexity and thus are good antigens.
- *Stability and degradability.* For immune cells to respond, stability of an antigen is an important feature. Flexible antigens, such as flagellin in a bacterium, are poor immunogens. However, when stabilized and rendered less flexible, as done in vaccine preparations, flagellin tends to be a potent immunogen. For an immune response to be initiated, the antigen ingested by phagocytic cells (e.g., macrophages) must be degraded and broken down into small peptides. Lymphocytes (T cells) will only respond to the peptides and not to large, native molecules. Antigens such as steel pins or plastic heart valves, even though large and complex, are inert and not degradable and thus are not good antigens.

Large, complex proteins (or lipoproteins or glycoproteins) that can be degraded and processed therefore tend to be excellent antigens. Other parameters that influence an individual's ability to respond to antigens include genetics (e.g., major histocompatibility complex genes), endogenous biomolecules that regulate and modulate immune responses (e.g., hormones, neuropeptides), and the level and route of exposure of antigens.

An antibody induced in response to an antigen will specifically bind to the antigen. Any minor alteration of the antigen will negatively impact the antibody's ability to bind to an antigen. Therefore invading microbes often alter their antigens to prevent the binding of induced antibodies, thus avoiding immune attack.

Body's Defense Against Invading Antigens

Both Nonimmune and Immune Mechanisms Defend Against Invading Antigens

The body is confronted with literally billions of antigens. Consequently, a unique challenge presented to the immune system is to respond effectively only to foreign antigens while refraining from response to "self" antigens. The induction of immune responses requires energy and protein, and antigens require extensive cellular division (and hence utilization of protein reserves); the body cannot mount immune responses to each of the innumerable antigens it encounters constantly. Instead, the body is well equipped to handle antigens effectively before resorting to a specific immune response.

Initially, most antigens are effectively handled by nonspecific defense mechanisms, such as impervious and formidable physical barriers (e.g., skin and other body surfaces) and antimicrobial body fluids (e.g., lysozymes in tears, saliva, gastric juices). These are considered the first line of defense and are discussed next. Should the antigen survive this "body armor," phagocytic cells (e.g., neutrophils, macrophages-monocytes) and natural killer (NK, NK-T) cells can effectively eliminate the invading antigens. These cells ingest and destroy a wide range of antigens and thus are nonantigen-specific. These cellular defenses constitute a second line of the body's defense. The body's initial defense (physical, chemical, and phagocytic antigen-presenting cells; natural killer cells) constitute the *innate immune system.* The antigen-presenting cells closely interact with specific T and B cells to induce a specific immune response. Thus specific immune responses by the "adaptive" immune system tend to be the last line of the body's defense (see Chapter 55). Collectively, both nonimmune and immune mechanisms effectively counter invading microbes.

A First Line of Defense Includes Physical and Chemical Barriers Such as the Skin and Internal Body Fluids

The physical nonimmune defense barriers include external body surfaces such as skin and internal body surfaces such as the gastrointestinal (GI), reproductive, respiratory, and the urogenital tracts. The skin plays a major role in preventing the entry of organisms through a variety of nonimmunological means, including the secretion of sebum from sebaceous glands, which maintains a low pH, and secretion of enzymes that are not conducive for the invading pathogens. Periodic natural desquamation of the skin also results in sloughing off any invading pathogens. Nonpathogenic bacteria also occupy skin surfaces, thereby preventing the adherence of pathogenic organisms to their target cells, which is prerequisite for entry into the body. Any changes in the skin, such as cuts, burns, and dry or very humid skin, will result in the entry of microbes. In addition to the nonimmunological mechanisms, skin is also rich in dendritic cells (Langerhans cells) and γ-δ T cells that contribute to warding off invading pathogens. The natural flushing action of urine and milk assist in elimination of infectious antigens, as evidenced by the infectious conditions that result from stasis of urine or milk.

Many body fluids are inhospitable to invading pathogens. For example, mucus in the mucosal tissues (respiratory, urogenital, and GI tracts), saliva, tears, gastric juices, and urine are rich in enzymes (e.g., lysozymes) and are low in pH. As with the skin, the GI tract is covered with nonpathogenic bacteria, which prevent the adhesion of pathogenic bacteria to their target cells. Furthermore, resident normal bacterial florae in gastric tissues secrete butyric or lactic acids, which not only maintain a low pH in gastric fluids, but also are bacteriostatic to other microbes. Vaginal epithelium is rich in glycogen and promotes the growth of *Lactobacillus,* which secretes lactic acids. In the respiratory tract, the antigen load is decreased by a variety of mechanisms, including the turbulence created when air is inhaled due to the anatomical construction of the lower respiratory tract, which narrows and branches. Microorganisms in inhaled air are carried by this turbulence and are forced onto the walls of the respiratory tract, which are rich in sticky mucus and bactericidal lysozymes. The ciliary action of the respiratory tract also eliminates antigens effectively.

A Second Line of Defense Consists of Phagocytic Cells of the Myeloid and Macrophage-Monocyte Lineages

Should an antigen survive the first line of the body's defense (i.e., body surfaces) and penetrate into blood vessels and tissues, the body's defense relies on cellular response. Key cells involved in cellular defense are phagocytic cells, which are an integral part of *innate immunity.* These cells, based on their cellular origin, are broadly divided into myeloid and macrophage-monocyte lineages. Included in the myeloid lineage are neutrophils, eosinophils, and basophils. The macrophage-monocyte series includes monocytes and macrophages. *Neutrophils* constitute the largest percentage of white blood cells in most species (60% to 65%), except in ruminants (20% to 25%). Neutrophils have a short life span in the blood (half-life, ~12 hours), but in tissues their longevity increases to several days. Neutrophils are approximately 12 μm in diameter, with multilobulated nuclei and a cytoplasm rich in granules. The population of granules is composed of both primary and secondary granules. The primary granules contain important bactericidal enzymes such as myeloperoxidase, lysozymes, acid hydrolases (e.g., β-glucuronidase, cathepsin), and neutral proteases against elastase hydrolases. Neutrophils also have *defensins,* small proteins that are inserted between the lipid bilayers and disrupt the interactions emanating from lipid membranes. The secondary granules include lysozymes, lactoferrin, and collagenases.

Neutrophils are considered the first responder cells to combat invading antigens. The primary function of neutrophils is to capture and destroy antigens. Neutrophils, unlike monocytes or macrophages, respond rapidly to invading antigens and readily phagocytose antigens. However, neutrophils lack the capacity to present antigens to lymphocytes. Neutrophils destroy antigens by two different but complementary mechanisms: (1) phagocytosis and (2) respiratory burst. Phagocytosis in turn is divided into four arbitrary stages: (1) chemotaxis, (2) adherence or attachment, (3) phagocytosis, and (4) destruction (Fig. 54.3).

Neutrophils are attracted to the site of infection in tissues by chemokines and chemical messengers released when tissues are damaged. In response to chemical signals and chemokines, vascular endothelial cells induce the expression of adhesion molecules. Neutrophils bind to these cellular adhesion molecules (CAMs) through specific receptors and are triggered to leave the circulation by crossing capillary walls (diapedesis) into tissues. When neutrophils move out of the circulation, they move toward the antigen. The contact between neutrophils and the antigen is greatly facilitated when antigens are coated or bound by a host's proteins, such as complement or antibodies. These proteins that enhance contact and phagocytosis by neutrophils or other phagocytes are called *opsonins.* The contact of neutrophils with antigens triggers infolding of the cell membrane (by the action of actin and myosin), and the antigen is trapped in a vacuole called a *phagosome.* The primary granules move toward the phagosome and fuse their membrane to become *phagolysosomes,* and in the process the granules release deleterious bacteriostatic and bactericidal biomolecules. Thus in the contained environment of phagolysosomes, the antigen is destroyed.

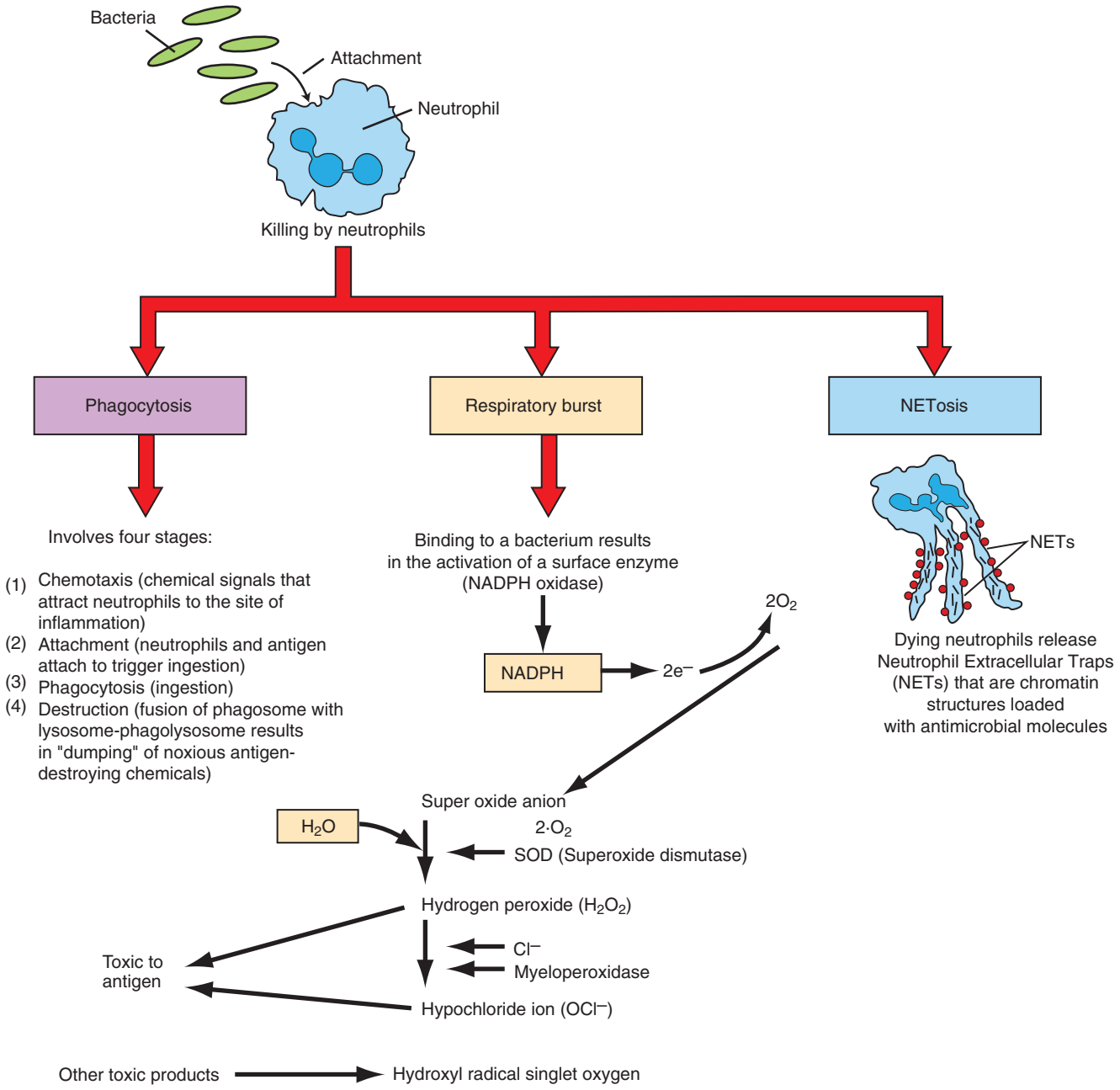
A concurrent mechanism by which neutrophils kill invading microbial antigens involves *respiratory burst* (see Fig. 54.3). On contact of neutrophils with an antigen, consumption of oxygen is immediately increased, 70–100-fold. This results in activation of an enzyme, nicotinamide adenine dinucleotide phosphate (NADPH) *oxidase,* which forms an electron transport chain with cytosolic NADPH as an electron donor of oxygen. A molecule of oxygen accepts two donated electrons to result in a superoxide anion (O_2^-). This O_2^-, under the influence of the enzyme superoxide dismutase and in the presence of water, will chemically react to yield hydrogen peroxide (H_2O_2), which is toxic to microbes. This H_2O_2, under the influence of myeloperoxidase and utilizing chloride ions (Cl^-), catalyzes oxidative reactions to form H_2O_2 and halide ions. All these products are highly toxic to the antigens. Neutrophils are also known to release hydroxyl radical singlet oxygen, which is toxic to bacteria.

Neutrophils have limited energy and a relatively short life. NETosis is a relatively recently identified form of neutrophil cell death, which is characterized by the release of neutrophil extracellular traps (NETs) loaded with antimicrobial molecules (Fig. 54.3). Activation of neutrophils by bacterial lipopolysaccharides (LPS) or proinflammatory cytokines such as tumor necrosis factor alpha (TNF-α) and IL-8 induces NET formation. These released NETs contain nuclear or mitochondrial chromatin, are histone-rich (citrullination of histones promotes NETs), and believed to trap pathogenic bacteria, protozoans, and fungi. NETs are like a spider throwing out nets to trap antigens on its "sticky" web to eliminate them and also stimulate B cells. There is some evidence that NET formation can occur in surviving neutrophils as well.

Dying neurophils can also release elastases and collagenases that serve as a powerful chemoattractant for another group of phagocytes called *macrophages,* and therefore neutrophils are sometimes referred to as "martyrs of the immune system." Macrophages are attracted by bacterial products as well but also through the chemotactic factors released from damaged tissues. Macrophages differ from neutrophils in several important aspects. Macrophages, even though not rapid-responder cells, have extensive ability to phagocytose antigens repeatedly. These long-lived cells secrete large quantities of cytokines and chemokines that play a key role in regulating immune responses. Some of these cells even have the ability to present antigens to the immune system. Macrophages are present throughout body tissues where entry of antigens is likely.

Macrophages are round or elongated cells and express many surface receptors that include major histocompatibility complex antigens I and II (MHC class I and II). MHC classes I and II play a major role in antigen recognition and presentation (see Chapter 55). Macrophages differ in their morphology based on tissues and thus are called by different names. For example, in the lymphoid organs, these cells are macrophages, whereas in the liver they are known as *Kupffer's cells* (Fig. 54.4). Macrophages are derived from the bone marrow hematopoietic cells and are initially called *monoblasts,* which mature and move into circulation, and are known as *monocytes.* When monocytes move into the tissues, they are called *macrophages.* Macrophages are larger than neutrophils and are rich in rough endoplasmic reticulum and Golgi bodies, indicating their extensive ability to produce and secrete immunoregulatory proteins.

Innate immune cells, predominantly macrophages, express Toll-like receptors (TLRs) that recognize specific sequences/molecules on microbes called pathogen-associated molecular patterns (PAMPs). These receptors serve as pattern recognition receptors that trigger an innate immune response to pathogens or other antigenic stimuli. Thus far, 13 TLRs have been identified in mammals. TLR1, 2, 4, 5, and 6 are located on the cell surface membrane, whereas TLR3, 7, 8, and 9 are found intracellularly. The combination of TLR1 with TLR2 recognizes a wide variety of fungal and protozoal products, including peptidoglycans and LPS. TLR2 alone, in addition to recognizing endogenous ligands

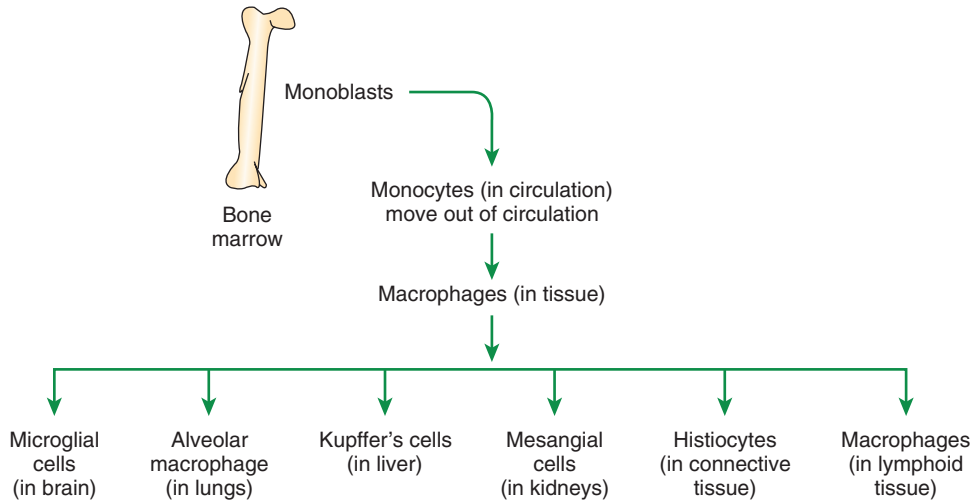


• **Fig. 54.3** Elimination of foreign material by neutrophils by three different but complementary mechanisms (phagocytosis, respiratory burst, and NETosis). *NADPH,* Nicotinamide adenine dinucleotide phosphate.

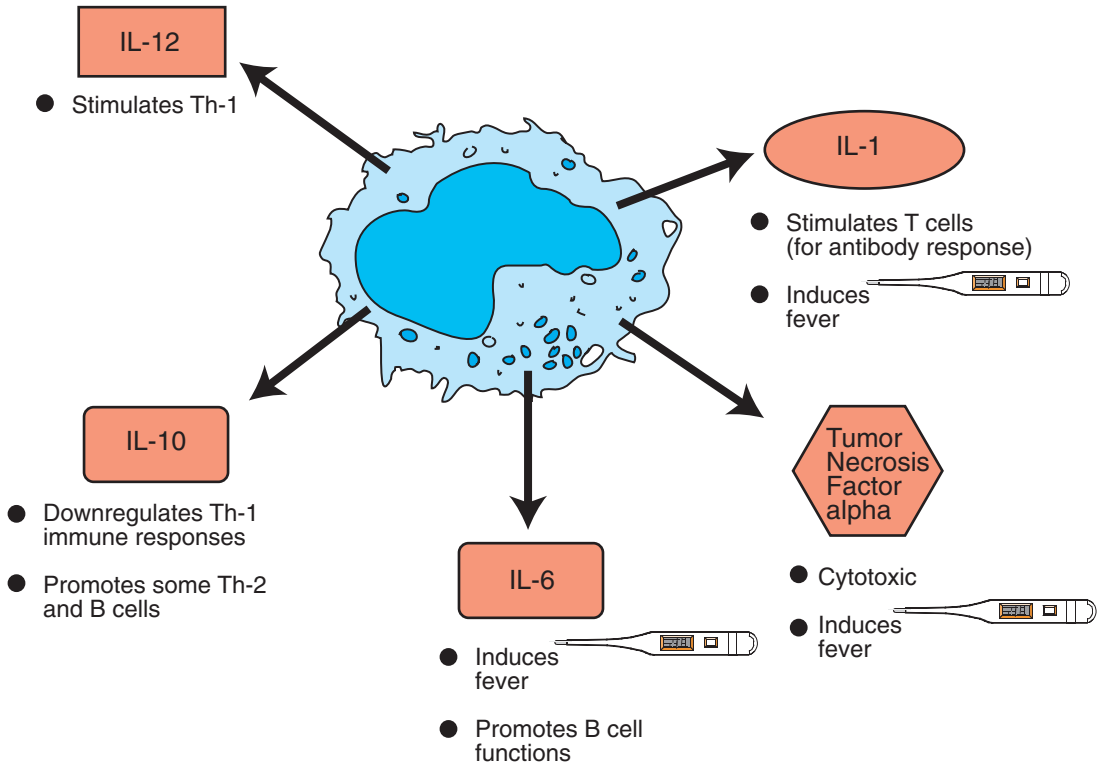
such as heat shock protein 70 (Hsp70), can recognize bacterial diacylated lipopeptides when the receptor is combined with TLR6. TLR4 is activated by LPS on gram-negative bacteria. TLR5 is expressed on the basolateral surface of intestinal epithelial cells and is activated by bacterial flagellin. Intracellular TLRs recognize intracellular pathogens such as viruses. With respect to activation of TLRs, activating TLR4 can result in initiation of two distinct signaling mechanisms, one of which is a myeloid differentiation factor 88 (MyD88)-dependent pathway. Genes induced by this pathway are mostly those coding for proinflammatory molecules (e.g., interleukins 1α, 1β, 6, 12; cyclooxygenase 2; TNF-α).

In addition to TLRs, other pattern recognition molecules have also been identified on innate immune cells. These include C-type lectin receptors (CLR) and NOD-like receptors (NLR). The CLR are a large family of proteins that contain C-type lectin domains. The Dectin receptor is a prototypical example of CLR present on the surface of neutrophils, macrophages, and dendritic cells. Dectin receptors recognize β-glucans to initiate proinflammatory responses especially against fungi whose cell wall contains β-glucans, and their role in recognizing other pathogens including *Mycobacterium tuberculosis* is also recognized. The NLRs (22 reported in humans) are cytosolic sensors for the peptidoglycan-rich bacterial cell wall that promote inflammatory cytokines, chemokines, defensins, and nitric oxide, among other antimicrobial products that counter pathogens.

Ingestion of antigen will activate the metabolic machinery of macrophages, including increasing lysosomal and bactericidal activity and upregulating the *inducible nitric oxide synthase* gene (iNOS)



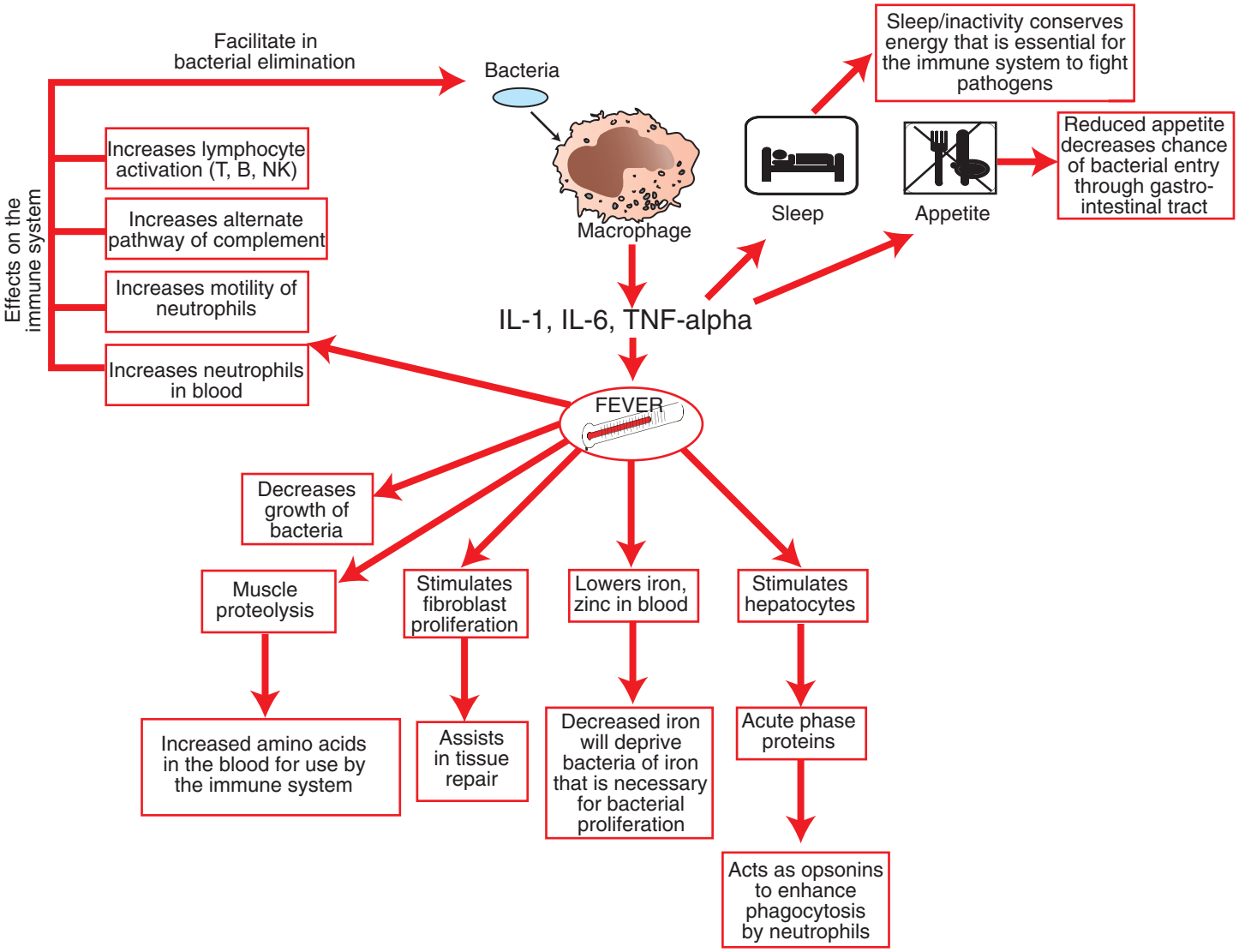
• **Fig. 54.4** Anatomical location of macrophages throughout the body.



• **Fig. 54.5** Macrophage-derived cytokines. *IL,* Interleukin; *Th-1,* T-helper cells type 1; *Th-2,* T-helper cells type 2.

that encodes for iNOS protein, which in turn is responsible for enhanced release of potent antimicrobial nitric oxide. Macrophages also release several oxygen free radicals that are also antimicrobial. Thus macrophages aid in killing of antigens by phagocytosis and by respiratory burst (oxygen and nitrogen free radicals) (see Fig. 54.3). Macrophages can potentially secrete more than 100 proteins. Some of these proteins, such as interleukin-1 (IL-1), interleukin-12 (IL-12), TNF-α, interleukin-18 (IL-18), and interleukin-27 (IL-27), play a central role in activating lymphocytes, especially naïve T lymphocytes (Fig. 54.5). After an effective immune response is initiated, macrophages also secrete cytokines such as interleukin-10 (IL-10) and transforming growth factor beta (TGF-β), which downregulate the immune response.

When the antigen is cleared by macrophages, these cells also play an important role in repairing the damaged tissues. Macrophages secrete angiogenic factors to enhance the blood supply. For example, IL-1 secreted by macrophages stimulates fibroblasts to secrete collagen to rebuild tissues.



• **Fig. 54.6** Pyrogenic cytokines, fever, and pathogen elimination.

Macrophage-Derived Cytokines Can Induce a Variety of Physiological Processes to Help Combat Infectious Antigens

Physiologically, a key component of combating infectious antigens is the induction of *fever,* which is mediated by the release of pyrogenic cytokines such as IL-1, IL-6, and TNF-α by macrophages (Fig. 54.6). These cytokines act on multiple tissues to mount a coordinated effort aimed at eliminating the invading microbes. For example, they act on thermoregulatory regions in the hypothalamus to induce fever, which is an integral part of this process. Fever accelerates the mobility of neutrophils, enhances their phagocytic ability, and activates lymphocytes and complement proteins while impeding the growth of bacteria. These cytokines also act on the liver to stimulate acute phase proteins that function as opsonins to promote phagocytosis. Further, these cytokines act on the sleep-associated regions in the hypothalamus to promote sleep in an effort to conserve energy and redirect energy toward the infectious challenge. Because major immunoprotective elements (e.g., antibodies, cytokines, complement) are all proteins, these cytokines increase the amino acid pool by acting on muscles and inducing mild proteolysis and release of amino acids, which are essential for synthesis of various immunoprotective elements.

Another important feature of some macrophages is their ability to present antigens to stimulate T cells in order to start the specific adaptive immune response. Other cells are part of the innate immune system and also serve as *antigen-presenting cells* (APCs); these include dendritic cells and B cells. *Dendritic cells* are considered the most potent cells in presenting antigens. These cells have long dendrites that physically allow them to interact simultaneously with many antigens. Dendritic cells are abundant in lymphoid organs, skin, and other tissues that frequently encounter antigens. The role of APCs in the adaptive immune response is discussed in Chapter 55.

CLINICAL CORRELATIONS

Swollen Lymph Nodes in a Colt

History

A 2-year-old colt that just returned from training has shown marked bilateral nasal discharge for at least 2 days. The colt has not been eating for the last day and appears depressed. The colt also seems to have considerable swelling under the jaw and in the throatlatch area.

Clinical Examination

On physical examination the colt has a temperature of 39.2° C (102.7° F) but a normal heart rate of 36 beats/min and normal respiratory rate of 12 breaths/min. The horse has bilateral mucopurulent nasal discharge. The lymph nodes underneath the jaw and in the throat area are both greatly swollen. No other abnormalities are present.

Comment

This colt was vaccinated for Eastern and Western encephalitis as well as rhinopneumonitis, influenza, rabies, tetanus, and West Nile virus. However, the colt was not vaccinated against *Streptococcus equi* (*S. equi*), the causative agent of strangles. Young horses not vaccinated against *S. equi* that have been traveling are at increased risk for disease. Based on the clinical signs of fever, bilateral mucopurulent discharge, and normal lung sounds and respiratory rate, this disease is likely an upper respiratory infection. The vaccination history and clinical signs indicate *S. equi* is the most likely pathogen.

Horses are exposed to *S. equi* through inhalation, ingestion, or exposure to conjunctival surfaces; association with other infected horses; or a contaminated environment. The bacteria enter the cells of the tonsillar crypts and ventral surface of the soft palate and migrate throughout the upper respiratory tract, infecting the lymph nodes as well and replicating there extracellularly. Most cases resolve over a prolonged period. In other horses, however, the guttural pouches can become infected, or the animal can become systemically infected and develop abscesses in other parts of the body (e.g., lymph nodes). In the latter situation, the syndrome is known as "bastard strangles."

The immune response associated with infection begins with the mucosal innate response. Horses will have increased levels of neutrophils in their blood (neutrophilia). The number of lymphocytes in the blood may vary, but there is a marked reaction in the local lymph nodes to the infection. These horses have severe lymphadenopathy of the regional nodes from increased numbers of neutrophils and lymphocytes. A combination of both the innate and the acquired immune response is capable of controlling the infection, and eventually the horse eliminates the pathogen. Again, however, in a small percentage of the infected horses, *S. equi* migrates to the systemic lymph nodes to cause bastard strangles at some later time point in life.

A diagnosis can be made based on history, clinical signs, and combined diagnostic testing. Traditionally, these horses have neutrophilia in their peripheral blood. Definitive diagnosis is based on culturing the agent, often from the regional lymph nodes using an aspirate. In many cases, there is concern about an outbreak within a particular barn because the agent is highly infectious, and infected horses and barns must be quarantined to prevent further spread of the disease. Another test for a definitive diagnosis is polymerase chain reaction (PCR).

Treatment

The treatment of *S. equi* infection varies based on clinical signs and the population at risk. In the early stages of infection, with clinical signs of fever, depression, anorexia, and bilateral mucopurulent nasal discharge, but no or limited lymphadenopathy, these horses can often be treated systemically with antibiotics, and the disease will usually resolve. However, when the lymph nodes begin to become enlarged or abscessed, antibiotic treatment will often prevent further spread of infection but will not aid in clearance. When antibiotics are discontinued, the disease continues to progress. Some horses become greatly depressed, with marked lymphadenopathy that can affect breathing. In these animals, the lymph nodes are often drained. Sometimes a tracheostomy is needed, in which case horses are treated with systemic antibiotics to prevent secondary infections. Disinfecting the environment is critical to the control of strangles within populations.

Heifer Calf With Recurrent Infection

History

An 8-week-old dairy heifer calf presents for coughing. She appeared to be fine at birth, but in the past 6 to 7 weeks, she has started to lose weight. None of the other calves of similar ages housed in pens are affected.

Clinical Examination

She has an increased temperature, and increased heart rate (tachycardic), respiratory rate (tachypnic), and effort. She has increased lung sounds consistent with pneumonia. Her feces are also a little looser than normal. Blood is collected for a complete blood count (CBC) as well as zinc sulfate turbidity test to examine immunoglobulin G (IgG) antibody levels. The CBC reveals very high neutrophil counts. The IgG levels are normal.

Comment

Bovine leukocyte adhesion deficiency (BLAD) is suspected in this calf. This is an autosomal recessive disease. There is impaired expression of the CD11/CD18 family (β2 integrins) of leukocyte adhesion molecules, and this prevents neutrophils from emigrating into tissues. Therefore the calves are likely to develop infection(s). Animals can be detected by PCR DNA testing. Cells can also be stained with antibodies and examined by flow cytometry for expression of CD18. Affected animals often have infections of the GI and respiratory systems as well as synovial structures (joints).

Treatment

Treatment is limited to supportive care, including antibiotics and fluid therapy, depending on the specific infection. Affected animals will succumb to infection. Euthanasia of a known affected BLAD animal should be considered as it is autosomal recessive and can be passed to future generations.

Practice Questions

1. Which of the following statements is correct in regard to macrophages?
- **a.** They are the first type of cells to appear in circulation and can repeatedly phagocytose in response to invasion by microbial antigens.
- **b.** They secrete anti-inflammatory cytokines that prevent fever.
- **c.** They have class I major histocompatibility complex (MHC) antigens, and selected macrophages also have class II MHC antigens and thus can serve as antigen-presenting cells.
- **d.** They have no ability to secrete proteins.
- **e.** They are actually lymphocytes.

2. Which of the following is correct about neutrophils?
- **a.** They can rapidly destroy antigen by phagocytosis, respiratory burst, and NETosis.
- **b.** They are lymphocytes that are the first cells to destroy antigens.
- **c.** They are major producers of IL-1.
- **d.** They are efficient antigen-processing cells.
- **e.** All the above.

3. Antigen-presenting cells (APCs) include which of the following?
- **a.** Dendritic cells
- **b.** Selected macrophages
- **c.** B cells
- **d.** T cells
- **e.** Answers a, b, and c

4. Which of the following is correct with regard to antigen?
- **a.** Highly degradable proteins and large, inert substances are not good antigens (i.e., not able to induce a strong immune response).
- **b.** Large molecules are good antigens, and therefore large molecules such as polysaccharides, with simple repeating sugars, are good antigens.
- **c.** Large complex proteins (especially foreign antigens) are good at inducing immune responses.
- **d.** Antigens do not drive an immune response.
- **e.** Lipids, but not lipoproteins, are excellent antigens.

Bibliography

Abbas AK, Lichtman AH, Pillai S. *Cellular and Molecular Immunology*. 8th ed. Philadelphia: Saunders; 2014.

Brubaker SW, Bonham KS, Zanoni I, Kagan JC. Innate immune pattern recognition: a cell biological perspective. *Ann Rev Immunol*. 2015;33:257–290.

Costa LR, Barrington GM. Immunologic disorders. In: Smith BP, eds. *Large Animal Internal Medicine*. 5th ed. St Louis: Mosby Elsevier; 2015.

Delves PJ, Martin SJ, Burton DR, Roitt IM. *Roitt's Essential Immunology*. 13th ed. Malden, MA: Wiley-Blackwell; 2017.

Janeway C. *Immunobiology: The Immune System in Health and Disease*. 8th ed. New York: Garland Science; 2011.

Jorch SK, Kubes P. An emerging role for neutrophil extracellular traps in noninfectious disease. *Nat Med*. 2017;23(3):279–287.

Kumar H, Kawai T, Akira S. Pathogen recognition by the innate immune system. *Int Rev Immunol*. 2011;30(1):16–34.

Nagahata H. Bovine leukocyte adhesion deficiency (BLAD): a review. *J Vet Med Sci*. 2004;66(12):1475–1482.

Parham P. *The Immune System*. 4th ed. New York: Garland Science; 2014.

Rao CV. *Immunology: A Textbook*. 3rd ed. Harrow, UK: Alpha Science; 2017.

Tizzard IR. *Veterinary Immunology: An Introduction*. 9th ed. Philadelphia: Saunders; 2013.

55

The Specific Immune Response: Acquired Immunity

S. ANSAR AHMED AND GERHARDT G. SCHURIG;
CONTRIBUTIONS BY XIN M. LUO

KEY POINTS

T Cells (T Lymphocytes)

1. Mature T cells develop from lymphoid stem cells that have migrated to the thymus.
2. T cells are a heterogeneous population of cytotoxic T cells and T-helper (Th) cells.

Interactions of Antigen-Presenting Cells and T Cells

1. Major histocompatibility complex (MHC) proteins are considered the central regulators of the immune system.
2. MHC class I antigens of infected nucleated cells play a major role in activating cytotoxic T cells.
3. MHC class II antigens on antigen-presenting cells play a major role in selective activation of Th cells.

Antibodies

1. Initial exposure to foreign antigen induces slow onset of antibody appearance, whereas subsequent exposure induces faster, longer-lasting antibody appearance.
2. Antibodies, or immunoglobulins, are glycoprotein molecules that can be divided into five isotypes, or classes.
3. The B-cell population produces antibodies to millions of different antigens, yet the antibody-antigen interaction is specific.
4. Cytokines promote the differentiation of B cells.
5. Expansion of an antigen-specific B–memory cell population on initial antigen exposure results in a faster, more pervasive, secondary immune response.

Regulation of Immune Responses

1. The actions, secretions, and surface molecule expression of immune cells play an important role in regulation of the body's immune response.

As discussed in Chapter 54, innate immunity offers effective defense against a wide range of pathogens. Key features of innate immunity include (1) rapid response against invading pathogens; (2) nonspecificity; and (3) physical, chemical, and cellular (phagocytic cells, natural killer [NK] cells) barriers. The response of the innate immune system, however, is not long-lasting and does not induce immunological memory (i.e., ability to recall previous exposure to antigens and respond to these effectively and specifically). For long-lasting immunity, another arm of the immune system must be activated. This is referred to as *acquired immunity,* which involves activation of T and B lymphocytes. *Antigen-presenting cells* (APCs), a part of the innate immune system, play a central role in activating lymphocytes. Activated *T lymphocytes* (T cells) secrete cytokines essential for defense against intracellular pathogens, activation of other cells, and coordination of immune responses. *B lymphocytes* (B cells) have two main functions: (1) secreting antibodies that bind specifically to the antigen that induced the antibody response and (2) acting as APCs.

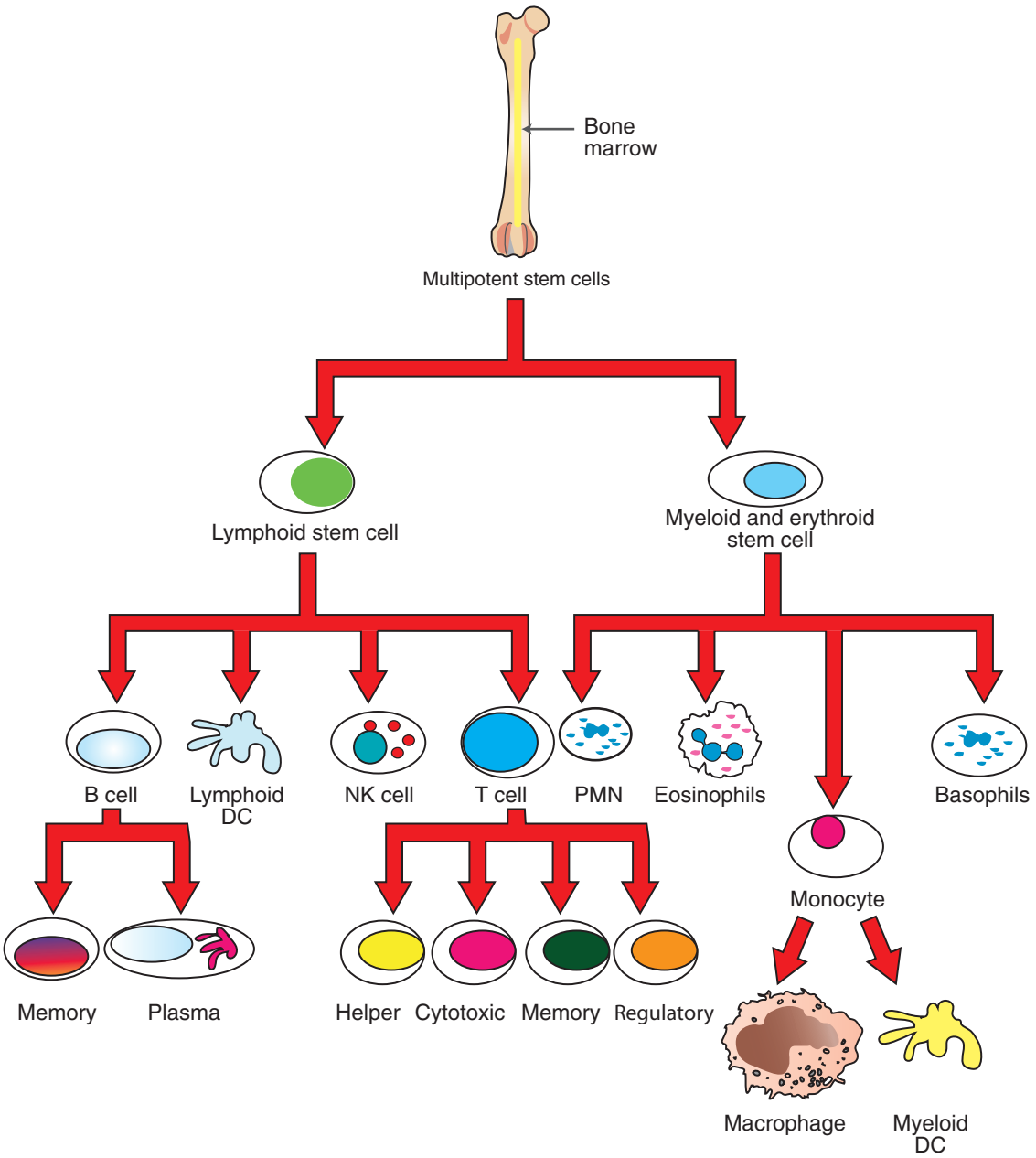
Before discussing how antigens are presented to specific lymphocytes, it is important to understand the different types of immune cells (Fig. 55.1). All cells of the immune system are derived from multipotent stem cells located primarily in the marrow of long bones. These multipotent stem cells subsequently give rise to primordial stem cells, such as lymphoid stem cells or myeloid stem cells. Myeloid stem cells give rise to monocytes, which mature in tissues to become macrophages or dendritic cells. Lymphoid stem cells give rise to T, B, NK, and lymphoid dendritic cells. Mature cells are found circulating throughout the body but concentrate in the peripheral lymphoid organs (e.g., lymph nodes, spleen) and gut-associated lymphoid tissues, where most of the complex interactions with antigens take place.

Birds, unlike mammals, have a unique lymphoid organ called the *bursa of Fabricius* where B cells develop. This round, sac-shaped organ is located above the cloaca. Analogous to the thymus, the bursa consists of lymphocytes embedded in epithelial tissues. Mammals have no precise lymphoid organ that is equivalent to this bursa. Bone marrow and ileal Peyer's patches are thought to be the principal mammalian organs where B cells develop.

T Cells (T Lymphocytes)

Mature T Cells Develop From Lymphoid Stem Cells That Have Migrated to the Thymus

Lymphoid stem cells destined to become T cells migrate to the thymus and are referred to as *thymocytes.* (The thymus extends approximately from the base of the trachea to the front of the heart.) The most recent immigrants from bone marrow arrive at the cortex of the thymus and lack important cell surface markers, such as T-cell receptors (TCRs), CD4, and CD8 markers, which are essential for T-cell activation. These immature thymocytes undergo a highly complex and tightly regulated development and



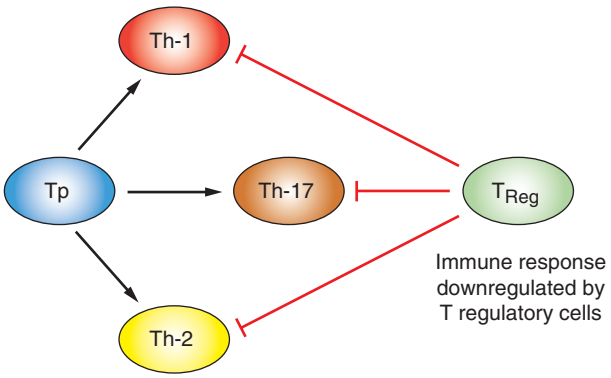
• **Fig. 55.1** Lymphopoiesis: development of various types of lymphocytes. *DC,* Dendritic cell; *NK,* natural killer; *PMN,* polymorphonuclear neutrophil leukocytes.

maturation process into mature T cells. During development, the cells begin to acquire both CD4 and CD8 surface markers (double positive) and TCRs. As the cells further mature, they lose either CD4 or CD8 markers. $CD4^+/CD8^+$ cells that lose the CD8 marker become $CD4^+/CD8^-$ cells and are known as T-helper (Th) cells, whereas those double-positive cells that lose the CD4 marker become $CD4^-/CD8^+$ cells, or cytotoxic T cells.

The selection for survival of T cells during this developmental process is extremely stringent and discriminating. During development, the thymocytes learn two important lessons: (1) T cells respond only to foreign antigens (positive selection), and (2) the cells will not respond to "self" antigens (negative selection). Learning these two critical lessons is essential for the survival of the organism. Therefore any developing thymocytes that deviate from learning these two key lessons are terminated by apoptosis (negative selection). Consequently, greater than 90% of developing thymocytes die within the thymus. Cells marked for intrathymic death include those cells that are defective (i.e., cannot bind to antigens or have truncated receptors) or autoreactive (bind strongly to "self" peptides). Thus only competent, positively selected T cells ($CD4^+$ or $CD8^+$) are allowed to emigrate out of the thymus as T cells.

T Cells Are a Heterogeneous Population of Cytotoxic T Cells and T-Helper Cells

All T cells express a T-cell antigen receptor (TCR), CD28 and related molecules, and either CD4 (helper cells) or CD8 (cytotoxic cells). TCR specifically binds to antigenic peptides presented by APCs. Based on discrete functions of T cells, these cells are subdivided into two major types: (1) helper cells and (2) cytotoxic cells. Th cells secrete proteins called *cytokines* that act on other immune cells to provide help and coordinate immune responses.



• **Fig. 55.2** T precursor cells (Tp) differentiate into T-helper (Th)-1, Th-2, or Th-17 cells based on the cytokine signals. Generation of T regulatory (Treg) cells will inhibit all three types of cells to downregulate immune responses.

The cells express the CD4 receptor. These cells express the CD8 molecule (but not CD4) and have granules rich in serine esterase granzymes. Cytotoxic T cells also have perforins and lymphotoxins that are important in initiating cytotoxicity and killing infected and abnormal cells.

Th cells, based on the predominant cytokines secreted, are further divided into three major types: Th-1, Th-2, and Th-17. Th-1 cells predominantly secrete interleukin-2 (IL-2), interferon-gamma (IFN-γ), and tumor necrosis factor beta (TNF-β). Th-1 immunity is critical for defense against intracellular pathogens (viral, bacterial, or protozoal) and certain types of tumors. Th-1 cells are preferentially generated when naïve $CD4^+$ cells are exposed to IL-12, a cytokine from APCs (Fig. 55.2). Failure to generate Th-1 cells creates susceptibility to these infections. Abnormal activation of Th-1 cells can result in a wide variety of inflammatory conditions, including autoimmune states.

Activation of naïve $CD4^+$ cells with interleukin-4 (IL-4) leads to differentiation into Th-2 cells (see Fig. 55.2). Th-2 cells predominantly secrete IL-4, interleukin-6 (IL-6), interleukin-5 (IL-5), and interleukin-10 (IL-10). Generation of Th-2 cells is essential for defense against extracellular pathogens, neutralization of toxins and viruses in body fluids, and activation of other cells of the immune system. Abnormal regulation of Th-2 cells leads to allergies.

Activation of CD4 Th cells with IL-6 and transforming growth factor beta (TGF-β) induces the differentiation of Th-17 cells, which secrete a powerful proinflammatory cytokine, interleukin-17 (IL-17; see Fig. 55.2). This cytokine is now recognized as an important mediator of inflammatory and autoimmune diseases. IL-17 acts on target cells to activate key signaling molecules to promote inflammation through several mechanisms including (1) recruiting inflammatory cells (e.g., neutrophils, monocytes, and macrophages) to the site of inflammation; (2) acting on target cells (e.g., fibroblasts, epithelial cells) to stimulate a broad range of strong proinflammatory molecules (e.g., IL-6, monocyte chemotactic protein 1, nitric oxide); and (3) synergizing with TLR ligands. Although IL-17 is protective in infection, overproduction of IL-17 is known to aggravate certain disease conditions (e.g., autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis).

Exposure of CD4 naïve cells to TGF-β alone (e.g., in the absence of IL-6) will drive the differentiation to T regulatory cells (Treg cells; see Fig. 55.2). Treg cells are a population of T cells that act as powerful suppressors of the T cell-mediated immune response and of self-reactive T cells in autoimmune diseases. Tregs utilize a broad range of suppressive mechanisms that include release of immunosuppressive cytokines TGF-β and IL-10, and cell-cell contact. IL-10 and TGF-β secreted by Tregs are critical for dampening immune responses in allergy, burns, pregnancy, cancer, viral diseases, and autoimmune diseases. Treg cells can inhibit and downregulate all three (Th-1, Th-2, and Th-17) subsets of CD4 cells (see Fig. 55.2). Dysregulation of Treg cells can lead to massive inflammatory diseases, whereas excessive numbers or functions of Tregs can lead to dampening of immune responses that lead to severe infections. Therefore physiologically, Treg cells must be finely balanced to maintain immune health status.

Thus it is clear that the immune system must initiate the correct type of immune response to maintain homeostasis and defend the host appropriately against the invasion of different types of pathogens.

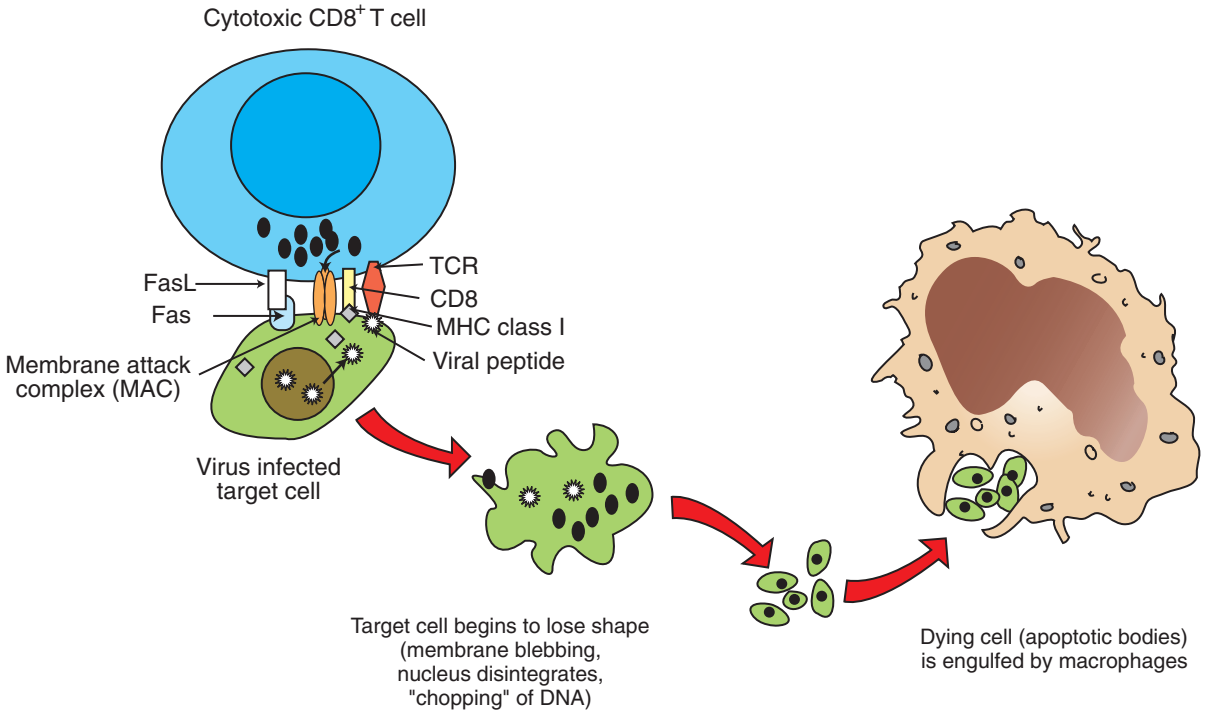
Interactions of Antigen-Presenting Cells and T Cells

Major Histocompatibility Complex Proteins Are Considered the Central Regulators of the Immune System

Activation of specific T cells is highly dependent on interactions with major histocompatibility complex (MHC) proteins, which have a unique ability to bind to processed antigenic peptides. Therefore MHC proteins are considered the central regulators of the immune system. MHC proteins are encoded by a number of genes that are clustered together on a chromosome and referred to as the *MHC locus*. The MHC gene complex is inherited as a block of genes and is known to encode three categories of proteins or antigens: class I, class II, and class III antigens. The number of genes that encode class I antigens varies from species to species, with a large number in humans (>30) at one end of the spectrum to a limited number of genes in pigs, turkeys, and cheetahs at the other end. In general, all nucleated cells express class I antigen, which is a single α-chain peptide of approximately 45 kilodaltons (kDa) linked to β2-microglobulin (a non-MHC protein thought to be essential for proper folding and stabilization of the α chain). Class I antigens can bind to peptides (e.g., viral peptides) and serve as receptors for CD8 molecules on cytotoxic T cells. Class I antigens have a high rate of mutation but no recombination. These mutations allow class I antigens to alter their ability to bind to endogenous, processed antigenic peptides.

MHC Class I Antigens of Infected Nucleated Cells Play a Major Role in Activating Cytotoxic T Cells

Cytotoxic killing of intracellularly infected, cancer, or autoreactive cells is an essential step in survival by containing infected cells or the spread of deleterious cells. For example, a viral infection of any cell in the body yields to viral replication within the cell, and some of these viral peptides will physically bind to intracellular MHC class I antigens (Fig. 55.3). This viral peptide–MHC class I complex is carried to the surface and displayed as an altered MHC class I molecule. TCR molecules of effector $CD8^+$ cytotoxic T cells will recognize the class I molecule–peptide complex to initiate cytotoxicity by at least four different, but complementary, mechanisms. First, contact of a cytotoxic $CD8^+$ cell with an infected cell displaying a MHC class I–peptide complex will immediately result in cytoplasmic reorganization within the $CD8^+$ cell. This



• **Fig. 55.3** Mechanism of CD8-mediated cytotoxicity. *FasL,* Fas ligand; *MHC,* major histocompatibility complex; *TCR,* T-cell receptor.

includes the alignment of granules and Golgi apparatus at the site of contact. Perforins in the cytotoxic cells polymerize to form tiny injectable tubes referred to as membrane attack complexes (MACs) that "drill" holes into the target cells. Granzymes are passed from the cytotoxic cells into the target cells through these perforin tubes to initiate apoptosis.

The other three mechanisms by which $CD8^+$ cells induce apoptosis of target cells are (1) the secretion of lymphotoxin α (tumor necrosis factor alpha, TNF-α), which binds to its specific receptor on the target cells to initiate apoptosis; (2) the interactions of CD95 ligand on T cells with the CD95 "death" receptor on target cells; and (3) secretion of granulysin, an antibacterial peptide found in granules that activates lipid-degrading enzymes (sphingomyelinases). This in turn results in an increase in saponins, including ceramide, which increase apoptosis. Granulysin kills not only infected target cells, but also bacteria, thus containing the infection.

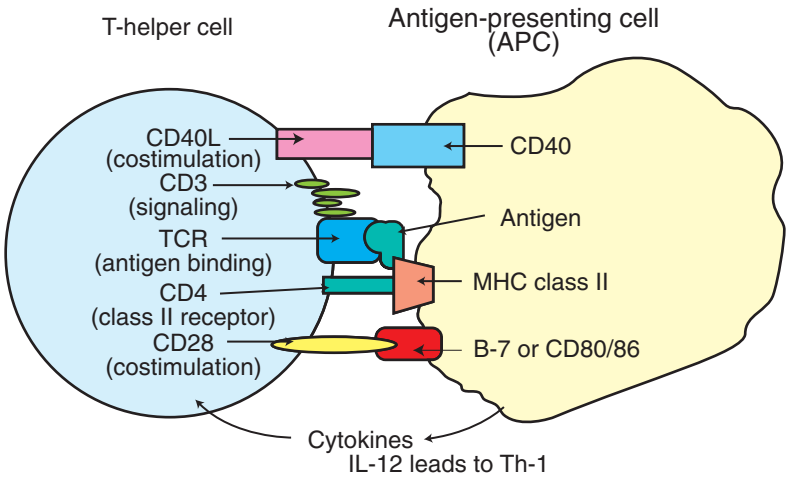
MHC Class II Antigens on Antigen-Presenting Cells Play a Major Role in Selective Activation of T-Helper Cells

The expression of cell surface MHC class II antigens is highly restricted. They are only present on select types of cells including dendritic cells, select macrophages, B cells, and keratinocytes. The presence of class II antigens on their surfaces endows these cells with a unique ability to present antigens to $CD4^+$ Th cells (Fig. 55.4). Therefore these cells are called professional APCs. Class II antigens are two-chain molecules composed of a glycoprotein, 33-kDa α chain and a shorter, 27-kDa β chain that form a groove onto which processed (exogenous) antigenic peptides bind. For example, when a macrophage phagocytoses an antigen and breaks it down into peptides within a vacuole, intracellular MHC class II antigens bind to these processed peptides, and this complex moves to the surface of the cell to be presented to $CD4^+$ T cells. The processed antigenic peptide specifically binds to the TCR on T cells, and the class II MHC proteins (on APCs) specifically interact with the CD4 molecule on T cells. These interactions are the first steps in activation of Th cells. Activation of T cells is highly regulated because their inadvertent activation has profound and widespread consequences; the cytokines secreted by activated Th cells can affect a wide range of both lymphoid and nonlymphoid cells.

Activation of $CD4^+$ T cells requires at least two signals for activation. The primary activation signal determines the specificity through interactions of antigenic peptides and MHC molecules on APCs with the TCR/CD3 complex on T cells. The second signal is referred to as a *co-stimulatory signal.* Co-stimulatory signals include interaction of CD28 and/or CD40L, both residing on T cells, with CD80/86 and/or CD40, both residing on APCs. Cytokines released from APCs, such as interleukin-12 (IL-12), interleukin-18 (IL-18), and interleukin-27 (IL-27), promote the generation of a Th-1 subset of cells.

Activation of T cells is strictly controlled, with two major restrictions. First, T cells cannot recognize free antigens; rather, they recognize short peptides that are a product of processed antigens by the APCs. Second, the processed antigen must be physically associated with the MHC molecules. This results in molecular interactions of antigenic peptide bonding to MHC class I or class II molecules on cells or APCs, with the TCR and CD8, or TCR and CD4 on T cells, respectively.

As mentioned earlier, T cells will specifically interact with antigenic peptides through specific recognition by the TCR on T cells, analogous to the B-cell receptor (BCR) on B cells. The TCR on T cells belongs to the immunoglobulin superfamily and thus has variable and constant regions along with transmembrane and cytosolic domains. Similar to B cells, the variable portion of the TCR chain determines the specific binding to the antigenic peptide.



• **Fig. 55.4** Interacting molecules of T-helper cells and antigen-presenting cells. *IL,* Interleukin; *MHC,* major histocompatibility complex; *TCR,* T-cell receptor; *Th-1,* T-helper cell type 1; *Th-2,* T-helper cell type 2.

Because more antigens exist in the universe than the actual number of T cells, these cells have adapted a variety of molecular mechanisms to interact specifically with an innumerable number of antigens. These mechanisms include recombination of TCR genes (similar to the BCR), unequal sister-chromatid exchange, and nucleotide insertion at various locations of the variable gene segments. An important difference between the TCR and BCR is that the TCR does not undergo somatic mutation. If the TCR could undergo somatic mutation, there would be an increased chance of inadvertently generating a TCR reactive against "self" antigens, resulting in devastating autoimmune conditions. This attribute is critical for survival because T cells, unlike B cells, can affect a large number of diverse lymphoid and nonlymphoid cells through secretion of cytokines.

Antibodies

Initial Exposure to Foreign Antigen Induces Slow Onset of Antibody Appearance, Whereas Subsequent Exposure Induces Faster, Longer-Lasting Antibody Appearance

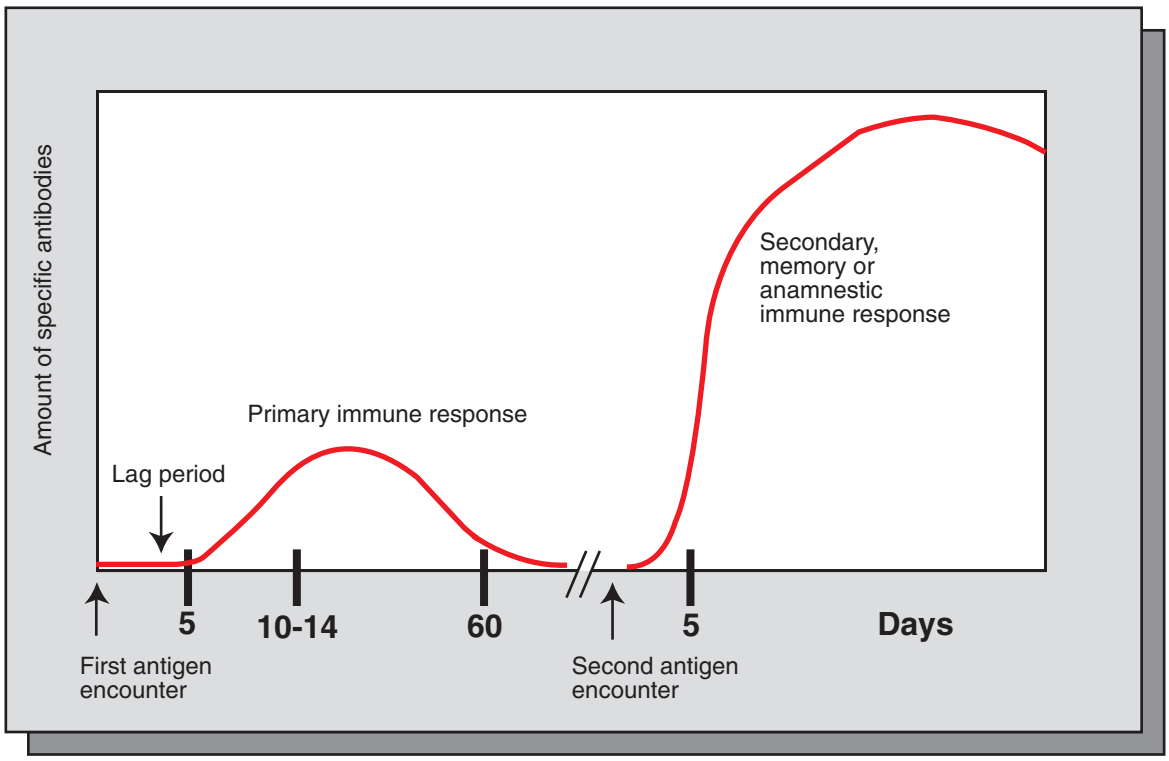
The exposure of an animal to a foreign antigen usually elicits a specific immune response. This response may involve the production of (1) specifically reactive T cells or (2) antibodies able to bind specifically with the foreign antigen. Typically, if an animal is exposed to a particular foreign antigen for the first time, no antibodies specific for that antigen will be detected in blood or secretions for several days. This "lag" period can last for up to 1 week, at which time antibodies capable of binding to the antigen appear in circulation and start increasing in quantity for the next 2 or 3 weeks. After that time, antibody quantities plateau and eventually decrease until they essentially disappear. The amount of antibodies produced and the duration of the response depend greatly on the nature of the antigen, quantity and route of exposure, and whether the antigen is given in combination with immune enhancers (adjuvants).

The type of response obtained after a first exposure to a specific antigen is called the primary immune response (Fig. 55.5). If the animal is reexposed to the antigen, the lag period is very short, much higher levels of specific antibodies are obtained, and the response usually lasts for a significantly longer period. This response to a second exposure to an antigen is called a secondary immune response or anamnestic (memory) immune response.

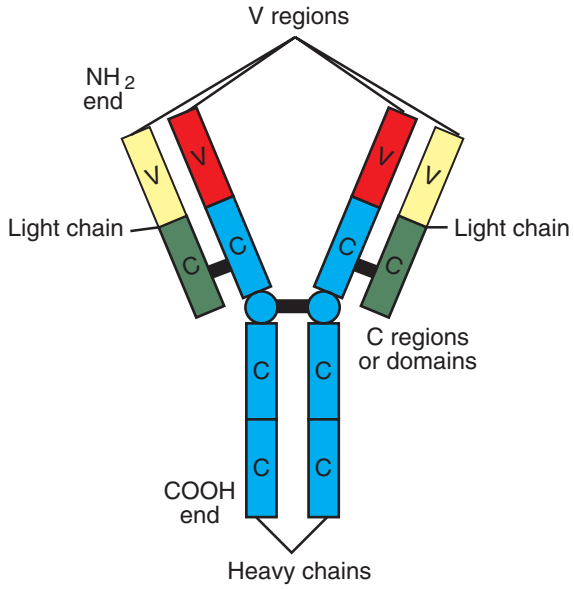
Antibodies, or Immunoglobulins, Are Glycoprotein Molecules That Can Be Divided Into Five Isotypes, or Classes

Antibodies are glycoprotein molecules that are the products of B lymphocytes. Antibodies, also called immunoglobulins, are basically made of four glycoprotein molecules. They are found on the surface of B cells, where they serve as antigen receptors (BCRs), or free in blood and secretions after being secreted by B cells. These free, or soluble, antibodies can neutralize antigens and assist in their removal. The basic structure of an antibody molecule has two identical short glycoprotein chains called light (L) chains and two identical longer chains called heavy (H) chains held together by disulfide bonds (Fig. 55.6). The L chains are made of two halves, or domains; the half located at the carboxyl end of the chain is called the constant part of the L chain (C_L), and the half located at the amino end is called the variable part (V_L). The H chain is made of one variable domain (V_H) and usually three constant domains (CH_1, CH_2, CH_3). The amino terminal ends of the L chain (V_L) and the H chain (V_H) come together to form an antigen-binding or combining site (Fig. 55.7). Therefore two identical antigen-combining sites exist per basic immunoglobulin molecule. The carboxyl end of the two H chains form the Fc portion of the molecules; this end is the portion able to bind to Fc receptors on specialized cells and is the part of the molecule attached to the membrane of B cells when the immunoglobulin serves as the antigen receptor (BCR) for the cell.

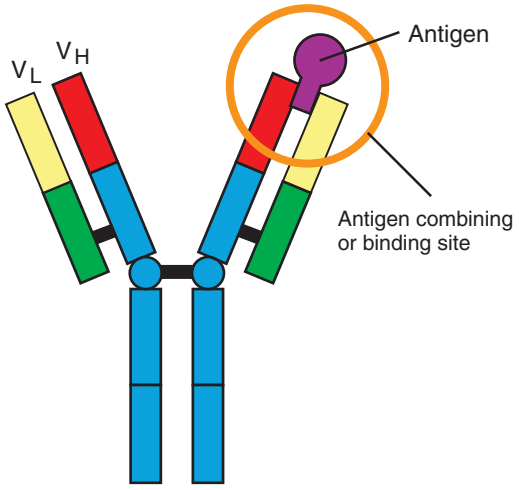
Depending on molecular weight and other characteristics, immunoglobulins can be divided into classes, or isotypes. Basically, there are five isotypes: IgM, IgG, IgA, IgE, and IgD. Soluble immunoglobulin M (IgM) consists of five basic antibody molecules that bind together by disulfide bonds and an additional short protein chain to form a pentamer. Therefore one IgM molecule has 10 identical antigen-combining sites. Its molecular weight is about 900 kDa. In primary immune responses, IgM is the predominant immunoglobulin. Because of its large size, IgM is rarely



• **Fig. 55.5** General dynamics and characteristics of the primary and secondary antibody responses.



• **Fig. 55.6** Basic structure of an antibody molecule. *Black bar,* Disulfide bond; *blue circle,* heavy chain hinge; *C,* constant region; *V,* variable region.



• **Fig. 55.7** Antigen-antibody combining site. V_H, Variable region of heavy chain; V_L, variable region of light chain.

found in body fluids other than blood. The BCR form of IgM is a monomer of 180 kDa.

Immunoglobulin G (IgG) has the structure of the basic antibody molecule (monomer) previously described, and its molecular weight is 180 kDa. IgG has two antigen-combining sites and is the predominant immunoglobulin detected in the secondary immune response. It is able to move out of the circulatory system and appears in body fluids and also in secretions.

Immunoglobulin A (IgA) is found in small amounts in circulation as a monomer and in much larger amounts in secretions, where it is found as the predominant immunoglobulin and primarily as a dimer. It is produced by plasma cells (mature B cells) located under body surfaces such as skin, mammary glands, and the intestinal, respiratory, genital, and urinary tracts. IgA is found in secretions and has an attached secretory molecule that protects IgA from intestinal proteases. Secretory IgA is the main immunoglobulin found on mucosal surfaces and has four antigen-combining sites in its dimeric structure. Its main role is to prevent antigen from attaching to these surfaces. Thus IgA blocks penetration of antigen into the body. IgA responses are mainly elicited if antigen exposure is through contact with mucosal surfaces, such as the upper respiratory and intestinal tracts. IgM and IgG responses are

elicited through parenteral contact with antigen (intradermal, subcutaneous, intramuscular, and systemic routes).

Immunoglobulin E (IgE) is a monomer, and its H chain contains four constant domains in addition to the variable domain. It is found in very low levels systemically, and most IgE is bound to basophils and mast cells (inflammatory and allergic reaction mediators) through its Fc portion. IgE is able to bind antigen while attached to these cells, thereby eliciting allergic reactions.

Immunoglobulin D (IgD) is a monomer and has only two constant domains on its H chain. IgD is primarily bound to the membrane of B cells and is secreted in negligible amounts in the serum. Negligible IgD is secreted.

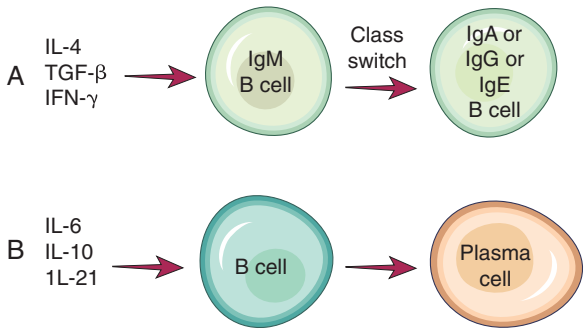
The B-Cell Population Produces Antibodies to Millions of Different Antigens, yet the Antibody-Antigen Interaction Is Specific

Antibodies bind to antigen through their antigen-binding sites. Each antigen-binding site is formed by the steric interaction of the V_L and V_H domains, which come in close contact because of the three-dimensional folding of the glycoprotein chains on which they reside. This steric interaction essentially forms a cleft, and any antigenic structure that fits into this cleft is recognized and binds to the antigen-binding site. Therefore if an antigen can bind to the antigen-combining site on the BCR, the B cell is eventually (after a complex set of signal interactions) triggered to replicate (clonally expand), giving rise to many "identical" B cells, which eventually produce and secrete the soluble immunoglobulin specific for that antigen.

It is important to stress that antibody-antigen responses are specific. Antibodies produced after exposure to an antigen will only bind to that antigen or to other antigens structurally similar to the original antigen (cross-reaction). The question is, "how does the immune system manage to respond to literally several million different foreign antigens in a specific way?" Three facts are crucial to understand this situation. First, the BCR is an immunoglobulin, and the specificity of the immunoglobulin secreted by a particular B cell is the same as the specificity of its BCR. Second, an individual B cell can only have BCRs of identical specificity. Third, essentially each B cell (and there are millions) in the body has a BCR with a different antigen-combining site, because B cells undergo random genetic mutations in the genes coding for their V_L and V_H domains during their early development (ontogeny). Because these domains make up the antigen-combining site, a single amino acid change in either of the regions changes the steric interaction of these domains, giving rise to different "clefts" with different antigen-combining abilities.

Cytokines Promote the Differentiation of B Cells

The activation of B cells requires several signals that include (1) antigen cross-linkage of surface IgM, or BCRs; (2) interaction between CD40 on B cells and CD40 ligand on Th cells; and (3) expression of cytokine receptors on B cells that recognize cytokines produced by Th cells. While B cells can proliferate with the first two signals, they would not differentiate into antibody-producing cells unless the third signal is present. B-cell differentiation involves three processes: class switching (changing of isotype), affinity maturation (selective perpetuation of cells with increased antigen affinity), and formation of plasma cells (often short-lived, high rate antibody secreting) and memory cells (longer surviving clones of antigen-exposed cells). All these processes require the functions



• **Fig. 55.8** Cytokines promote the differentiation of B cells in class switching (A) and plasma cell formation (B). *IFN,* Interferon; Ig, immunoglobulin; *IL,* interleukin; *TGF,* transforming growth factor.

of Th cytokines. IL-4, for example, stimulates the proliferation and affinity maturation of B cells as well as the differentiation of and class switching to IgG1- or IgE-producing B cells in mice. TGF-β, on the other hand, promotes the differentiation of and switching to IgA- or IgG2b-producing B cells. IFN-γ stimulates the differentiation of and switching to IgG2a- or IgG3-producing B cells in mice. Finally, the formation of plasma cells requires the functions of IL-6 (mouse), IL-10 (human), and IL-21 (mouse and human) (Fig. 55.8).

Expansion of an Antigen-Specific B–Memory Cell Population on Initial Antigen Exposure Results in a Faster, More Pervasive Secondary Immune Response

As the host is exposed to a foreign antigen, all existing B cells that have a BCR capable of binding the foreign antigen are able to react to such an antigen. Initially, among the millions of different B cells, the BCR of a few will bind to the antigen. This binding allows those matching B cells to multiply, creating many more B cells with the same BCR. This rather rapid expansion of antigen-specific B cells is often referred to as specific clonal expansion. The newly generated B cells start producing specific antibodies and secreting them. Specific antibodies begin appearing in circulation, and as clonal expansion increases, the quantity of specific antibodies in circulation increases. All B cells start their specific antibody production by producing IgM isotypes. Given the right milieu (e.g., appropriate sequence of signals), B cells stop producing IgM and switch production to IgG or another isotype (IgA or IgE). For example, IL-4 secreted by Th-2 cells will induce the switch from IgM to IgE. On the other hand, Treg cells can block the production of cytokines from Th-1 and Th-2 cells, which could lead to the decrease of specific antibody isotype switching in B cells. It is important to understand, however, that the antigenic specificity of the antibody does not change. These dynamics are consistent with those observed in a primary and secondary immune response.

As the host encounters an antigen for the first time, few of the existing B cells are able to recognize it, but once the antigen is recognized by those few B cells, they undergo clonal expansion. Therefore antibody production starts increasing, shows up in circulation, and IgM predominates during the primary immune response. The long "lag" period represents the time needed to process significant numbers of antibody-producing B cells through clonal expansion. Some B cells start switching to IgG production, which

is why some IgG appears during the later parts of a primary immune response. As antigen is neutralized by the antibodies, B-cell stimulation stops, antibody production declines until it eventually stops, and many B cells within the expanded population become long-lived memory B cells.

If the host confronts the same antigen for a second time, the antigen is recognized by the vastly expanded, antigen-specific B–memory cell population, many cells of which have switched their antibody production capability to an isotype other than IgM, mainly IgG. This expanded cell population starts antibody production quickly because of the larger population of antibody-producing cells and greater likelihood of being in an IgG-producing state. The lag period is therefore very short, and a large amount of IgG is produced, typical of a secondary immune response. Eventually, as in a primary immune response, clonal expansion stops, and antibody production decreases and eventually stops. Future confrontations with the same antigen will lead to secondary immune responses, which are characterized by short "lag periods" and high sustained production of IgG or other isotypes (IgA or IgE).

Antigen stimulation of the BCR is not a sufficient signal to initiate B-cell clonal expansion and antibody production. Many other signals must reach the B cell after its BCR has recognized antigen. These signals, often represented by interleukins, come from Th cells that have recognized the antigen through complex interaction mechanisms with APCs (see earlier discussion). Interestingly, B cells that have bound antigen through their BCR and have internalized the BCR-bound antigen are also able to interact with Th cells, helping them to recognize antigen. A B cell–T cell interaction is also necessary for B cells to switch their production from IgM to other isotypes. The B cells must interact with interleukins from Th cells. The timing and nature of the interleukins reaching the B cells play an important role in deciding to which isotype the B cell is switching.

Regulation of Immune Responses

The Actions, Secretions, and Surface Molecule Expression of Immune Cells Play an Important Role in Regulation of the Body's Immune Response

When the antigen is cleared by antibodies or T cells, it is imperative that the immune response return to its normal state to maintain homeostasis. Failure to downregulate the heightened immune reactivity will likely result in a number of pathological conditions, including autoimmunity, lymphoid tumors, allergies, amyloidosis, and abortion. The body has multiple mechanisms to downregulate the immune system. Immune cells themselves secrete various biomolecules that downregulate immune activity. These include prostaglandin E_1 (PGE_1), which increases cyclic adenosine monophosphate (cAMP) to suppress physiological activity of cells. As previously noted, immune cells also secrete cytokines, such as TGF-β and IL-10, which downregulate immune responses. The importance of these cytokines in downregulating immune responses is evident in the immune-mediated inflammation that results in mice with a deleted TGF-β gene. Antibodies themselves can downregulate their immune responses by binding to Fc receptors on B cells. This cross-ligation of Fc receptors results in delivery of inhibitory signals. Immune responses are antigen-driven, so clearing of antigens by immune mechanisms leads to a decrease in antigenic load and thus diminishes the antigen-induced activation of lymphocytes.

As mentioned earlier, activation of T cells requires at least two signals. The second signal is altered by the increased secretion of molecules or the T cells that deliver a negative signal to dampen T-cell activation. Important in this regard are the aforementioned, recently discovered Treg cells, a small percentage of T cells (<5%) that are $CD4^+/CD25^+/FoxP3^+$ and that have powerful downregulatory effects. The importance of these cells in downregulation of immune responses is evident because defects in these cells lead to widespread autoimmune diseases in many experimental models. Conversely, administration of Treg cells leads to prevention of inflammatory attacks. These cells hold promise in understanding the biology of immunoregulation and in therapy of various inflammatory diseases. There is a growing body of knowledge on these subsets of T cells in domestic animals.

CLINICAL CORRELATIONS

Unthrifty Foal

History

You are presented with a 4-week-old Arabian filly for depression, coughing, and nasal discharge. The foal was born with no apparent problems. Immunoglobulin levels were checked and were normal. However, the foal has been unhealthy and seems to be repeatedly sick with skin and respiratory infections.

Clinical Examination

The foal has a temperature of 38.8° C (102° F) (elevated), heart rate of 60 beats/min (elevated), and a respiratory rate of 48 breaths/min (elevated). On auscultation, the foal has both crackles and wheezes (abnormal lung sounds). The capillary refill time (CRT) is prolonged, and the mucous membranes are darker pink than normal. The foal also has some abrasions and cellulitis in these areas on the skin.

Comment

With the increased temperature, abnormal lung sounds, poor perfusion (prolonged CRT, darker-than-normal mucous membranes), and increased respiratory rate, this foal probably has a respiratory infection, most likely bacterial in origin. Additionally, based on the skin abrasions, the filly probably has a skin infection, which is uncommon in normal, healthy foals. Based on the age, breed, and recurrent infections, this foal likely has combined immunodeficiency (CID), a genetic autosomal recessive disorder. Foals with CID have a defect in a DNA-dependent protein kinase. This results in inability to produce mature B and T cells. As a result, the foal's immune responses

CLINICAL CORRELATIONS—cont'd

are limited. A normally functioning innate response is present, composed of neutrophils, macrophages, dendritic cells, and NK cells, but a normal adaptive immune response is lacking. The B cells do not make antibodies, and insufficient T cells are present to produce cytokines to help with the immune response and control infection. This foal has evidence of recurrent infections based on the history and appears to have concurrent skin and respiratory infections on presentation.

Foals with this type of history, including those with low lymphocyte counts and low immunoglobulin levels, are suspect candidates for CID. However, a definitive diagnosis is based on genetic testing and necropsy. Typically, foals with CID have rudimentary thymuses and lymph nodes because of the lack of lymphocytes and germinal centers. The spleen is usually reduced in size as well.

Treatment

Because CID is a genetic disorder that the filly cannot overcome, the prognosis is poor. In many cases, until a definitive diagnosis is made, these foals are treated with antibiotics for the infection and with supportive care. When a definitive diagnosis is made based on genetic testing, however, the foals are usually euthanized because of the poor long-term prognosis for life.

Swelling Under Jaw in 20-Year-Old Rocky Mountain Horse Gelding

History

A client calls you because they noticed that their older horse had a swelling under his jaw. At first they thought it might just be from a mosquito bite or sting, but now they have noticed it has gotten larger in the last week or two. The horse does not have any other symptoms. He has been in good health and is current on vaccinations, deworming, and Coggins testing (a test for equine infectious anemia).

Clinical Examination

The horse's temperature, pulse, and respiratory rate are normal. The only thing remarkable on your physical examination is the swelling underneath his jaw, which is caudal to his submandibular lymph nodes. It is firm, fairly homogenous-feeling, nonpainful, and about 4 cm × 2 cm × 2.5 cm. You also notice that he has many deer ticks (*Ixodes scapularis*) on his abdomen. The owner comments that she has been having a really difficult time keeping the ticks off. Routine blood work demonstrates increased white blood cell count, including lymphocytes, increased fibrinogen, and increased globulins supporting an inflammatory response. You recommend biopsying the mass and running Lyme multiplex assay for antibodies.

Comment

The mass comes back as many lymphocytes and histiocytes present, which is nonspecific for some sort of inflammatory process, potentially due to an infection. He does not have any other masses present. The Lyme multiplex assay measures the differing antibody response to shifts in antigen expression on the bacterial spirochete that causes Lyme disease, *Borrelia burgdorferi*. over the course of infection. The Lyme test comes back as high antibodies to OspF (outer surface protein F), which is more consistent with chronic infection. DNA profiling using polymerase chain reaction (PCR) of the mass is positive for *B. burgdorferi.* In rare cases in *B. burgdorferi*-infected horses, they can develop pseudolymphoma, which is a benign lymphoproliferative process that mimics true lymphoma. Therefore identifying the cause is critical. In the few cases reported, horses respond well to treatment. The inflammatory response is thought to be associated with infection, in this case *B. burgdorferi,* but it is not known why some of these horses develop pseudolymphoma.

Treatment

The horse is treated with minocycline, a broad-spectrum antibiotic, for 28 days. At that time, the mass had resolved. No additional treatment was recommended.

Practice Questions

1. On their surface, T cells have:
- **a.** T-cell antigen receptor
- **b.** CD3 molecules
- **c.** CD4 or CD8
- **d.** CD28 molecules
- **e.** All the above
- **f.** Both a and b

2. Which of the following statements is correct with regard to $CD4^+$ T cells?
- **a.** They are regarded as a suppressor/cytotoxic type of T cell.
- **b.** CD4 molecule on T cells is a receptor for class I molecules on antigen-presenting cells.
- **c.** They provide help to T, B, and NK cells and to macrophages by secreting critical cytokines such as IL-2.
- **d.** CD4 molecules recognize CD8.
- **e.** CD4 molecules bind to class I MHC.

3. Which of the following is correct concerning cytotoxic cells?
- **a.** Viral-infected target cells undergo necrotic cell death by cytotoxic cells.
- **b.** Cytotoxic cells release granzymes (through holes "drilled" by perforin) into target cells.
- **c.** Cytotoxic cells express CD4 molecules but lack TCR.
- **d.** Cytotoxic cells do not secrete interferon-γ and are not T cells.

4. Cytotoxic T cells can kill their target cells through:
- **a.** Perforins and granzymes
- **b.** Phagocytosis
- **c.** Respiratory burst
- **d.** All of the above
- **e.** None of the above

5. A simple antibody molecule consists of:
- **a.** One heavy chain and one light chain.
- **b.** Two heavy chains and two light chains bound by peptide bonds.
- **c.** Two heavy chains and two light chains bound by dipeptide bonds.
- **d.** Two heavy chains and two light chains bound by disulfide bonds.

6. Which of the following is correct regarding primary and secondary antibody response?
- **a.** Primary response has a short lag phase, rapid exponential phase, short steady/peak phase, and slow decline.
- **b.** Secondary response has a long lag phase, slow exponential phase, short steady/peak phase, and slow decline.
- **c.** Primary response has a long lag phase, slow exponential phase, short steady/peak phase, and rapid decline.
- **d.** Secondary response has a short lag phase, slow exponential phase, long steady phase, and slow decline.

Bibliography

Abbas AK, Lichtman AH. *Cellular and Molecular Immunology*. 8th ed. Philadelphia: Saunders; 2014.

Delves PJ, Martin SJ, Burton DR, Roitt IM. *Roitt's Essential Immunology*. 13th ed. Malden, Mass: Wiley-Blackwell; 2017.

Janeway C. *Immunobiology: The Immune System in Health and Disease*. 8th ed. New York: Garland Science; 2011.

Jonuleit H, Tuettenberg A, Steinbrink K. Research in practice: regulatory T cells—targets for therapeutic approaches? *J Dtsch Dermatol Ges*. 2011;9(1):8–11.

LeRoith T, Ahmed S. Regulatory T cells and viral disease. In: Khatami M, eds. *Open Access Book Project; Inflammatory Diseases/Book 1*. Rijeka, Croatia: INTECH Open Access Publisher; 2011.

Parham P. *The Immune System*. 4th ed. New York: Garland Science; 2014.

Rao CV. *Immunology: A Textbook*. 3rd ed. Harrow, UK: Alpha Science; 2017.

Sakaguchi S. Regulatory T cells: history and perspective. *Methods Mol Biol*. 2011;707:3–17.

Sears KP, Divers TJ, Neff RT, et al. A case of Borrelia-associated cutaneous pseudolymphoma in a horse. *Vet Dermatol*. 2012;23:153–156.

Tizzard IR. *Veterinary Immunology: An Introduction*. 9th ed. Philadelphia: Saunders; 2013.

Appendix A

Answers to Practice Questions

Chapter 1	1. b	2. c	3. d	4. b	5. d	
Chapter 2	1. d	2. b	3. b	4. e	5. c	
Chapter 3	1. d	2. a	3. c	4. d		
Chapter 4	1. e	2. c	3. a	4. b	5. a	
Chapter 5	1. d	2. b	3. c	4. b	5. a	
Chapter 6	1. d	2. a	3. d	4. c	5. d	6. b
Chapter 7	1. c	2. c	3. a	4. a	5. d	
Chapter 8	1. a	2. c	3. b	4. d		
Chapter 9	1. c	2. a	3. c	4. a	5. e	6. e
Chapter 10	1. c	2. e	3. b	4. c	5. e	6. a
Chapter 11	1. c	2. b,e	3. c	4. a	5. c,e	
Chapter 12	1. c 7. d	2. b	3. e	4. b	5. a	6. b
Chapter 13	1. d	2. b	3. a	4. e	5. a	
Chapter 14	1. b	2. a	3. c	4. d	5. c	
Chapter 15	1. a	2. c	3. a	4. e	5. a,e	6. d
Chapter 16	1. c	2. d	3. c	4. a	5. a	
Chapter 17	1. d	2. c,e	3. b	4. c	5. a	
Chapter 18	1. d 7. b	2. c	3. a	4. a	5. e	6. e
Chapter 19	1. b 7. d	2. c	3. a	4. c	5. e	6. a
Chapter 20	1. b 7. d	2. e	3. d	4. b	5. d	6. a
Chapter 21	1. d 7. a	2. b	3. b	4. c	5. d	6. e
Chapter 22	1. a 7. d	2. b 8. e	3. c	4. e	5. b	6. d
Chapter 23	1. c 7. a	2. b 8. d	3.c	4. d	5. e	6. a
Chapter 24	1. d 7. a	2. e 8. b	3. d	4. b	5. a	6. c
Chapter 25	1. e 7. d	2. c	3. a	4. e	5. b	6. c
Chapter 26	1. c 7. e	2. b 8. c	3. a 9. d	4. e 10. d	5. a 11. a	6. c 12. b
Chapter 27	1. d 7. b	2. d 8. d	3. a 9. b	4. c 10. c	5. d	6. b
Chapter 28	1. d 7. d	2. d 8. a	3. b 9. c	4. c 10. b	5. e 11. d	6. d
Chapter 29	1. d	2. a	3. b	4. a	5. c	
Chapter 30	1. b	2. c	3. e	4. b	5. a	6. c
Chapter 31	1. d	2. c	3. a	4. c	5. d	
Chapter 32	1. c	2. d	3. e	4. b	5. a	
Chapter 33	1. c	2. c	3. c	4. e	5. b	
Chapter 34	1. a	2. c	3. b	4. d	5. d	
Chapter 35	1. d	2. d	3. d	4. c	5. c	
Chapter 36	1. b	2. b	3. a	4. e	5. c	
Chapter 37	1. e	2. e	3. c	4. e	5. f	6. f
Chapter 38	1. a,c,d,e,f	2. b	3. c	4. e	5. f	
Chapter 39	1. d	2. d	3. b	4. c	5. c	
Chapter 40	1. c	2. b	3. d	4. d	5. d	
Chapter 41	1. c	2. d	3. a	4. d	5. e	
Chapter 42	1. a	2. d	3. d	4. d,e	5. a,b,e	
Chapter 43	1. a	2. e	3. c	4. d	5. e	
Chapter 44	1. c	2. a	3. a	4. c	5. e	
Chapter 45	1. e 7. e	2. b 8. c	3. d 9. e	4. c	5. e	6. b
Chapter 46	1. d	2. e	3. a	4. b	5. c	
Chapter 47	1. a	2. d	3. a	4. a	5. c	6. d
Chapter 48	1. e	2. b	3. d	4. a	5. c	6. e
Chapter 49	1. b	2. a	3. d	4. a	5. a	6. b
Chapter 50	1. d	2. d	3. c	4. b	5. a	
Chapter 51	1. b 7. c	2. d	3. d	4. e	5. a	6. b
Chapter 52	1. a	2. d	3. c	4. b	5. e	6. d
Chapter 53	1. b	2. d	3. c	4. e	5. d	
Chapter 54	1. c	2. a	3. e	4. c		
Chapter 55	1. e	2. c	3. c	4. a	5. d	6. c

Index

Page numbers followed by "*f*" indicate figures, "*t*" indicate tables, and "*b*" indicate boxes.

A

Abdominal muscles, 523
Abductor muscles, 523
Abomasum, 591
  torsion of, 592*b*–594*b*, 593*f*–594*f*
Absolute refractory period, 192
Absorption, intestinal, 316–338
  of electrolytes, 326–329, 329*t*
    bicarbonate, 328, 328*f*
    chloride, 327–328, 327*f*
    potassium, 328–329, 328*f*
    sodium, 327, 327*f*
  of fats, 316, 331–334, 333*f*–335*f*
  metabolism during, 362
    amino acids in, 363, 365*t*
    fatty acids in, 363
    glucose in, 362–363
    ketone bodies in, 363, 363*f*
  paracellular, 325
  passive transport, 325, 326*f*
  transcellular, 324–325, 324*f*–325*f*
  transport systems in, 323–324
  of volatile fatty acids, 342–343, 342*b*, 343*f*
  of water, 326–329
Absorptive phase, 363–364
Acceleration/deceleration, of head, 110, 112*f*–113*f*
Accommodation, 139–140, 139*f*
ACE. *See* Angiotensin-converting enzyme.
Acetate, 362
Acetic acid, 342, 342*b*, 343*f*
Acetyl coenzyme A (acetyl-CoA), 362*f*, 364
Acetylcholine
  as autonomic neurotransmitter, 129–130, 130*f*
  in bronchoconstriction, 527, 528*f*
  in cardiac regulation, 194, 194*f*
  in cardiovascular regulation, 263–265, 264*t*
  in catecholamine release, 409
  pancreatic secretions and, 311–312
  in peripheral autonomic function, 131
  in synaptic transmission, 68–70, 68*f*, 70*f*
Acetylcholinesterase, 70
Acid-base balance, 509–517, 584–589
  acid excretion and, 510, 510*f*, 514–515, 514*f*
  ammonium generation/excretion and, 510, 511*f*
  bicarbonate and, 510–511, 511*f*, 515*f*
  buffers and, 509, 585–586, 586*f*
  disturbances of, 589–592. *See also* Acidosis; Alkalosis.
    compensatory adjustments for, 591
    diagnosis of, 592, 592*t*
  hydrogen ion concentration measurement and, 584, 586*f*–588*f*
  metabolic alkalosis and, 516*b*
  practice questions on, 516–517, 595
  respiratory acidosis and, 516*b*
  respiratory system and, 591
Acid-base homeostasis, 584–595
Acidosis
  definition of, 584
  metabolic, 590, 590*f*
    ammonia excretion during, 511–513, 512*f*
    ventilatory regulation and, 564
  respiratory, 589
    with renal compensation, 516*b*
  as respiratory stimulant, 563
Acinar cells, 310*f*, 311
Acquired immunity, 617–626
  antibody response in, 621–624, 622*f*–623*f*
  overview of, 617
  practice questions on, 625
  regulation of, 624
  T cells and. *See* T cells.
Acromegaly, 392
Acrosin, 450
Acrosome reaction, 450
Actin, 4, 5*f*, 73–75, 74*f*, 187*f*
  myosin, 77*f*
  sarcomeres and, 76–77, 186–187
Action potentials, 52, 60*f*, 64*f*
  cardiac
    antiarrhythmic drugs and, 199–201
    in contractions, 186, 187*f*, 187*t*
    duration of, 192–193, 192*f*
    length of, 189, 191*f*
    membrane calcium channels and, 190–192, 191*f*
    propagation of, 188, 188*f*, 190*f*, 193–194, 193*f*
    reentrant, 199
    tachyarrhythmias and, 198–199, 200*f*
  cochlea and, 170, 170*f*
  in gastrointestinal smooth muscle, 298–300, 299*f*–300*f*
  of sarcolemma, 75–76, 76*f*
  spindle sensory neurons and, 88–89, 88*f*–89*f*
  in synaptic transmission, 68–70, 68*f*
  visual cortex and, 144–145, 144*f*
Activating caspases, 39
Active hyperemia, 278
Active transport, 8, 10, 10*t*, 11*f*, 14, 175
  of calcium, 421
  secondary, 324, 491
  tertiary, 325
Active zone, of presynaptic membrane, 66, 67*f*
Actomyosin, 5, 5*f*–6*f*
Acute protein-losing enteropathy, 250*b*–251*b*
Adaptive relaxation, 297, 297*f*
Adenohypophysis, 385
  hormones of, 389
  hypothalamic releasing hormones and, 390–392, 390*f*
  in reproductive control, 431
Adenosine diphosphate (ADP), 4, 5*f*
Adenosine triphosphate (ATP), 3, 5*f*
Adenosinetriphosphatase (ATPase), 4, 6*f*
Adenylyl cyclase, 21
Adipose tissue, fatty acid mobilization by, 369, 372
ADP. *See* Adenosine diphosphate.
Adrenal cortex, 394
  anatomy of, 402–409, 402*f*
  fetal
    estrogen production, 452–453
    in parturition initiation, 453, 454*f*
  hormones of, 404. *See also* Glucocorticoids; Mineralocorticoids; *specific hormones*.
    metabolism of, 404
    plasma transport of, 404
    synthesis of, 403*f*
    types of, 403
Adrenal glands, 394, 402, 402*f*
Adrenal medulla, 127–128
  anatomy of, 409–412, 410*f*
  catecholamines in
    factors stimulating, 412
    mechanism of action of, 410–412, 411*f*, 412*t*
    synthesis of, 409–410, 410*f*–411*f*
Adrenalin. *See* Epinephrine.
β-adrenergic agonists, 197
β-adrenergic antagonists, 197, 201
Adrenergic neurotransmitters. *See* Norepinephrine.
Adrenergic receptors
  types of, 263, 264*t*, 410
  vasoconstriction and, 263, 264*t*
β-adrenergic receptors
  bronchodilation and, 527
  cAMP as second messenger, 21
  cardiac contractile force and, 196
  in cardiac regulation, 194, 196, 263, 264*t*
  saliva and, 308
Adrenergic synapses, 129
Adrenocortical hormones, 380–381, 380*f*, 404
Aerosols, 568–569, 569*f*
Afferents, 52
Agalactia, 392*b*–393*b*
Airflow, turbulent *vs.* laminar, 526
Airflow resistance, 526–527, 527*f*–528*f*
  dynamic compression and, 528
Airway disease, chronic, 530*b*
Airway obstruction, right-to-left vascular shunts and, 546–547
Airway resistance, 526
Airways
  anatomic dead space and, 521, 521*f*
  dynamic compression of, 528
  receptors in, 562
Alanine, from BCAA catabolism, 372*f*
Albumins, 382, 396, 461
Aldosterone, 403, 407*f*–408*f*, 601
  regulation of secretion of, 484, 484*f*
  renal effects of, 497
Alkaline tide, 310

Alkalosis
  definition of, 584
  metabolic, 590*f*, 591
    with paradoxical aciduria, 516*b*
  respiratory, 589–590, 590*f*
Allantochorion, 576
Allosteric shape changes, in proteins, 3, 3*f*
  in muscle contraction, 4–6, 6*f*
Alpha (α) motor neurons, 86, 88, 88*f*, 93, 94*f*
  coactivation with gamma neurons, 100, 104–105
  in lower motor neuron disease, 94
  in movement, 97
α-γ coactivation, 100, 104–105
Altitude
  acclimatization, 565–566
  hypoxic vasoconstriction and, 536, 536*f*
  ventilatory response to, 565–566
Alveolar capillaries, gas exchange in, 175
Alveolar cells, type II, 524, 580
Alveolar gas composition, 540–541, 540*f*–541*f*
Alveolar hyperventilation, 542, 589
Alveolar hypoventilation, 542, 543*f*, 589
Alveolar hypoxia, pulmonary vasoconstriction and, 535–536, 536*f*
Alveolar macrophages, 570, 571*f*
Alveolar septum, 532
Alveolar ventilation, 520–521, 540–541, 544*f*–546*f*
Alveolar ventilation/perfusion ratio, 544–546, 545*f*–546*f*
Alveolar vessels, pulmonary, 533, 533*f*
Alveolar-to-arterial oxygen tension difference, 542
Alveolus, 379
  lung, gas exchange in, 175, 176*f*, 544–546, 544*f*–546*f*
  mammary gland, 459, 459*f*
Amacrine cells, 140, 140*f*
Amines, 385
Amino acids. *See also specific amino acids.*
  branch-chain, 371, 371*f*–372*f*
  classification of, 365, 365*t*
  hepatic destruction, 366*f*, 368
  hydrophilic, 2
  insulin-stimulated uptake of, 368, 368*f*
  keto-analogues of, 366, 367*f*
  metabolism of, 363, 366*f*–367*f*
  mobilization of, 368*f*, 370–371
  muscle release of, 371
  during undernutrition, 372–373, 372*f*
Ammonia
  fermentative digestion and, 344, 345*f*, 347*f*
  renal excretion of, 511–513, 512*f*, 588
Ammoniagenesis, renal, 511–512, 512*f*
Amplitude, of sound waves, 166
Ampulla, 110, 113*f*
Anabolic nervous system, 131. *See also* Parasympathetic nervous system.
Anabolic steroids, 478
Anaerobic threshold, 566
Anal sphincter, 302
Anatomic dead space, 521, 521*f*
Anchorage dependence, of growth, 33
Androgens. *See also* Testosterone.
  anabolic steroids, 478
  in external genitalia formation, 430
  insensitivity, 434*b*–435*b*
Anemia
  cardiac output and, 183
  oxygen capacity and, 552–553
  parasitism and, 184*b*
Anesthesia
  hypoventilation and, 566*b*–567*b*
  hypoxemia during, 548*b*
  local, for arrhythmias, 199
Anestrus phase, 442
Aneuploidy, 30
Angiogenesis, 537
Angiotensin, pulmonary metabolism of, 573
Angiotensin I, 407–408, 484
Angiotensin II, 407–408, 484, 497
Angiotensin-converting enzyme (ACE), 484
Angiotensinogen, 407–408
Anion gap, 417, 592
ANP. *See* Atrial natriuretic peptide.
ANS. *See* Autonomic nervous system.
Anterior chamber, 138
Antiarrhythmic drugs, 199–201
Antibodies, 621–624
  antibody-antigen interaction, 623
  antigen exposure and, 621, 622*f*
  antigen-specific B-memory cell, 623–624
  B cells and, 623, 623*f*
  classes of, 621–623, 622*f*
Antibody-antigen interaction, 623
Anticoagulants, 181, 181*f*
Antidiuretic hormone (ADH), 21
  antidiuretic mechanism of, 388*f*
  renal sodium transport and, 498
  secretion of, 387
  water balance and, 506–507, 506*f*
Antigenicity, 609–610
Antigen-presenting cells (APCs), 614, 617, 619–621, 620*f*
Antigens, 608–616
  body's defense against, 610–614, 612*f*–614*f*
  classification of, 608, 609*f*
  definition of, 608
  in degree of immune response, 609–610
  inducing immune response, 608–609, 609*f*
  practice questions on, 615–616
  structures of bacteria, 608, 609*f*
Antiperistalsis, in colon, 301
Antiporters, 11, 325
Antithyroglobulin autoantibody test, 400
Antrum, 297–298
Aorta
  blood pressures in, 178, 230, 234
  oxygenation and, 177
Aortic depressor nerve, 265*f*, 266
Aortic incompetence/insufficiency, 223
Aortic regurgitation, 223, 226
Aortic stenosis, 222, 223*f*, 225–226
APCs. *See* Antigen-presenting cells.
Apical membrane, 317
Apocrine glands, 601
Apoptosis, 39, 40*f*–41*f*
  extrinsic pathway, 42
  intrinsic pathway, 39–42, 41*f*
  mechanism of, 39–43
  regulation of, 30–31
Aquaporins, 11, 506–507
Aqueous humor, 138, 145
Arachidonic acid, 454
Arachnoid, 54
Arachnoid villi, 153, 153*f*
Archicerebellum. *See* Vestibulocerebellum.
Arcuate nucleus, 432–433
Arginine vasopressin. *See* Antidiuretic hormone.
Aromatization, 430
Arrhythmias. *See* Cardiac arrhythmias.
Arterial baroreceptor reflex, 263, 265–267, 265*f*–266*f*
  in blood pressure maintenance after hemorrhage, 272, 275–276, 276*f*
  in exercise initiation, 278–279
  in heart failure compensation, 272–273, 273*f*
Arterial baroreceptors, 265, 265*f*
Arterial blood gas measurement, 592, 592*t*
Arterial blood pressure
  pulmonary, 178
  pulsatile nature of, 236–237, 238*f*
  regulation of, 265–267, 265*f*–266*f*
  systemic, 542*t*, 546
Arteries
  pulmonary
    blood pressure in, 178
    species variation in, 532–533
    vasoconstriction of, 535–536, 535*f*–536*f*
  umbilical, 579
  walls of, 241, 242*f*
Arteriolar vasodilation, 244
Arterioles
  blood flow regulation in, 233, 233*f*. *See also* vascular resistance.
  resistance, 233*f*
  walls of, 241, 242*f*
Arteriovenous anastomoses, 598, 599*f*
Arteriovenous fistulae, 223
Ascites, 248
Astrocytes, 158–159
Asynergia, 123
Ataxia, 123
Atelectasis, 524–526, 535, 546–547
ATP. *See* Adenosine triphosphate.
ATPase. *See* Adenosinetriphosphatase.
Atrial cells, 193
Atrial depolarization, 205–207, 205*f*–206*f*
Atrial fibrillation, 199, 213*b*–214*b*, 214*f*, 558*b*
Atrial flutter, 199
Atrial natriuretic peptide (ANP), 408, 486, 498
Atrial repolarization, 207
Atrial systole, 217
Atrial tachycardia, 198–199
  paroxysmal, 221
Atrial volume receptor reflex, 263, 267–268, 267*f*
  in blood pressure maintenance after hemorrhage, 272, 275–276, 276*f*
Atrioventricular (AV) node, 188, 188*f*
  auxiliary pacemaker function of, 195–196, 195*f*
  electrical characteristics of, 196*t*
Atrioventricular node block, 198, 201*b*, 211, 212*f*
Atrioventricular valves, 222*t*
Atrophy, 93, 94*f*
  juvenile pancreatic, 337*b*
Atropine, colic from administration of, 135*b*–136*b*
Auditory cortex, 170, 170*f*
Auerbach plexus, 284. *See also* Myenteric plexus.
Auscultation, cardiac, 223
Autocrine effector, 379, 379*f*
Autonomic nervous system (ANS), 126–137
  blood pressure, 133
  in cardiac regulation, 194–195, 194*f*
  cardiovascular regulation by, 263–265, 264*t*
  definition of, 126
  enteric nervous system and, 287*f*
  extrinsic innervation from, 286–287
  functions of, 52, 126
  gastrointestinal regulation by, 286–287
  immune system and, 135
  micturition, 134, 134*f*
  practice questions on, 136–137
  preganglionic *vs.* postganglionic neurons in, 126, 127*f*, 135
  pupillary light reflex, 134
  pupils and, 145
  smooth muscle innervation by, 126
  somatic nervous system vs, 126–127, 127*f*
  subdivisions of, 127. *See also* Parasympathetic nervous system; Sympathetic nervous system.
  thoracolumbar spinal cord, 127–128
  visceral afferent (sensory) neurons, 131–133
Autonomic reflexes, 131
AV node. *See* Atrioventricular (AV) node.
Avian species. *See* Birds.
Axons, 52, 57, 58*f*, 431
  cerebellum, 118, 119*f*–120*f*
  corticospinal tract, 102–103, 102*f*
  ganglion cell, 138, 140, 140*f*, 143–144, 144*f*
  reticulospinal tract, 100

Axons *(Continued)*
rubrospinal tract, 101
tectospinal tract, 101
Axoplasmic transport, 57
Azotemia, prerenal, 499*b*–500*b*

B

B cells, 617, 618*f*, 623, 623*f*
Backward heart failure, 275
Bacteremia, 174
Bacteria
antigenic structures of, 608, 609*f*
fermentative digestion a, 339–340
BAER. *See* Brainstem auditory-evoked response.
Balance
cerebellum and, 121, 122*f*
of Starling forces, 247
Baroreceptors, arterial, 265–267, 265*f*. *See also* Arterial baroreceptor reflex.
Basal ganglia
function of, 54
in modulation of movement, 97, 106–107, 106*f*
Basal metabolic rate (BMR), 597
Basilar membrane, 167, 168*f*
Basket cells, 119, 159*f*
Basolateral membrane, 317
Basolateral surface, 311
BBB. *See* Blood-brain barrier.
BCAAs. *See* Branch-chain amino acids.
Bicarbonate
as buffer, 586–587, 587*f*
gastric acid production and, 310
intestinal absorption of, 328, 328*f*
renal reabsorption of, 513, 514*f*, 588
renal secretion of, 515, 515*f*
as renal tubular buffer, 509–511
Bilateral epistaxis, 538*b*
Bilateral heart failure, 272
Bilateral receptor system, vestibular system as, 109, 110*f*
Bile
composition of, 312–313
secretion, 312–313
initiation of, 313, 313*f*
Bile acids, 312–313, 313*f*, 332
Bile ductules, 312
Bile pigments, 313
Bilirubin, 313
Binding sites, of proteins, 2, 3*f*
Bipolar cells, 140, 140*f*, 143–144
Birds
gastrointestinal motility in, 303, 304*f*
immune systems, 617
BLAD. *See* Bovine leukocyte adhesion deficiency.
Blind spot, 138
Blood. *See also* Plasma.
buffers in, 586–587, 587*f*
cardiovascular transport and, 173
cellular components of, 181–182, 181*f*, 182*t*. *See also* Erythrocytes; Leukocytes.
clotting of, 182
hormone transport in, 174, 381–382
oxygen transport in, 182–183, 183*f*
Blood cell counts, 182, 182*t*
Blood flow. *See also* Circulation; Vascular resistance.
arterioles in regulation of, 230
autoregulation of, 255–256, 255*f*–256*f*
in fetal circulation, 578*f*, 579
gastrointestinal, 330–331, 331*f*
intrinsic and extrinsic control mechanisms of, 253
local control of, 253–261
mechanical compression, 256–258, 257*f*–258*f*
metabolic control of, 253–255, 254*f*–255*f*, 254*t*
to muscle during exercise, 235*f*
organ distribution of, 234–236, 235*f*
paracrines and, 253, 254*t*, 256
Blood flow *(Continued)*
perfusion pressure and, 232*f*
practice questions on, 260, 270–271
precapillary sphincter in regulation of, 241
in pulmonary circulation, 236, 532–539, 535*f*
Blood gas measurement, arterial, 592, 592*t*
Blood pressure
arterial baroreceptor reflex and, 265–267, 265*f*–266*f*
atrial volume receptor reflex and, 267–268, 267*f*
autonomic control of, 133
measurement of, 178–179
neural and hormonal control of, 262–263
practice questions on, 270–271
pulmonary, 178–179, 532
systemic, 178–179, 230–231, 231*f*, 236–237, 236*f*
Blood vessels. *See also* Arteries; Arterioles; Capillaries; Veins.
perfusion *vs*. transmural pressure in, 175, 175*f*
placental, 576–577, 577*f*
of pulmonary circulation, 533, 533*f*
retinal, 138, 139*f*
types of, 179–180, 179*t*, 180*f*
walls of, 241
Blood volume
after hemorrhage, restoration of, 276–277, 277*f*
atrial volume receptor reflex and, 267–268, 267*f*
distribution of, 179, 179*t*
increase in heart failure, 273
neural and hormonal control of, 262–263
venous system as reservoir of, 231
Blood-brain barrier (BBB), 149–157, 155*f*, 243
Bloodworms, 174
BMR. *See* Basal metabolic rate.
Body fluids, as a defense against antigens, 610
Bohr effect, 553–554
Bone
calcium homeostasis and, 421–425, 422*f*
phosphate and, 421
Bony labyrinth, 109, 166–167
Bötzinger complex, 560, 562*f*
Bovine leukocyte adhesion deficiency (BLAD), 615*b*
Bovine viral diarrhea virus (BVDV), 124*b*
Bowman's capsule, 480, 482*f*
Bowman's space, 480, 482*f*
Bradycardia, 197, 210–211, 211*f*
Bradykinin
pulmonary metabolism of, 573
in vascular resistance regulation, 256
Brain. *See also* Brainstem; Cerebellum; Cerebral cortex.
anatomic regions of, 51–52
barriers between blood and, permeability, 154–155, 155*f*
ventricles of
anatomy of, 149, 150*f*
CSF production and, 149–150, 150*f*–151*f*
Brain tumors, 164*b*–165*b*
Brainstem
descending motor pathways of, 98–99, 99*f*
lateral, 99*f*, 101
medial, 99–100, 99*f*
function of, 53
PNS and, 128–129, 129*f*
respiratory regulation by, 560–562, 562*f*
transmission of action potentials from cochlea through, 170, 170*f*
vomiting and, 299–300
Brainstem auditory-evoked response (BAER), 163, 164*f*
Branch-chain amino acids (BCAAs), 371, 371*f*–372*f*
Breathing
frequency, 520
initiation of, 580
Brisket disease, 536, 536*f*
Bromocriptine, 433
Bronchi, 527. *See also* Tracheobronchial tree.
Bronchial circulation, 532, 537
Bronchial glands, 569–570, 570*f*
Bronchial smooth muscle, 527, 528*f*
Bronchioles, 527. *See also* Tracheobronchial tree.
Bronchoconstriction, 527, 528*f*
Bronchovascular bundle, 533, 533*f*
Brown fat, 598
Bruce effect, 447
Brush border
of intestinal mucosa, 317, 318*f*
of proximal renal tubule epithelium, 489–490, 490*f*
Buffers
acid-base balance and, 509, 585–586, 586*f*
in blood, 586–587, 587*f*
intracellular, 588–589
Bulk flow, 153
in cardiovascular system, 175, 175*f*–176*f*
water movement by, across capillary walls, 244–246, 245*f*–246*f*
Bundle of His, 189
Bursa of Fabricius, 617
Butyric acid, 342, 344*f*
BVDV. *See* Bovine viral diarrhea virus.

C

C cell, 395
C fibers, 562
Cabergoline, 433
Cachexia, 45–46
Calcitonin, 395
in calcium/phosphorus homeostasis, 422–423, 424*f*
renal calcium reabsorption and, 498–499
Calcium
body pool of, 421–425, 422*f*
calcitonin in, 422–423, 424*f*
parathyroid hormone in, 421–422, 422*f*
renal reabsorption of, 498–499
as second messenger, 20–21, 22*f*
in skeletal muscle contraction, 76, 76*f*
vitamin D and, 423–425, 424*f*
Calcium channel blockers, 80
for arrhythmias, 199–200
Calcium ion channels
in cardiac muscle, 190–192, 191*f*, 199–201
in skeletal muscle, 191–192
Calmodulin, 20–21, 22*f*
Calorigenic effect, 397–398
cAMP. *See* Cyclic adenosine monophosphate.
CAMP-dependent protein kinase, 21–23
Canaliculi, 312, 312*f*
Cancer, 29–50. *See also specific tumors.*
apoptosis and, 29, 39–43, 40*f*–41*f*
cell cycle regulation and, 29–31, 30*f*
cellular life span determinants, 29, 42–43
death by, 44–46
genetic dysfunction and, 30–31, 30*f*
growth factor pathway in, 29, 33–37, 34*f*
origin, spread, and growth of, 29, 30*f*, 43–47, 43*f*
tumor suppressors and, 29, 32*f*, 37–38, 38*f*–39*f*
Capacitation, of sperm, 450
Capillaries, 241–252
alveolar
gas exchange in, 175
systemic arterial gas content and, 542*t*, 546
blood flow in, 179, 179*t*, 230
of brain
blood-brain barrier and, 154, 155*f*
and oxygen delivery, 176, 176*f*
diffusion in, 175–177, 176*f*
fenestrated, 243
glomerular, 178, 480–482, 482*f*–483*f*
peritubular, 490–491, 491*f*
practice questions on, 251–252
pulmonary, 572
walls of, 241
water exchange in, 244–246, 245*f*–246*f*

Capillary endothelium, 480–482
Capillary recruitment, 544
Capillary refill time, 179
Capnography, 556, 557*f*
Capsule, bacterial, 608
Carbamino compounds, 181, 556
Carbohydrates, 319. *See also specific carbohydrates*.
  in colostrum, 461
  digestion of, fermentative, 342
  in mammalian milk, 466
  in neonates, 463–466
Carbon dioxide
  alveolar exchange of, 540–541
  blood content of, 556, 556*f*
  diffusion of, 542–544
  production of, 518
  transport of, 555–556, 555*f*
Carbon dioxide tension
  buffer system and, 586, 588*f*
  ventilatory control and, 563–564, 563*f*–564*f*
Carbon dioxide transport, 555–556, 555*f*
Carbon monoxide, hemoglobin and, 554
Carbonic anhydrase, 310, 556, 586
Carboxypeptidase, 321*t*
Carcinomas, 31
Cardiac action potentials
  antiarrhythmic drugs and, 199–201
  in contractions, 186, 187*f*, 187*t*
  duration of, 192–193, 192*f*
  length of, 189, 191*f*
  membrane calcium channels and, 190–192, 191*f*
  propagation of, 188, 188*f*, 190*f*, 193–194, 193*f*
  reentrant, 199
Cardiac afterload, 221
Cardiac arrhythmias
  atrioventricular node block and, 198
  electrolyte abnormalities and, 174
  origins of, 197–198
  reentrant, 199
  sinus arrest, 197
  tachyarrhythmias, 198–199, 200*f*
  treatment of, 197, 197*t*, 199–201
Cardiac cycle, 188–189, 188*f*, 190*f*, 216–229, 217*f*
Cardiac decompensation, 274
Cardiac defects
  of heart valves, 222–224, 222*t*, 223*f*
  murmurs and, 222–224, 222*t*, 223*f*
  patent ductus arteriosus, 223, 223*f*, 226
  pathological consequences of, 216, 225–226, 225*f*
Cardiac failure. *See* Heart failure.
Cardiac glycosides, 200–201
Cardiac hypertrophy, 224–225
Cardiac mucosa, of stomach, 309
Cardiac muscle, 4, 78–79
  contraction
    force of, 196–197, 196*t*
    skeletal muscle contraction *vs*., 186–187, 187*f*, 187*t*
  excitation-contraction coupling in, 79–80
  functional syncytium of, 187–188, 187*f*
  initiation and organization of heartbeats, 188–189, 188*f*, 190*f*
  innervation of, 126, 127*f*
  membrane calcium channels and, 190–192, 191*f*
  structure of, 78–79, 186, 187*f*
Cardiac output, 173, 178, 216
  arterial blood pressure and, 235*f*
  determinants of, 218, 218*f*
  during exercise, 234, 278–279
  in fetus, 578*f*, 579
  heart rate and, 218, 218*f*, 221–222, 221*f*, 222*t*
  neural-hormonal control of, 262
Cardiac pacemakers
  artificial, 197–198
  atrioventricular node. *See* Atrioventricular (AV) node.
Cardiac pacemakers *(Continued)*
  ectopic, 198, 200*f*
  sinoatrial node, 188
Cardiogenic shock, 174
Cardiomyopathy, dilative, 213*b*–214*b*
Cardiovascular reflexes, 263. *See also* Arterial baroreceptor reflex; Atrial volume receptor reflex.
Cardiovascular system. *See also* Circulation; Heart.
  autonomic regulation of, 263–265, 264*t*
  blood vessels of, 179–180, 179*t*, 180*f*. *See also* Blood vessels.
  blood volume distribution in, 179*t*
  bulk flow and diffusion in, 175, 175*f*–176*f*
  cardiac output and, 178
  dysfunction of, 173–174. *See also specific disorders*.
    congenital *vs*. acquired, 174
    primary *vs*. secondary, 174
  function of
    overview of, 173–185
    primary, 173
  practice questions on, 185
  pulmonary *vs*. systemic circulation, 177–178, 177*f*
  substances transported by, 174–175
Caretakers, 37
Carnitine, 373
Carnitine palmitoyltransferase I (CPT-I), 373
Carotid bodies, 563, 563*f*
Carotid sinuses, 265, 265*f*
Carrier membrane proteins, 11, 12*f*
Caruncles, 452, 576
Caseins, 461, 466
  iodinated, 467
Caspases, 39
Catabolism, 371, 371*f*–372*f*
Catalysis, proteins and, 2
Cataracts, 140
Catecholamines
  effects of, 410, 411*f*, 412*t*
  secretion of, 412
  synthesis of, 3, 4*f*, 409–410, 410*f*–411*f*
Catechol-O-methyltransferase (COMT), 129–130
Cats
  cerebellar hypoplasia in, 124*b*
  chronic renal failure, 487*b*–488*b*
  flea infestation anemia in, 558*b*
  hepatic lipidosis in, 375*b*–376*b*
  hyperthyroidism in, 400–401
  pregnancy maintenance in, 452
Caveolae, 80, 573
CCK. *See* Cholecystokinin.
CDKs. *See* Cyclin-dependent kinases.
Cecocolic orifice, 356
Cecum
  equine, 301*f*, 302
  motility function of, 356–358, 356*f*–357*f*
Cell adhesion, in cell cycle regulation, 34*f*, 35–37, 36*f*
Cell body, 51, 57, 58*f*
Cell cycle, 29
  checkpoints in, 37
  control of, 29, 31–33, 31*f*
  dysfunction in regulation of, 30–31, 30*f*
  gatekeepers, 32*f*, 37–38, 38*f*–39*f*
  by growth factors, 33–34, 36*f*
  inhibitors of, 32*f*, 37–38, 38*f*–39*f*
  MAP kinase pathway in, 34*f*, 35–37, 36*f*
  phases of, 31, 32*f*
  therapy, 47–48, 48*b*–49*b*
Cell life span, 39–43
  dysfunction in regulation of, 30–31, 30*f*
Cell suicide. *See* Apoptosis.
Cell walls
  of bacteria, 608
  of plants, 341–342
Cells
  membranes of. *See* Plasma membranes.
  proliferation of, in cancer, 31–33, 31*f*–32*f*
Cellular life span, determinants of, 42–43
Cellulase, 341
Cellulose, 341
Central circulation, 177–178, 179*t*, 230
Central command, in exercise, 278
Central command neurons, 566
Central diabetes insipidus, 387, 507*b*–508*b*
Central nervous system (CNS)
  afferent neurons in, 287–288
  anatomical and functional regions of, 51, 52*b*, 52*f*
  autonomic division. *See* Autonomic nervous system (ANS).
  meninges and cerebrospinal fluid, 54
  in movement. *See* Movement.
  relationships of ventricles, CSF, and meninges, 150*f*
  in reticulorumen regulation, 339, 352–353
  spindle sensitivity and, 89–90
Central pattern generator (CPG), 560, 561*f*, 565*f*
Central venous catheter, 218
Central venous pressure, 218
Centroacinar cells, 311
Cephalic phase
  of gastric secretion, 298, 310
  of pancreatic secretion, 311
Cerebellar cortex, 118
  function of, 120*f*
Cerebellar disease, 123
Cerebellar hypoplasia, 124*b*
Cerebellar nuclei, 118, 119*f*
Cerebellar peduncles, 118
Cerebellum, 118–125
  anatomy of, 118, 119*f*
  histology and phylogeny of, 119–121, 120*f*–122*f*
  input/output organization of, 121*f*
  in modulation of movement, 97, 106–107, 106*f*
  molecular layer of, 119, 120*f*–122*f*
  in motor learning, 123
  movement and, 118
    eye, balance and, 121, 122*f*
    limb, muscle tone and, 121–123, 122*f*
    sequences in, 122*f*, 123
  practice questions on, 124–125
Cerebral cortex
  function of, 54
  in hearing, 170, 170*f*
  histological features of, 158–159, 159*f*
Cerebral hemispheres, 54
Cerebral infarction, 176–177
Cerebrocerebellum, 121, 122*f*, 123
Cerebrospinal fluid (CSF), 54, 149–157, 150*f*
  absorption of, 153–154, 153*f*
  flow of, 150*f*–151*f*, 151
  formation of, 149, 150*f*–151*f*
  functions of, 149
  glymphatic system of, 152, 152*f*
  hydrocephalus and, 149, 154
  practice questions on, 156–157
Cervix, 450, 454
cGMP. *See* Cyclic guanosine monophosphate.
Channels, 11, 12*f*
Checkpoints, in cell cycle, 33, 37
Chemical digestion, 319, 320*f*
Chemical driving force, 60
Chemically gated ion channels, 61
Chemokines, 571
Chemoreceptor trigger zone, 300
Chemoreceptors, lungs, 560, 563–566, 563*f*–565*f*
Chest leads, 208
Chief cells, 309, 421–422
Chloride, intestinal absorption of, 327–328, 327*f*
Cholecystokinin (CCK)
  pancreatic secretions and, 311–312
  site of, 288, 289*t*

Cholesterol
  bile acid formation from, 312–313, 313*f*
  in steroid hormone synthesis, 380–381, 380*f*, 403–404, 403*f*
  transport of, 14–15, 15*f*
Cholesterol esterase, 332
Cholesteryl ester, 332, 333*f*
Cholinergic neurotransmitters. *See* Acetylcholine.
Cholinergic receptors
  muscarinic, 263, 264*t*
  types of, 263
Cholinergic synapses, 129
Chorion, 576
Chorionic girdle, 453
Chorionic gonadotropin, 451–452
Choroid, 138
Choroid plexus, 149–150, 151*f*
Chromaffin cells, 127–128, 409, 410*f*
Chronic airway disease, 530*b*
Chronic myeloid leukemia (CML), 47
Chronic renal failure, 487*b*–488*b*
Chronic renal insufficiency, 507*b*–508*b*
Chronic respiratory acidosis, 516*b*
Chylomicrons, 332–334, 335*f*
Chymosin, 321*t*
Chymotrypsin, 321*t*
CID. *See* Combined immunodeficiency.
Cilia, of vestibular system, 109
Ciliary body, 138
Ciliary process, 145
Circadian rhythms, 385
Circulation. *See also* Blood flow.
  blood volume distribution in, 179, 179*t*
  bronchial, 532, 537
  capillary blood flow in, 179, 179*t*
  central, 177–178, 179*t*, 230
  fetal, 577–579, 578*f*
  pulmonary. *See* Pulmonary circulation.
  splanchnic, 178
  systemic
    arrangement of, 177–178, 177*f*
    blood pressure in, 231*f*
    blood vessel types in, 179*t*, 241
    organ blood flow regulation in, 234–236
    perfusion pressure of, 178, 230
    total peripheral resistance of, 234
    vascular resistance in. *See* Vascular resistance.
Circulatory convection, 598
Circumventricular organs, 155
Circus movement, 199
Cisterna magna, 151, 152*f*
Cisterns, mammary, 459, 460*f*
Citrate, 364, 365*f*
Clara cells, 518, 569, 570*f*
Climbing fiber axons, 120*f*, 121
Clitoris, 430
Clostridium tetani, 71–72
CML. *See* Chronic myeloid leukemia.
CNS. *See* Central nervous system.
Coagulation, 180, 182
Cochlea
  action potential from, 170, 170*f*
  anatomy of, 109, 110*f*, 166–167, 168*f*
  sound waves and, 167–170, 168*f*–169*f*
Coelomic epithelium, 428
Cold stress, 604
Colic, 174
  from atropine administration, 135*b*–136*b*
  Strongylus parasitism, endotoxic shock and, 184*b*
Colipase, 332
Collapsing trachea, 528
Collateral ventilation, 529
Collecting duct system
  reabsorption/secretion by, 496, 496*f*, 514–515, 514*f*–515*f*
Collecting duct system *(Continued)*
  urea recycling in, 503–504, 504*f*
  urine pH and, 513
  water permeability of, 506–507, 506*f*
Colligative properties, of solutions, 9
Colloid, of thyroid gland, 395
Colloid osmotic pressure, 9, 244
Colon, motility of, 293, 295, 300, 356–358, 356*f*–357*f*
Colony-stimulating factors, 34
Color blindness, 143
Colostrum, 458, 461
  components of, 461, 463*t*
  formation of, 461
  passive immunity and, 461
Combined immunodeficiency (CID), 624*b*–625*b*
Communicating hydrocephalus, 154
Compartmentalization, 14–15, 15*f*
Compensatory response, of baroreceptor reflex, 266
Compliance
  definition of, 231, 241
  pulmonary, 524–525, 524*f*
  vascular, blood vessel types and, 232*f*
  ventricular, 218–220, 219*f*–220*f*
Compression and blood flow, mechanical, 256–258, 257*f*–258*f*
COMT. *See* Catechol-O-methyltransferase.
Concentration difference, 175
Conduction deafness, 171
Conductive heat loss, 597, 600
Cones, 140, 140*f*
  photoreception by, 141–143, 142*f*–143*f*, 143*t*
Conformation, of proteins, 2
Congenital deafness, 171*b*
Congestive heart failure, 279*b*–281*b*
Conn's syndrome. *See* Hyperaldosteronism.
Consensual pupillary light reflex, 145
Contact inhibition, of growth, 33
Continuous murmurs, 222, 222*t*
Contractility, ventricular, 216, 220–221, 220*f*
Contraction
  cardiac
    force of, 196–197, 196*t*
    initiation of, 188
    premature, 198
    ventricular, 210, 211*f*
    vs. in skeletal muscle, 186–187, 187*f*, 187*t*
  muscle
    actin and myosin sliding, 76–77, 77*f*
    body movement and, 73–80, 74*f*
    gastrointestinal, 294
    ischemia during, 257, 258*f*
    sarcoplasmic reticulum, 76, 76*f*
    strength, 77–78
Convective heat loss, 599–600
Copulation, ovarian cycle modification by, 440
Core temperature, 597*t*, 598, 601
Cornea, 138
Corneal ulcers, 135*b*–136*b*
Coronary arteries, oxygen transport in, 177
Coronary artery disease, 177
Coronary capillaries, 177
Corpus luteum, 436–439
  practice questions on, 441
  in pregnancy, 437–438, 451–452, 451*f*–452*f*
Cortex
  auditory, 170, 170*f*
  cerebellar, 118
  cerebral. *See* Cerebral cortex.
  visual, 144–145, 144*f*
    primary, 144, 144*f*
Cortical neurons, behavior of, 160–163
Corticobulbar tract, 103
Corticospinal tract, 97, 101–102, 102*f*
Corticosteroid-binding globulin, 404
Corticosteroids. *See* Glucocorticoids; Mineralocorticoids.
Corticosterone, 403
Corticotropin, 389–390, 389*t*, 405–406, 405*f*
Corticotropin-like intermediate lobe peptide, 389
Corticotropin-releasing hormone (CRH), 390, 391*t*
Cortisol, 403, 405*f*
  fetal, in parturition, 453–455, 454*f*–455*f*
  synthesis of, 382*f*
Cotransmitters, 131
Cotransport, 11, 324
Coughing, 528, 570
Countercurrent mechanism
  in renal medullary hypertonicity, 504–505, 504*f*–505*f*
  in temperature regulation, 599, 600*f*
Coupled sodium chloride transport, 327
Cows
  abomasal torsion in, 592*b*–594*b*, 593*f*–594*f*
  brisket disease in, 538*b*
  BVDV cerebellar damage in, 124*b*
  diarrhea in, 269*b*–270*b*
    with dehydration and acidosis, 337*b*
  hardware disease in, 279*b*–281*b*
  hypocalcemia in, 80*b*–81*b*
  hypomagnesemia in, 107*b*
  inadequate nutrition and ovarian inactivity in, 448*b*
  metabolic alkalosis in, 516*b*
  obturator nerve paralysis in, 91*b*
  prolonged pregnancy in, 456*b*
  recurrent infections in, 615*b*
CPG. *See* Central pattern generator.
CPT-1. *See* Carnitine palmitoyltransferase I.
Cranial nerve nuclei, 54
Cranial nerve VIII, 109, 111*f*
Craniosacral system. *See* Parasympathetic nervous system.
Creatinine clearance, 487
CRH. *See* Corticotropin-releasing hormone.
Cricopharyngeal muscle, 296
Crista ampullaris, 110, 111*f*
Crop, avian, 303, 304*f*
Cross-bridging, 4–5
Cross-reactive epitopes, 608
Cryptorchidism, 472
Crypts of Lieberkühn, 317, 318*f*
CSF. *See* Cerebrospinal fluid.
Cud chewing, 351
Cumulus oophorus, 434
Cupula, 110, 113*f*
Curd, 466
Cushing's syndrome, 406
Cyanosis, 554
Cyclic adenosine monophosphate (cAMP), 19, 23*f*
Cyclic guanosine monophosphate (cGMP), 23–24
Cyclic nucleotide phosphodiesterase (PDE), 21
Cyclin-dependent kinases (CDKs), 31–33, 32*f*
Cyclins
  in cell cycle regulation, 32
  MAP kinase pathway and, 35–36
Cytokines, 34
  in immune system regulation, 613*f*, 614
  in respiratory defenses, 571
  T cells and, 618–619, 619*f*
Cytosol, 189
Cytotoxic T cells, 618–619, 619*f*–620*f*

D

DAG. *See* Diacylglycerol.
Dead space ventilation, 520, 521*f*, 547
Dead space/tidal volume ratio, 521
Deafness, 171, 171*b*
Death receptor pathway, 39, 41*f*
Decerebrate rigidity, 101
Decibel scale, 166
Decompensation, cardiac, 274
Defense-alarm reaction, 194, 195*f*, 263, 268, 268*f*
Defibrillation, 199

Degenerative myelopathy, 55*b*
Deglutition, 295
Dehydration
  arterial blood pressure and, 234
  diarrhea with, and acidosis, 337*b*
Dehydroepiandrosterone, 452–453
Delivery. *See* Parturition.
Dendrites, 51, 57, 58*f*
Dendritic cells, 614
Dendritic spines, 57–58, 58*f*, 71
Dense bodies, 79
Deoxyhemoglobin, 556
Deoxyribonucleic acid (DNA)
  binding, 16–17
  cellular life span and, 42–43
Depolarization, 61, 69*f*
  atrial, 205–207, 205*f*–206*f*
  ventricular, 205–207, 206*f*
    premature, 210
Desensitization, of receptors, 17
Dextrins, 321
Diabetes insipidus, 387, 507*b*–508*b*
Diabetes mellitus, 416–418, 426, 426*b*
  ketone bodies in, 374
  type 1, 416
  type 2, 416, 418
Diabetic ketoacidosis, 416
Diacylglycerol (DAG), 19, 24*f*
Diagnostic procedures
  arterial blood gas measurement, 592, 592*t*
  brainstem auditory-evoked responses, 163, 164*f*
  electrocardiogram. *See* Electrocardiogram (ECG).
Diaphragm, 523
Diarrhea
  with dehydration and acidosis, 337*b*
  neonatal, 592*b*–594*b*, 593*f*–594*f*
  pathophysiology of, 316, 336–337, 336*f*
  secretory, 337
Diastasis, 217
Diastole, ventricular, 216–218, 223*f*
Diastolic aortic pressure, 178
Diastolic arterial pressure, 236
Diastolic filling time, 220, 221*f*
Diastolic murmurs, 222–223, 222*t*, 223*f*
Diencephalon, 53*f*, 54
Diestrus, 442, 443*f*
Diet, proteins in
  digestion and, 311
  fermentative digestion of, 344, 345*f*
Diffusion. *See also* Passive transport.
  alveolar, 175, 542–544, 542*f*–544*f*
  capillary, 175, 175*f*–176*f*
  facilitated, 11, 12*f*
  Fick's law, 243–244, 244*f*
Diffusion coefficient, 243–244
Digestion, 317–323, 317*f*
  chemical, 319, 320*f*
  of fats, 318–323, 321*f*–323*f*, 331–334
  fermentative. *See* Fermentative digestion.
  hydrolysis in, 329
  luminal-phase of, 319–321, 321*t*
  membranous-phase of, 322, 323*f*, 326*f*
  in neonates, 335–336
  of peptides, 323, 324*f*
  of proteins, 321–322
Digestive enzymes, 6, 311, 319, 322
Digestive tract. *See* Gastrointestinal (GI) tract.
3,4-Dihydroxyphenylalanine, 4, 4*f*, 409–410
Diiodotyrosine, 395
Dilative cardiomyopathy, 213*b*–214*b*
Dilution rate, of rumen, 339, 352
Dipalmitoyl phosphatidylcholine, 524
2,3-diphosphoglycerate (2,3-DPG), 579
Dipole, electrical, 203, 204*f*
Direct pupillary light reflex, 145, 146*f*
Dirofilaria immitis, 174
Disaccharides, 319–320
Discontinuous capillaries, 243
Distal renal tubules, 494–496, 494*f*–495*f*
Diurnal rhythms, 385
DNA. *See* Deoxyribonucleic acid.
Dogs
  acromegaly in, 392
  anesthetic hypoventilation in, 566*b*–567*b*
  atrioventricular block in, third-degree, 201*b*
  brain tumors in, 164*b*–165*b*
  chronic renal insufficiency in, 507*b*–508*b*
  congenital deafness in, 171*b*
  degenerative myelopathy in, 55*b*
  diabetes insipidus in, 507*b*–508*b*
  diabetes mellitus in, 426*b*
  dilative cardiomyopathy in, 213*b*–214*b*
  exercise intolerance in, 279*b*–281*b*
  femoral nerve mononeuropathy in, 91*b*
  glomerulonephritis in, 487*b*–488*b*
  glucosuria in, 499*b*–500*b*
  heartworm disease in, 174, 239*b*
  heat stroke in, 606*b*
  hemangiosarcoma in, 48*b*–49*b*
  herniated intervertebral disk, 95*b*
  homonymous hemianopia in, 147*b*
  Horner's syndrome, 135*b*–136*b*
  hyperadrenocorticism in, 406–407
  hypoadrenocorticism in, 408–409, 499*b*–500*b*
  hypoglycemia in, 63*b*–64*b*
  hypothyroidism in, 399–400
  hypoventilation in, 548*b*
  hypoxemia with hyperventilation in, 566*b*–567*b*
  infertility in, 440*b*–441*b*
  intracranial pressure in, increased, 155*b*–156*b*
  intraoperative hemorrhage in, 269*b*–270*b*
  lower motor neuron disease in, 95*b*
  lung fibrosis in, 530*b*
  mitral insufficiency in, 573*b*–574*b*
  motor cortex focal lesion in, 107*b*
  myasthenia gravis in, 71*b*–72*b*
  pancreatic tumor in, 426*b*
  pancreatitis in, 314*b*
  patent ductus arteriosus in, 258*b*–260*b*, 581*b*–582*b*
  prolonged pregnancy in, 456*b*
  pulmonic stenosis in, 226*b*–227*b*
  respiratory acidosis in, 516*b*
  sexual attraction in spayed bitch, 448*b*
  upper airway obstruction in, 592*b*–594*b*, 593*f*–594*f*
  upper motor neuron disease in, 95*b*
  vestibular syndrome in, 116*b*
DOPA. *See* 3,4-Dihydroxyphenylalanine.
Dopamine, 431
  biosynthesis of, 3–4, 4*f*
  as hypothalamic hormone, 390
  in Parkinson's disease, 106
  prolactin inhibition by, 462, 464*f*
Dormitory effect, 447
Driving force, 7–8, 13
Driving pressure, for gas diffusion, 543
Ductus arteriosus, 455, 579
  patent, 223, 223*f*, 226, 258*b*–260*b*
Ductus venosus, 455
Dura mater, 54
Dynamic compression, of airways, 528
Dysmetria, cerebellar disease and, 123
Dyspnea, 581*b*–582*b*

E

Eardrum, 166
Ears
  inner, 109, 110*f*
  middle, 166, 167*f*
  outer, 166, 167*f*
Eccentric implantation, 452
Eccrine glands, 601
ECF. *See* Extracellular fluid.
ECG. *See* Electrocardiogram.
Eclampsia, 468–469
Ectopic cardiac pacemakers, 198, 200*f*
Ectotherms, 596–597
Edema
  in chronic liver disease, 9–10
  electrolyte abnormalities and, 174
  heart failure and, 273–275, 274*f*
  peripheral, 27
  pulmonary
    formation of, 572
    mitral regurgitation and, 225
  pulmonary, electrolyte abnormalities and, 174
EEG. *See* Electroencephalogram.
Efferents, 52
Egestion, 303
Einthoven, Willem, 207
Einthoven's triangle, 207–208
Ejection fraction, 216
Elastase, 321*t*
Elastic vessels, 241
Elasticity, 241
Elastin fibers, 524
Electrical dipole, 203, 204*f*
Electrical driving force, 60
Electrical signaling, potassium transport and, 13
Electrocardiogram (ECG), 203–215
  abnormal patterns in, 209–211, 209*f*–212*f*, 214*f*
  abnormal voltages in, 208–209
  basic principles of, 203–205, 204*f*, 205*t*
  intervals and segments in, 207, 207*f*
  leads used in, 207–208, 208*f*, 210*f*
  practice questions on, 214–215, 215*f*
  variability in large animals, 211–213
  voltage deflections in, causes of, 205–207, 206*f*
Electrochemical potential, 7
Electroencephalogram (EEG), 158–165, 159*f*–163*f*, 164*b*–165*b*
Electrolytes. *See also specific electrolytes.*
  abnormalities of, in hypoadrenocorticism, 408–409
  absorption of, 326–329
  active secretion of, 329–330, 330*f*
  cardiovascular transport of, 174–175
  disorders of. *See also specific abnormalities.*
    cardiac arrhythmias, 174
    edema, 174
  intestinal secretion of, 329–330
  in plasma, 180–181, 181*t*
  in saliva, 308–309, 308*f*
Electromagnetic radiation, in infrared range, 600
Electromyography (EMG)
  in lower motor neuron disease, 94
  in upper motor neuron disease, 95
Electroretinogram (ERG), 144
Electrotonic current, 62
Embden-Meyerhof pathway, 342
Embryo development, 450–451
EMG. *See* Electromyography.
Emulsification, 332
Encephalopathy, hypoxemic ischemic, 239*b*
End-diastolic volume, ventricular, 216
  determinants of, 218, 219*f*–220*f*
  stroke volume and, 218, 218*f*–219*f*, 221–222
  ventricular contractility and, 216, 220–221, 220*f*
Endocrine cells, 288, 288*f*
  distribution of, 289*f*
Endocrine pancreas, 311
Endocrine system, 378–393. *See also* Hormones; *specific endocrine glands.*
  feedback control mechanisms in, 378, 384–385, 385*f*
  of gastrointestinal tract, 284
  hormone synthesis in, 380–381, 380*f*
  nervous system and, 379–380
  pituitary gland and. *See* Pituitary gland.

Endocrine system *(Continued)*
  postreceptor cell responses, 383, 383*f*
  practice questions on, 393
  signal amplification in, 379–380
Endocytosis, 14, 15*f*, 243, 382
Endolymph, 109, 166–167
Endometrial cups, 453
Endopeptidases, 321, 321*t*
Endosome, 14–15, 15*f*
Endothelin
  renal sodium transport and, 498
  in vascular resistance regulation, 256
Endothelium-derived constricting factors, 486
Endothelium-derived relaxing factors, 486
Endotherms, 596–597
Endotoxemia, 258*b*–260*b*
Endotoxic shock, 174, 184*b*
End-product inhibition, 4
End-systolic volume, 216
Enteric nervous system (ENS), 127, 284–286
  autonomic nervous system and, 287*f*
  fibers and receptors within, 286*f*
  intrinsic, 284
  motor neurons of, 284
  organization of, 286*f*
  plexus of, 284
  regulation of, 285*f*
Enterochromaffin, 288–289
Enterocytes, 317, 319*f*
Enterogastric reflexes, 298
Enterohepatic circulation, 313, 313*f*
Enterohepatic cycle, 397
Enterotoxins, 337
Envelope, bacterial, 608
Environment, heat exchange with, 599–601
Enzymes. *See also specific enzymes*.
  calcium and, 421–425
  catalysis by, 2
  in catecholamine synthesis, 3–4, 4*f*
  salivary, 307–308
Epinephrine
  in cardiovascular regulation, 263–265, 264*t*
  cardiovascular transport of, 174–175
  synthesis of, 3–4, 4*f*, 409–410, 411*f*
Epistaxis, bilateral, 538*b*
Epithelial slit diaphragm, 482, 482*f*
Epithelial transport, 14
Epithelium
  coelomic, 428
  visceral, 482, 483*f*
Epitopes, 608
EPSP. *See* Excitatory postsynaptic potential.
Equilibrium, electrochemical, 8
Equilibrium potential, 13
  resting membrane potential and, 60
Equine chorionic gonadotropin, 389
Equine large hindgut, function of, 347*f*, 355–358
Erythrocyte sedimentation rate (ESR), 182
Erythrocytes, 181, 181*f*, 182*t*
  in oxygen transport, 176, 183*f*
Esophageal groove, 355
Esophagus, motility of, 296–297, 296*f*–297*f*
Estriol, 452–453
Estrogen receptor, 25–26
Estrogen response element, 25–26
Estrogens
  embryonic synthesis of, 451
  gestational production of, 452–453, 453*f*
  negative-feedback inhibition, 444
  in preovulatory surge of gonadotropins, 436–437
Estrous cycle, 442–449, 443*f*
Eupnea, 560–561
Euthyroid sick syndrome, 400
Evaporative heat loss, 597, 601
Exchange vessels. *See* capillaries.
Exchangers, 325
Excitation-contraction coupling, 5–6, 76, 79–80
Excitatory postsynaptic potential (EPSP), 61, 61*f*, 88–89, 160
Executioner caspases, 39
Exercise
  arterial blood pressure and, 234, 235*f*
  cardiac changes during, 221–222, 222*t*
  cardiovascular response to, 278–279, 278*f*–279*f*
  gas transport during, 556–557, 557*f*
  pulmonary circulation during, 537
  ventilatory response to, 566
Exercise intolerance, 174, 226*b*–227*b*
  heart failure and, 273–275, 274*f*, 279*b*–281*b*
Exercise reflex, 278
Exhalation, 521, 528
Exocrine effectors, 379
Exocrine pancreas, 311
Exocytosis, 14, 15*f*, 243, 380, 421
Exopeptidases, 321
Extension, 73
External genitalia, development of, 429–430, 429*f*
External intercostal muscles, 523
External sphincter, 302
Extra-alveolar vessels, pulmonary, 533, 533*f*
Extracellular fluid (ECF). *See also* Interstitial fluid; Plasma.
  calcium homeostasis and, 421–425, 422*f*
  colloid osmotic pressure and, 9
  substance concentrations in, 10*t*
Extrafusal muscle fibers, 77–78, 86, 87*f*–88*f*, 88–89
Extraocular muscles, 139
Extrapyramidal motor system, 98–99
Extravasation, 44
Eyes, 138–148. *See also specific structures*.
  accommodation of, 139–140, 139*f*
  anatomy of, 138–139, 139*f*
  diagnostic testing of, 144
  intraocular pressure in, 145
  movement, coordination of, 121, 122*f*
  in photoreception, 141–143, 142*f*–143*f*, 143*t*
  practice questions on, 147–148
  pupillary size regulation, 145
  retina
    anatomy of, 139–140
    cell types of, 140, 140*f*
    fovea, 140–141, 141*f*–142*f*
    tapetum, 139*f*, 141
    visual image processing in, 143–144

F

Facilitated diffusion, 11, 12*f*
F-actin, 4
Fainting. *See* Syncope, vasovagal.
Fak (focal adhesion kinase) protein, 36–37
Far-field potentials, 163. *See also* Sensory-evoked potentials.
Farnesyl transferase inhibitors (FTIs), 48
Fasting, 372–373, 372*f*
Fast-twitch muscle fibers, 77
Fats. *See also* Fatty acids.
  digestion and absorption of, 331–334
  in milk, 462, 464*f*, 466, 466*t*
  during undernutrition, 372–373, 372*f*
Fatty acid-binding proteins, 332, 335*f*
Fatty acids, 362–363
  adipose tissue mobilization of, 372
  glucose and, conversion of, 364
  in glucose synthesis, 374, 374*f*
  during undernutrition, 373, 373*f*
  volatile. *See* Volatile fatty acids.
Feedback control mechanisms
  in endocrine system, 378, 384–385
  for regulation of body temperature, 602, 603*f*
Female reproductive system. *See also* Lactation; Ovulation; Pregnancy; Reproductive cycles; *specific organs*.
  corpus luteum, 436–439
  ovarian cycles, 440, 440*f*
  ovarian follicle, 433–434
  puberty and, 443–445, 444*f*
    mammary gland development and, 460–461, 461*f*–462*f*
Femoral nerve mononeuropathy, 91*b*
Fenestrated capillaries, 243
Ferguson reflex, 454, 455*f*, 463
Fermentative digestion, 339–360
  in equine hindgut, 339, 355
  microbial ecosystem of, 339–341
  practice questions on, 359–360
  of protein, 344, 345*f*
  sites of, 340
  substrates and products of, 339, 341–346
    plants cell walls, 341–342
    protein, 344–346, 346*f*
    volatile fatty acids, 342–343, 343*f*
Fertilization, 450
Fetus
  circulation of, 577–579, 578*f*
  estrogen production by, 452–453
  lung development in, 580
  oxygen transport in, 576–583, 577*f*, 577*t*, 579*f*
  in parturition initiation, 453–455, 454*f*–455*f*
Fever, 604–605, 604*f*–605*f*
Fibers, 319
Fibrinogen, 180
Fibrosis, lung, 530*b*
Fick's law of diffusion, 241, 243–244, 244*f*
Fight or flight system, 131. *See also* Sympathetic nervous system.
Filtration, 245
Fimbria, 428
First breath, 580
First nursing, 463–466
First-degree AV node block, 198, 211, 212*f*
Flagella, as antigens, 608
Flea infestation, 558*b*
Flexion, 73
Fluid exchange, 241–252
Fluid mosaic structure, of biomembranes, 6
Foals
  cervical trauma in, 85*b*
  combined immunodeficiency in, 624*b*–625*b*
  diarrhea in, 592*b*–594*b*, 593*f*–594*f*
  hypoxemic ischemic encephalopathy in, 239*b*
  seizures in, 155*b*–156*b*
Focal lesion, of motor cortex, 107*b*
Follicle cells, 428
  of thyroid gland, 395, 395*f*
Follicle-stimulating hormone (FSH), 389–390, 389*t*
  in male, 476
Follicular phase, 440, 442
Food. *See also* Diet; Nutrients.
  deglutition, 295
  mastication, 295
  prehension, 295
  stomach's processing of. *See* Stomach, motility of.
Foot processes, 482, 482*f*
Foramen ovale, 455, 579
Forced convection, 599–600
Forebrain, 52–53
Forestomach, 340–341
Forward heart failure, 275
Fovea, 140–141, 141*f*–142*f*
Foveola, 141, 141*f*
Fractional composition, of gas mixtures, 540
Fractional excretion rate, 489
Fractionation, of movement, 103
Free energy, 7
Frequency coding, by sensory receptors, 82

Frictional resistance, in airways, 526–527, 526*f*–527*f*
Frontal lobe, motor cortices of, 103–104, 104*f*
Frostbite, 606
Fructose, spermatozoal utilization of, 450
Fructose-1,6-bisphosphatase, 369–370
Fuels, metabolic, 362–369
  homeostasis and, 362
Functional residual capacity, 520
Functional specific gravity, 349
Functional syncytium, of cardiac muscle, 187–188, 187*f*
Fungi and fermentative digestion, 340–341

G

G proteins, 18, 19*f*, 35, 35*f*
G-actin, 5*f*
Gain-of-function mutation, 33
Galactose, 466
Galactosyl transferase, 466
Gallbladder, 313
Gamma (γ) motor neurons, 86
  coactivation with alpha neurons, 100, 104–105
  in spindle sensitivity control, 89–90
Ganglion cells, 126, 138, 140–141, 140*f*, 143–144, 144*f*
GAP. *See* GnRH-associated peptide.
Gap junctions, 188, 453
Gas exchange, 540–549
  abbreviations used for, 541*t*
  alveolar, 542, 542*f*–543*f*
    systemic arterial gas content and, 542*t*, 546
  defense mechanisms and, 568
  disorders of, 548*b*
  evaluation of, 547–548
  in hypoxemia, 541–542
  metabolism and, 518–520
  placental, 576, 577*f*, 577*t*
  practice questions on, 549
  processes involved in, 518, 519*f*
  right-to-left shunts and, 546–547, 547*f*
  in tissues, 544
Gas mixtures, composition of, 540, 540*f*
Gases
  in plasma, dissolved, 181*t*
  transport. *See* Carbon dioxide transport; Oxygen transport.
Gastric emptying, 298–300, 299*f*
Gastric inhibitory polypeptide (GIP), 289*t*, 298
Gastric mucosa, 309, 309*f*
Gastric phase, of gastric secretion, 310
Gastric pits, 309
Gastric secretion, 309–310
  cephalic phase of, 298, 310
  gastric phase of, 310
Gastric ulcers, 314*b*
Gastrin, 289*t*, 298
Gastrointestinal blood flow, 330–331
Gastrointestinal motility, 293–306
  in bird, 303, 304*f*
  of colon, 300–302, 301*f*–302*f*
  control of, 298
  of esophagus, 296–297, 296*f*–297*f*
  in herbivores, omnivores and carnivores, comparison of, 301*f*, 302
  practice questions on, 305–306
  slow waves in, 293–294, 294*f*
  of small intestine, 298–300, 299*f*–300*f*
  smooth muscles and, 294–295, 295*f*
  of stomach, 297
    distal region in, 297–298, 298*f*
    indigestible material and, 299
    proximal region in, 297, 297*f*
  in vomiting, 299–300
Gastrointestinal (GI) tract. *See also* Stomach.
  absorption of. *See* Absorption, intestinal.
  avian, 303, 304*f*
Gastrointestinal (GI) tract *(Continued)*
  digestion in. *See* Digestion; Fermentative digestion.
  regulation of, 284–291, 290*f*
    by enteric nervous system, 284–286
    by immune system, 290
    by intrinsic endocrine system, 288–290, 289*t*
    by intrinsic enteric nervous system, 284
  secretions of, 307–315, 314*b*
    bile secretion, 312–313
    gastric, 309–310, 309*f*
    pancreatic, 311–312
    salivary glands and, 307–309
Gated channels, 11, 12*f*
Gatekeepers, 37
GBM. *See* Glomerular basement membrane.
GDP. *See* Guanosine diphosphate.
Gefitinib, 47
General adaptation syndrome, 402
Genetic dysfunction and cancer, 30–31, 30*f*
Genitalia, external, development of, 429–430, 429*f*
Germ line, 30–31
Gestation. *See* Pregnancy.
GFR. *See* Glomerular filtration rate.
GHRH. *See* Growth hormone-releasing hormone.
GI tract. *See* Gastrointestinal (GI) tract.
GIP. *See* Gastric inhibitory polypeptide.
Gizzard, 303
Gland acini, 308, 308*f*
Glandular mucosa, 309
Gleevec, 47
Glial cells, 52, 57, 158–159
Glomerular basement membrane (GBM), 480–482
Glomerular capillaries, 178, 480–482, 482*f*–483*f*
Glomerular filtrate, 480
Glomerular filtration rate (GFR), 480, 588
  determinants of, 482–483, 483*f*
  measurement of, 486–487
  practice questions on, 488
  regulation of, 484–486, 484*f*–485*f*
Glomerular tuft, 480, 482*f*
Glomerulonephritis, 487*b*–488*b*
Glomerulus, 480, 481*f*
  blood filtration, 480
  structure of, 480–482, 482*f*–483*f*
Glomus cells, 563, 563*f*
Glucagon, 418–420
  functions of, 418–419, 419*f*
  in ketone body formation, 373, 373*f*
  regulation of, 420*f*
  secretion of, 367, 367*f*, 419
  synthesis of, 418–419, 420*f*
Glucocorticoids, 403
  effects of, 404, 404*t*–405*t*
  inflammatory response and, 406–409, 406*f*
    hyperadrenocorticism and, 406–407
    hyperaldosteronism, 409
    hypoadrenocorticism and, 408–409
    insulin and, 415*t*
    mineralocorticoids and, 407–408, 407*f*, 407*t*
  metabolism of, 404
  potencies of, 404*t*
  regulation of, 405–406, 405*f*
  synthesis of, 403–404, 403*f*
Glucokinase, 418
Gluconeogenesis, 418–419, 419*f*
  glucocorticoids and, 404–405
  hepatic, 362–363, 370–371
  in ruminants, 375, 375*f*
Glucose
  conversion of, in rumen, 342*b*, 364
  in lactose, 466
  metabolism of
    catecholamines and, 410–412, 411*f*, 412*t*
    glucagon, 418–420
    hepatic, 363–364
    insulin and, 413–418, 415*b*
Glucose *(Continued)*
  transport of, 10–11, 11*f*
  during undernutrition, 372–373, 372*f*
Glucosuria, 499*b*–500*b*
Glucuronides, 397, 404
Glycocalyx, 317
Glycogen, decreased synthesis of, 418–419, 419*f*
Glycogen phosphatase, 369–370
Glycogen synthase, 369–370
Glycogenolysis, 362, 418–419, 419*f*
Glycolysis, 362–363
Glycoproteins, 389, 431
Glycosaminoglycans, 454
"Glymphatic" system, 152, 152*f*
GnRH-associated peptide (GAP), 431
Goats
  neurological disease in, 164*b*–165*b*
  parasitism in, 184*b*
Goblet cells, 318
Golgi tendon organs, 86, 87*f*, 90, 90*f*
Gonadal activity, 431
Gonadotropin receptors, 434
Gonadotropin-releasing hormone (GnRH), 390, 391*t*
  in males, 476
  puberty and, 444
  pulsatile release of, 432, 432*f*, 446
Gonadotropins. *See also* Chorionic gonadotropin; Follicle-stimulating hormone; Luteinizing hormone.
  modifications of release of, 432–433, 432*f*
  preovulatory surge of, 436–437, 438*f*
Gonads
  embryonic development of, 429*f*, 430
  organization of, 428, 430*f*
G-protein-coupled receptors (GPCRs), in signaling, 17–18, 17*f*, 18*t*
Grain engorgement toxemia, 359*b*
  microbial ecosystem of, 340*b*
  substrates and products of volatile fatty acids, 342*b*
Granule cell layer, 120*f*, 121
Granulosa cells, 428
Gray ramus, 127
Grit, 303
Growth factor pathway, in cell proliferation, 33–37, 34*f*–35*f*
Growth hormone (GH), 389–390, 389*t*
  changes in, peripheral fuel utilization and, 374–375
  for enhancement of lactation, 467–468, 468*f*
  in lactogenesis, 462
Growth hormone-inhibiting hormone, 391*t*
Growth hormone-releasing hormone (GHRH), 390, 391*t*
GTP. *See* Guanosine triphosphate.
GTPase. *See* Guanosinetriphosphatase.
Guanosine diphosphate (GDP), 18
Guanosine triphosphate (GTP), 18
Guanosinetriphosphatase (GTPase), 17, 17*f*
Gular flutter, 601
Gut wall
  afferent neurons in, 287–288
  anatomy of, 285*f*
  innervation from autonomic nervous system in, 286–287
  multiple cell types in, 290–291
  neurohumoral regulatory molecules of, 286*b*
Guttural pouches, 599

H

Hair cells, 109, 111*f*
Hardware disease, 27, 279*b*–281*b*
Harvey, William, 173
Haustra, 301, 356

Head
  acceleration/deceleration of, 110, 112*f*–113*f*
  movements, vestibular reflexes for, 114–115, 115*f*
Hearing, 166–172. *See also* Ears.
  practice questions on, 171–172
  receptor organs for, 109, 110*f*
Heart. *See also cardiac entries.*
  autonomic regulation of, 196–197, 196*t*
  conduction system, 188–189, 188*f*, 190*f*
    dysfunction of, 197, 197*t*, 200*f*
  contraction of cardiac muscles
    force of, 196–197, 196*t*
    initiation of, 188
    vs. in skeletal muscle, 186–187, 187*f*, 187*t*
  electrical activity of, 186–202
  practice questions on, 202, 228, 228*f*
  as pump, 216–229
  workload of, cardiac defects and, 224–225
Heart failure, 174
  backward *vs.* forward, 275
  causes of, 272
  complications of, 273–275, 274*f*
  congestive, 279*b*–281*b*
  consequences and compensations for, 272–273, 273*f*
  decompensation in, 274
  definition of, 272
  edema formation and, 273–275, 274*f*
  left-sided, 272
  practice questions on, 281–282
  right-sided, 272, 538*b*
    hypoxia of altitude and, 536, 536*f*
  ventricular contractility in, 220–221
Heart murmurs, 216, 222–224, 222*t*, 223*f*
Heart rate, 218
  autonomic regulation of, 196, 196*t*
  cardiac output and, 218, 218*f*, 221–222, 221*f*, 222*t*
Heart sound
  murmurs. *See* Heart murmurs.
  normal, 217
Heart valves, defects of, 222, 223*f*
Heartworm disease, 174
Heat loss, 597, 599–601
Heat production, 598, 598*f*, 598*t*, 604–605
Heat stress, 602–604, 603*f*
Heat stroke, 605
Heat transfer, 598–599, 599*f*–600*f*
Heaves, 530*b*
Hemangiosarcoma, 48*b*–49*b*
Hematocrit, 181, 181*f*, 182*t*, 183, 183*f*
  hemorrhage and, 276
Hematology, 182, 182*t*
Hemicellulose, 341
Hemiparesis, 104
Hemoglobin, 181. *See also* Oxyhemoglobin.
  as blood buffer, 586–587
  carbon dioxide and, 553–554, 553*f*
  carbon monoxide affinity and, 554
  color of, 554
  fetal, 579–580
  indices for, 182*t*
  oxygen binding of, 551, 551*f*
  in oxygen transport, 176, 182, 183*f*, 550–551, 551*f*
Hemoglobinopathies, 182
Hemolytic crisis, neonatal, 468–469
Hemoptysis, 537
Hemorrhage, 174
  arterial blood pressure and, 234
  blood volume restoration after, 272, 276–277, 277*f*
  cardiovascular responses to, 272, 275–276, 275*f*–276*f*
  intraoperative, 269*b*–270*b*
Hemorrhagic shock, 174
Hemostasis, 182
Henle's loop, 494, 494*f*, 513
  countercurrent mechanism in, 504, 504*f*
Hepatic lipidosis, in cat, 375*b*–376*b*
Hepatocytes, 312
Hering-Breuer reflex, 562
Herniated intervertebral disk, 95*b*
Heterometric autoregulation, 219
Heterotrimeric G protein, 17–18, 17*f*
High-pressure, high-resistance side of circulation, 236–237
Hindgut, 341
  anatomy of, 358
Hippocampus, 54
Histamine, 256, 310
  arteriolar dilation of, 253
Homeostasis
  autonomic nervous system and, 126
  fuel, 362
Homeotherms, 597
Homonymous hemianopia, 147*b*
Horizontal cells, 140, 140*f*
Hormones. *See also specific hormones.*
  blood transport of, 174–175, 381–382
  classes of, 380
  definition of, 378
  feedback control mechanisms on, 378, 384–385, 385*f*
  lactation and, 467–468, 468*f*
  metabolism of, 378, 384
  in plasma, 181*t*
  pulmonary metabolism of, 573
  receptors for, 382–383
  synthesis of, 380–381, 380*f*
Hormone-sensitive lipase (HSL), 372
Horner's syndrome, 135*b*–136*b*
Horses
  acute protein-losing enteropathy in, 250*b*–251*b*
  agalactia in, 392*b*–393*b*
  androgen insensitivity in, 434*b*–435*b*
  atrial fibrillation in, 213*b*–214*b*, 558*b*
  bilateral epistaxis in, 538*b*
  bloodworms in, 174
  chronic airway disease in, 530
  colic in
    from atropine administration, 135*b*–136*b*
    endotoxic shock and, secondary to *Strongylus* parasitism, 184*b*
  combined immunodeficiency in, 624*b*–625*b*
  exercise intolerance in, 226*b*–227*b*
  gastric ulcers in, 314*b*
  hyperlipemia in, 375*b*–376*b*
  hypoxemia during anesthesia in, 548*b*
  infertility in, 478*b*–479*b*
  insufficient milk/udder development in, 469*b*
  neonatal isoerythrolysis in, 469*b*
  neurological disease in, 55*b*
  persistent luteal phase, 440*b*–441*b*
  pituitary pars intermedia dysfunction in, 392*b*–393*b*
  pleuritis in, 573*b*–574*b*
  rabies in, 304*b*–305*b*
  stimulating ovulatory activity in mares, 448*b*
  strangles in, 615*b*
  tetanus in, 71*b*–72*b*
  uveitis in, 147*b*
  vasculitis in, 250*b*–251*b*
  vestibular disease in, 116*b*
HSL. *See* Hormone-sensitive lipase.
Hyaluronidase, 450
Hydrocephalus, 149, 154
Hydrochloric acid, gastric secretion of, 309–310
Hydrogen bonding, 2
Hydrogen ions
  acid-base homeostasis and, 591–592
  measurement of, 585, 586*f*–588*f*
  renal excretion of, 509, 510*f*, 588
  sources of, 584–585
Hydrolysis, 319
Hydrophilic amino acids, 2
Hydrophobic amino acids, 2
Hydrophobic interaction, 2
Hydrostatic pressure, 9–10, 231*f*
Hydroxyapatite, 421
Hyperadrenocorticism, 406–407
Hyperaldosteronism, 409
Hypercalcemia, 424–425
Hypercapnia, 546, 563
Hyperemia
  active, 253, 255*f*, 278
  reactive, 254–255, 255*f*
Hyperglycemia, 419*f*
Hyperkalemia, 591
Hyperkalemic periodic paralysis (HYPP), 27*b*
Hyperlipemia, 375*b*–376*b*
Hyperpolarization, 61
Hyperreflexia, 94–95
Hypersensitization, of receptors, 17
Hypertension, 234
Hyperthermia, malignant, 80–81
Hyperthyroidism, 400–401
Hypertrophy, cardiac, 224–225
Hyperventilation
  alveolar
    causes of, 542
    respiratory alkalosis and, 589
  buffer system and, 588
  hypoxemia with, 566*b*–567*b*
Hypoadrenocorticism, 408–409, 499*b*–500*b*
Hypocalcemia, 425, 425*f*
  in cows, 80
  lactation and, 468–469
Hypoglycemia, 63*b*–64*b*
  catecholamine release and, 412
  glucagon secretion and, 419*f*
Hypokalemia, 591
Hypomagnesemia, 107*b*
Hyponatremia, 499*b*–500*b*
Hypoparathyroidism, 425
Hypophysiotropic hormones, 390, 390*f*, 391*t*
Hypoproteinemia, 184*b*, 572
Hypothalamic-hypophyseal portal system, 178
Hypothalamic-pituitary-gonadal system, 477–478
Hypothalamic-pituitary-testicular axis, 476, 477*f*
Hypothalamoneurohypophyseal system, 386*f*
Hypothalamus
  function of, 54
  gender differentiation and, 430*f*, 431
  pituitary gland and, 385–392, 386*f*
  sexual behavior and, 445
  thermoregulatory set point and, 602, 603*f*
Hypothermia, 605
Hypothyroidism, 399–400
Hypoventilation, 527–528, 548*b*
  alveolar
    buffer system and, 589
    causes of, 542, 543*f*
  anesthesia and, 566*b*–567*b*
Hypoxemia, 273–274, 563
  during anesthesia, 548*b*
  gas exchange in, 541–542
  high altitude ascent and, 565–566
  with hyperventilation, 566*b*–567*b*
Hypoxemic ischemic encephalopathy, 239*b*
Hypoxic pulmonary vasoconstriction, 535
Hypoxic vasoconstriction, 536, 536*f*

I

ICC. *See* Interstitial cells of Cajal.
Ileocecal sphincter, 300
Ileum, 300
IMCD. *See* Inner medullary collecting duct.
Immune response
  antigens inducing, 608–609, 609*f*
  degree of, antigens in, 609–610
  regulation of, 624

Immunity
  acquired, 617–626
    antibody response in, 621–624, 622*f*–623*f*
    overview of, 617, 618*f*
    practice questions on, 625
    regulation of, 624
  innate, 568, 608–616, 612*f*–614*f*
  passive, colostrum and, 461
Immunogenicity, 609–610
Immunogens. *See* Antigens.
Immunoglobulins. *See* Antibodies.
Impaction colic, 359*b*
Implantation, 451
Indirect pupillary light reflex, 145, 146*f*
Induced ovulators, 440
Inducible nitric oxide synthase, 612–613
Infarction, 173
  cerebral, 176–177
  myocardial, 177
Infectious antigens, 608, 609*f*
Inferior colliculus, of midbrain, 170
Infertility
  in bulls, 478*b*–479*b*
  in mares, 440
  in stallions, 478*b*–479*b*
Influenza, in pigs, 606*b*
Information, definition of, 2
Infrared radiation, heat exchange and, 600
Inhalable particulates, 568–569
Inhalation, 521
Inhibin, 436
Inhibitory postsynaptic potentials (IPSPs), 61, 61*f*, 89, 160
Innate immunity, 568, 608–616, 612*f*–614*f*
Inner ear. *See also* Vestibular system; *specific structures*.
  anatomy of, 109, 110*f*, 166–167
Inner medullary collecting duct (IMCD), 503–504, 504*f*
Inorganic phosphate, 421
Inositol 1, 4, 5-triphosphate, 19, 24*f*
Insulin, 413–418
  deficiency. *See* Diabetes mellitus.
  functions of, 413–416, 414*f*
  protein synthesis and, 368
  regulation of, 420*f*
  secretion of, 413, 415*b*, 415*f*
  sites of action and effects of, 415*t*
  synthesis of, 413, 414*f*
Insulinomas, 63
Integrins, 36–37
Intention tremor, cerebellar disease and, 123
Intercalated cells, 496, 514
Intercalated disks, 78–79, 188
Intercostal muscles, 523
Interdigestive motility complex, 299
Interleukins, in respiratory defenses, 571
Internal sphincter, 302
Interneurons, 82, 159*f*
Internodes, 62
Intersegmental reflexes, 83–84, 84*f*
  in lower motor neuron disease, 93–94
Interstitial cells of Cajal (ICC), 293–294, 294*f*
Interstitial fluid, 152. *See also* Extracellular fluid.
  capillary diffusion and, 175, 242, 242*f*
  edema and, 244, 248–250, 248*f*–250*f*
  hydrostatic pressure, 245
  reabsorption of, after hemorrhage, 276–277, 277*f*
  solute exchange in, 241, 242*f*
Interstitial implantation, 452
Intervals, in electrocardiograms, 207, 207*f*
Intestinal absorption. *See* Absorption, intestinal.
Intestinal epithelium, growth and development of, 334–335
Intestinal secretion, of water and electrolytes, 329–330
Intestinal villi, 334–335
Intracellular amino acid pool, 368*f*
Intracellular buffers, 588–589
Intracellular fluid, substance concentrations in, 10*t*
Intracranial pressure, increased, 155*b*–156*b*
Intrafusal muscle fibers, 86, 87*f*
Intraluminal zymogen, 322*f*
Intraocular pressure, 145
Intravasation, 44
Involution, 466
Iodinated casein, 467
Iodine, 395
Iodotyrosine dehalogenase, 396
IPSPs. *See* Inhibitory postsynaptic potentials.
Iris, 138, 145
Irritant receptors, 562, 570
Ischemia, 173
  muscle contraction and, 257, 258*f*
  of organs in heart failure, 274, 274*f*
Islets of Langerhans, 412, 413*f*
Isoerythrolysis, neonatal, 469*b*
Isomaltase, 322
Isomaltose, 321, 321*f*
Isovolumetric relaxation, 217
Isthmus, 309

J

Jejunum, 329*t*
Joints, sensory receptors in, 90–91
Junctional folds, 66–68, 67*f*
Junctional tachycardia, 198–199
Juvenile pancreatic atrophy, 337*b*
Juxtacapillary receptors, 562
Juxtaglomerular apparatus, 485
Juxtaglomerular cells, 484
Juxtamedullary nephrons, 502

K

Kallikrein, 256
Ketone, 362
Ketone bodies, 363, 363*f*
  in diabetes mellitus, 374
  in mitochondria, 373*f*, 374
  during undernutrition, 373, 373*f*
Kidneys, 480. *See also renal entries*.
  acid-base regulation. *See* Acid-base balance.
  ammonia excretion by, 511–513, 512*f*, 588
  bicarbonate
    reabsorption by, 513, 514*f*
    secretion by, 515, 515*f*
  in calcium homeostasis, 421
  calcium reabsorption by, 498–499
  collecting duct system of. *See* Collecting duct system.
  in electrolyte regulation, 175
  glomerular filtration rate and. *See* Glomerular filtration rate.
  glomerulus of. *See* Glomerulus.
  heart failure and, 274, 274*f*
  hydrogen ion excretion by, 509, 510*f*
  portal system of, 178
  potassium reabsorption by, 496–497, 497*f*
  sodium reabsorption by, 497, 503
  water balance and, 501. *See also* Water balance.
Knee jerk reflex, 88–89, 88*f*
Knudsen's two-hit hypothesis, 37
Koilin, 303
Krebs cycle, 362
  acetyl coenzyme A and, 374, 374*f*
  glycolysis and, 362–363
Kupffer's cells, 611, 613*f*
Kwashiorkor, 248–249

L

Labia, 430
Labyrinth, 109
Labyrinthitis, 116*b*
Lacrimal gland, 139
Lactase, 466
Lactation, 458–470. *See also* Lactogenesis.
  cycle of, 458, 466–468, 467*f*–468*f*
  hormonal induction of, 467–468, 467*f*–468*f*
  hypocalcemia and, 468–469
  reproductive cycles and, 447
Lactogen, placental production of, 453
Lactogenesis, 458, 462–463, 464*f*
Lactose, 319–320, 321*f*
  in colostrum, 461
  in milk, 466, 466*t*
Lactose synthetase, 466
Lamina densa, 480–482, 482*f*
Lamina rara externa, 480–482
Lamina rara interna, 480–482
Laminar airflow, 526
Landscaping genes, 44–45
Laryngeal hemiplegia, 523
Lateral brainstem motor pathways, 99, 99*f*, 101
Lateral corticospinal tract, 102–103
Lateral geniculate nucleus, 144, 144*f*
Lateral space, 317–318
Leads, electrocardiographic, 207–208, 208*f*, 210*f*
Leak channels, 11, 12*f*
  potassium ion, 13
Lens, accommodation by, 139–140, 139*f*
Leukemias, 43–44
Leukocytes, 181–182, 181*f*, 182*t*
Leukotrienes, pulmonary metabolism of, 573
Leydig cells, 428–429, 478
LH. *See* Luteinizing hormone.
Libido, testosterone and, 446
Ligand binding, 2, 3*f*
Ligand-gated channels, 11, 12*f*
Light, circadian rhythms and, 385
Light transduction, 141–143, 142*f*–143*f*, 143*t*
Lignin, 341
Limit dextrins, 321
Lingual glands, 308
Lingual lipase, 307
Lipase, 332
Lipidosis, hepatic, 375*b*–376*b*
Lipids. *See* Fats; Fatty acids; Phospholipids.
Lipid-soluble/insoluble substances, capillary diffusion of, 242–243, 242*f*
Lipolysis, 398
Lipoprotein lipase (LPL), 368–369
β-Lipotropin, 389, 389*f*
Liver
  amino acid uptake by, 371, 371*f*–372*f*
  bile secretion and, 312, 312*f*
  fatty acid metabolism in, 373, 373*f*
  glucose metabolism in, 363–364
  ketone body formation in, 373, 373*f*
  microanatomy of, 312, 312*f*
Liver disease, chronic, 9–10
Lobar bronchus, resistance in, 526
Local anesthetics, for arrhythmias, 199
Long-loop feedback system, 391, 391*f*
Long-loop nephrons, 502–503, 503*f*
Long-loop reflexes, 84
Low pressure, 236–237
Low-density lipoproteins (LDLs)
  in cholesterol transport, 14–15, 15*f*
  thyroid hormones and, 398
Low-dose dexamethasone suppression (LDDS) test, 407
Lower motor neurons, 93–96, 94*f*, 104–105
  disease of, 93–94, 94*f*
  practice questions on, 95–96
Low-pressure, low resistance side of circulation, 236–237
Lung compliance, 523–524

Lungs. *See also pulmonary entries.*
acid-base balance and, 509
airways of. *See* Airways.
birth transition of, 580–581
blood flow distribution in, 534, 535*f*
chemoreceptors in, 560, 563–566, 563*f*–565*f*
defense mechanisms of, 568–571, 569*f*–571*f*
edema. *See* Pulmonary edema.
elasticity of, 523–524, 524*f*
fetal development of, 580
fibrosis of, 530*b*
gas exchange in, 176, 176*f*, 544–546, 544*f*–546*f*
hypoxic vasoconstriction in, 535–536, 536*f*
lymphatics of, 572–573, 572*f*
mechanical properties of, 528–529, 529*f*
metabolic functions of, 573
nonrespiratory functions of, 568–575
practice questions on, 531, 575, 582–583
surfactant in, 524–525, 525*f*
ventilation. *See* Ventilation.
Luteal oxytocin, 439
Luteal phase, 440, 451
Luteinization, 437, 453
Luteinizing hormone (LH), 389–390, 389*t*
in corpus luteum maintenance, 438
in males, 476
in ovulation, 436–437, 438*f*
preovulatory surge of, 436–437
Luteinizing-inhibiting factor, 436–437
Luteolysis, 436, 437*f*, 453
Luteotrophs, 451, 452*f*
Luteotropin, 438, 451
Lymph nodes, swollen, 615*b*
Lymphatic system, pulmonary, 572–573, 572*f*
Lymphatic vessels, 572
Lymphedema, 249
Lymphocytes. *See* B cells; T cells.
Lymphoid stem cells, 617–618
Lymphopoiesis, 618*f*
Lysozymes, 307

M

Machinery murmurs, 223, 223*f*
Macroelectrodes, use of, 160–163
Macrophages, 611–613, 613*f*
Macula, 110, 111*f*
Macula densa, 485
Magnesium, in diabetic ketoacidosis, 417
Magnocellular neurons, 386
Malabsorptive diarrhea, 336, 336*f*
Male reproductive system, 471–479
anabolic steroids and, 478
functional anatomy of, 471–473, 472*t*
hypothalamic-pituitary-testicular axis, 476, 477*f*
practice questions on, 479
puberty, 476–478
spermatogenesis, 473–475, 474*f*
Malignant hyperthermia, 80*b*–81*b*
Malleus, 166
Malnutrition, 372–375
Malonyl CoA concentrations, 373
Maltase, 466
Mammary buds, 459
Mammary gland, 458–470. *See also* Colostrum; Lactation.
anatomical aspects of, 458–459, 459*f*
development of, 459, 461*f*
diseases associated with, 458, 468–469
involution of, 466
lobes/lobules of, 461, 462*f*
practice questions on, 470
Mammogenesis, 458, 460–461, 461*f*–462*f*
Mandibular glands, 308
MAO. *See* Monoamine oxidase.
MAP. *See* Mean arterial pressure.
MAP kinase pathway, 34*f*, 35–37, 36*f*
Mastication, 295
Mastitis, 468
MCH. *See* Mean corpuscular hemoglobin.
MCHC. *See* Mean corpuscular hemoglobin concentration.
Mean aortic blood pressure, 178, 234, 237
Mean arterial pressure (MAP), 265–266
Mean circulatory filling pressure, 230–231, 231*f*–232*f*
Mean corpuscular hemoglobin (MCH), 182, 182*t*
Mean corpuscular hemoglobin concentration (MCHC), 182, 182*t*
Mechanical compression and blood flow, 256–258, 257*f*–258*f*
Mechanoreceptors
in gut wall, 284, 286*f*
pulmonary, 560
Medial brainstem motor pathways, 99, 99*f*
Medial geniculate nucleus, 170
Medial longitudinal fasciculus (MLF), 114
Medial preoptic area, 445
Median eminence, 431
Medulla
function of, 53*f*, 54
reticular nuclei of, 100
Megakaryocytes, 182
Meiosis, 474
Meissner plexus, 284. *See also* Submucosal plexus.
α-Melanocyte-stimulating hormone (α-MSH), 389
Melatonin, 446, 467
Membrane fusion, in compartmentalization and transport, 14–15, 15*f*
Membrane potential, resting, 13, 60, 60*f*
Membrane proteins
channel *vs.* carrier, 11, 12*f*
functions of, 10–11, 10*f*, 10*t*
in membrane structure, 6, 7*f*
Membranous labyrinth, 109, 166–167
Membranous-phase digestion
intestinal surface membrane and, 322
of peptides, 323, 324*f*
of polysaccharide, 322, 323*f*
in unstirred water layer, intestinal mucus and glycocalyx microenvironment, 322, 323*f*
Meninges
anatomy of, 53*f*, 54
cerebrospinal fluid and, 150*f*, 151
Meningitis, 155*b*–156*b*
Menopause, 445
Menstrual cycle, 442–449, 443*f*
Menstruation, 442
Mesangial cells, 484
Mesenchyme, 460, 580
Mesonephros, 428–429
Messenger ribonucleic acid (mRNA), 397, 398*f*
Metabolic acidosis, 590, 590*f*
ammonia excretion during, 511–513, 512*f*
diabetic ketoacidosis and, 417
ventilatory regulation and, 564, 564*f*
Metabolic alkalosis, 590*f*, 591
with paradoxical aciduria, 516*b*
Metabolic fuels, 362
Metabolic rate
blood flow and, 253–255, 254*f*–255*f*, 254*t*
heat production and, 598
Metabolism
absorption during, 362
amino acids in, 363, 365*t*
fatty acids in, 363
glucose in, 362–363
ketone bodies in, 363, 363*f*
of adrenocortical hormones, 404
calcium and phosphate, 421–425
calcitonin and, 421–422, 424*f*
parathyroid hormone and, 421–422, 422*f*
vitamin D and, 423–425, 424*f*
Metabolism *(Continued)*
catecholamines and, 410–412, 411*f*, 412*t*
glucocorticoids and, 404, 404*t*
heat production and, 598, 598*f*, 598*t*
insulin and, 413, 414*f*, 415*b*, 415*t*
lungs and, 573
steroid hormones, 378, 384
thyroid hormones and, 397–399, 398*f*
Metastasis, definition of, 33
Metastatic cascade, 44, 45*f*
Metestrus, 442, 443*f*
Methane, fermentative production of, 342, 343*f*
Methanobacterium ruminantium, 342–343
Methemoglobin, 554
Micelles, 332, 334*f*–335*f*
Microbial ecosystem, of fermentative digestion, 340–341
Microbial protein, 345*f*, 346
Microbial yield, 345–346
Microcirculation. *See* Capillaries.
Microcotyledon, 576, 577*f*
Microglia, 158–159
Micropuncture studies, 534, 535*f*
Microvilli, 317, 318*f*
of proximal renal tubule epithelium, 489–490, 490*f*
Micturition, 134–135
Midbrain, 53*f*, 54
Middle ears, 166, 167*f*
Middle latency recordings, 163
Migrating motility complex (MMC), 299
Migrating myoelectric complex, 300
Milk
composition of, 458, 466, 466*t*
insufficient development of, in pregnant mares, 469*b*
prepartum secretion of, 459, 461
production of. *See* Lactation; Lactogenesis.
removal of, 458, 463, 465*f*
Milk letdown, 463, 465*f*
Mineral metabolism, 379
Mineralocorticoids, 403
effects of, 407–408, 407*f*, 407*t*
potencies of, 404*t*
Minute ventilation, 520
Mitochondrial acetyl CoA, 374
Mitochondrion, 41
Mitosis phase, of cell cycle, 31, 31*f*
Mitral insufficiency, 573*b*–574*b*
Mitral regurgitation, 222, 223*f*, 225
Mitral stenosis, 217, 226
Mixed venous blood, 543
MLF. *See* Medial longitudinal fasciculus.
MMC. *See* Migrating motility complex.
Monoamine oxidase (MAO), 129–130
Monoblasts, 611, 613*f*
Monocytes, 611–613, 613*f*
Monoiodotyrosine, 395
Monosaccharides, 7
Morula, 451
Mossy fiber, 121
Motilin, 289*t*
Motility, 293
Motor cortices, 103–104, 104*f*
focal lesion of, 107*b*
Motor learning, cerebellum in, 123
Motor neurons
alpha. *See* Alpha (α) motor neurons.
gamma. *See* Gamma (γ) motor neurons.
lower, 93–96, 94*f*, 104–105
disease of, 93–94, 94*f*
practice questions on, 95–96
pools, 77–78, 78*f*, 98, 98*f*
in reflex arcs, 82
upper, 94, 94*f*
disease of, 94–95
practice questions on, 95–96

Motor system, 97, 98*f*
  autonomic. *See* Autonomic nervous system.
  basal ganglia in, 97, 106–107, 106*f*
  brainstem in, descending motor pathways of, 99–100, 99*f*
  cerebellum in, 97, 106–107, 106*f*
  corticospinal tract, 97, 101–102, 102*f*
  motor cortices in, 103–104, 104*f*
  practice questions on, 108
  sensory system and, 105–106
  somatic, 126–127, 127*f*
  spinal cord in, 97, 98*f*
Motor units, 77–78, 78*f*, 97
Movement, 97–108. *See also* Motor system.
  cerebellar disease and, 123
  eye, 121, 122*f*
  forms of, 97
  fractionation of, 103
  hierarchical organizational of, 97, 98*f*
    basal ganglia, 97, 106–107, 106*f*
    brainstem, descending motor pathways of, 98–99, 99*f*
    cerebellum in, 97, 106–107, 106*f*
    corticospinal tract, 97, 101–102, 102*f*
    motor cortices in, 103–104, 104*f*
    spinal cord, 98, 98*f*
  inappropriate, in upper motor neuron disease, 94
  limb, cerebellum and, 121–123, 122*f*
  sequences, 122*f*, 123
  skeletal muscle and, 73–80, 74*f*
mRNA. *See* Messenger ribonucleic acid.
Mucociliary clearance, 569
Mucociliary system, particle deposition onto, 569–570, 570*f*
Mucous neck cells, 309
Mucus
  intestinal, 322, 323*f*
  respiratory tract, 570, 570*f*
Müllerian ducts, 428
Multiple-drug resistance (MDR), 47–48
Multiunit smooth muscle, 79
Muscarinic acetylcholine receptors, 130
Muscarinic cholinergic agonist, 197
Muscarinic cholinergic antagonist, 197, 197*t*
Muscarinic cholinergic receptors
  in cardiac regulation, 194
  in cardiovascular regulation, 263, 264*t*
Muscle
  cardiac. *See* Cardiac muscle.
  contraction of, calcium and, 421
  ischemia during, 257, 258*f*
  physiology of, 73–81, 80*b*–81*b*
  respiratory, 523
  skeletal. *See* Skeletal muscle.
  smooth. *See* Smooth muscles.
  types of, 4, 73
Muscle fibers, 73
  acetylcholine and, 68–70
  cardiac, 186
  intrafusal *vs.* extrafusal, 86
Muscle pump, 279, 279*f*
Muscle spindle stretch receptors, 562–563
Muscle spindles, 87, 87*f*–88*f*
Muscle stretch (myotatic) reflex, 88–89, 88*f*
Muscle tone, coordination of, 121–123, 122*f*
Muscular vessels, 241
Mutations, in cancer, 30–31
Myasthenia gravis, 71*b*–72*b*
Myelin, 126–127
Myelin sheath, 51
Myelography, 149
Myenteric plexus, 284, 286*f*
Myocardial infarction, 177
Myoepithelial cells of mammary gland, 463, 464*f*
Myofibrils, 73, 74*f*, 186
Myogenic hypothesis, 256
Myogenic reflex, 484
Myoglobin, 557
Myosin, 4, 5*f*
Myosin kinase, 21, 22*f*
Myosin phosphatase, 21
Myotatic reflex, 88–89, 88*f*

N

Nares, airflow resistance of, 526
Nasal insufflation, 239*b*
Necrosis, 39, 173
NEFAs. *See* Nonesterified fatty acids.
Negative-feedback control, 384–385, 385*f*
Neonatal isoerythrolysis, 469*b*
Neonates
  carbohydrate stores in, 463–466
  colostrum ingestion in, 461
  diarrhea in, 592*b*–594*b*, 593*f*–594*f*
  digestion in, 335–336
  hemolytic crisis in, 468
Nephrogenic diabetes insipidus, 506–507
Nephrons, 480, 481*f*
  short-loop and long-loop, 502–503, 503*f*
  superficial, 502, 503*f*
Nernst equation, 13
Nerve deafness, 171
Nervous system, 51–56. *See also* Autonomic nervous system; Enteric nervous system; Neurons.
  endocrine system and, 379
  function of, 51–52, 53*f*, 55*f*
  practice questions on, 56
  subdivisions of, 51–52, 52*b*, 52*f*. *See also* Central nervous system; Peripheral nervous system.
Net change, 8
Net dipole, 205–206
Net increase, 368
Net resistance. *See* Total peripheral resistance.
Neural stalk, 431
Neural-hormonal mechanisms, of cardiovascular control, 262–263
Neurocrines, of gut wall, 284, 286*f*
Neurohumoral regulatory molecules, 288
Neurohypophysis, 431
  embryonic development of, 386*f*
  hormones of, 385–386
Neurological disease
  in goats, 164*b*–165*b*
  in horses, 55*b*
Neuromuscular junction, 66, 67*f*
  anatomy of, 66–68, 67*f*
  in neuron-to-neuron synapses, 70–71
Neurons, 57–65, 63*b*–64*b*
  action potentials of, 61–62, 62*f*–64*f*
  anatomy of, 51, 52*f*, 57–58, 58*f*–59*f*
  central command, 566
  motor. *See* Motor neurons.
  neurosecretory, 379*f*
  oxygen delivery to, 176–177
  preganglionic *vs.* postganglionic, 126, 127*f*
  premotor, 98, 98*f*
  resting electrical membrane potential of, 59–60
  resting membrane potential of, 60, 60*f*
  sensory
    in reflex arcs, 82, 83*f*
    of vestibular system, 109–110, 112*f*
    visceral afferent, 52
  synaptic transmission and, 66, 67*f*
Neuron-to-neuron synapses, neurotransmitters in, 70–71
Neurophysins, 431
Neurosecretory neurons, 379*f*
Neurotransmitters, 379, 379*f*. *See also specific neurotransmitters*.
  in neuromuscular junction, 66, 67*f*, 70*b*
  in neuron-to-neuron synapses, 70–71
  in peripheral autonomic function, 131
Neurovascular unit, 155, 155*f*
Neutrophils, 611, 612*f*
Nexus, 188
Nicotine, 130
Nicotinic acetylcholine receptors, 70
Nicotinic cholinergic receptors, 187*t*, 263
Nipple, 459
Nitric oxide
  in nonadrenergic noncholinergic inhibitory nervous system, 527
  in pulmonary artery vasodilation, 535
  sodium transport and, 498
  in vascular resistance regulation, 256
  vasodilation and, 264
Nitrite poisoning, 554
Nodes of Ranvier, 57
Noncommunicating hydrocephalus, 154, 154*f*
Nonesterified fatty acids (NEFAs), 372
Nonglandular mucosa, 309
Noninfectious antigens, 608, 609*f*
Nonketotic hyperosmolar diabetes, 416
Nonpropulsive pattern, 300
Nonruminant period, 355
Nonshivering thermogenesis, 598, 604
Norepinephrine
  cardiac contractile force regulation by, 196
  in cardiac regulation, 194, 194*f*
  in cardiovascular regulation, 263–265, 264*t*
  cardiovascular transport of, 174–175
  synthesis of, 3–4, 4*f*, 409
Nuclear receptors, steroid hormone interactions and, 25–26, 26*f*
Nuclear sclerosis, 140
Nuclei, 431
Nucleic acids, 608
Nursing, first, 458, 463–466
Nutrient utilization
  during absorptive phase, 363, 364*f*
  postabsorptive phase during, 369–372, 369*f*
  during prolonged energy malnutrition, 372–375
Nutrients. *See also* Absorption, intestinal; Digestion; Food.
  cardiovascular transport of, 174–175
  placenta and, 576
  in plasma, 181, 181*t*
Nystagmus, 115

O

Obturator nerve paralysis, 91*b*
Ocular fundus, 139*f*
Oligodendrocytes, 57, 158–159
Oligosaccharides, 319–320
Omasal function, 353
Omasal transport failure, 353
Omasum, 353
Oncogenes, 33
  growth factor/oncogene pathway, 35
Oncotic pressure, 9
Oocyte proliferation, 433
Oocyte-inhibiting factor, 436–437
Opsonins, 611
Optic chiasm, 144
Optic disk, 138, 139*f*
Optic nerve, 144
Optic radiations, 144
Optic tracts, 144
Organ of Corti, 167
Organa vasculosum, of lamina terminalis, 605
Organic phosphate, 421

Osmolality, renal medullary adaptation to, 504–505, 504*f*
Osmolar (osmotic) concentration, 9
Osmolytes, organic, 507
Osmoreceptors, 387
Osmosis, 8, 8*f*, 244
Osmotic coefficient, 9
Osmotic pressure, 9–10, 244
Ossicles, auditory, 166, 167*f*
Osteoblasts, 421
Osteocytes, 421
Otolith organs, 109–110
Outer ears, 166, 167*f*
Ovarian cycle, 436–442
  inadequate nutrition in cattle and, 448*b*
  lactation and, 447
Ovarian follicle, 433–434
Ovarian follicular phase, 434
Ovarian remnant syndrome, 434*b*–435*b*
Overhydration, 387
Oviductal reservoirs, 450
Ovulation, 428, 434, 436–437, 437*f*
  practice questions on, 441
  stimulating, in mares, 448*b*
Oximetry, 554
Oxygen
  alveolar exchange of, 540–541
  consumption of, 518, 520*f*
    during exercise, 556–557, 557*f*
  diffusion of, 542–544, 542*f*–544*f*
Oxygen tension, ventilatory control and, 563–564, 563*f*
Oxygen transport, 175, 176*f*, 550–555
  during exercise, 556–557
  in fetus, 576–583, 577*f*, 577*t*, 579*f*
  hemoglobin and, 176, 182, 183*f*, 550–551, 551*f*
  in plasma, 550, 551*f*
  practice questions on, 558–559
  veterinarian's management of, 554–555, 555*t*
Oxyhemoglobin, 181
Oxyhemoglobin dissociation curve, 546, 551–553, 552*b*, 552*f*–553*f*
Oxyphil cells, 421–422
Oxytocin, 385–386, 431
  effects of, 387, 388*f*
  luteal, 439
  in milk removal, 463, 464*f*–465*f*
  in parturition, 454, 455*f*
  synthesis and release of, 378, 387*f*

P

P waves, 205, 206*f*, 207
P53 proteins, 37–38, 39*f*
  in intrinsic pathway of apoptosis, 39, 41*f*
PACAP. *See* Pituitary adenylate cyclase-activating peptide.
Pacemaker cells, 188, 191*f*
  action potentials and, 193–194, 193*f*
Pacemaker potential, 193, 193*f*
Pacemakers, cardiac
  artificial, 197–198
  atrioventricular node. *See* Atrioventricular (AV) node.
  ectopic, 198, 200*f*
  sinoatrial node, 188
Pain
  referred, 133
  weight loss with, in horses, 314*b*
Pancreas, hormones of, 394, 412–420, 413*f*. *See also specific hormones*.
Pancreatic atrophy, juvenile, 337*b*
Pancreatic exocrine insufficiency, 337*b*
Pancreatic islets, 420*f*
Pancreatic polypeptide, 420
Pancreatic secretions, 311–312
  exocrine, 311
  intestinal phase of, 311–312
  pancreatic cells, 311–312, 312*f*
Pancreatic tumor, 426*b*
Pancreatitis, in dog, 314
Panting, 601
Paracellular absorption, 325
Paracellular pathway, 490
Paracrine effectors, 379, 379*f*
Paracrines, 288
  blood flow regulation and, 254*t*, 256
  distribution of, 289*f*
Paradoxical aciduria, 516*b*
Parafollicular cell, 395
Parallel fibers, of cerebellum, 119, 120*f*
Parallel processing, 105–106
Paralysis
  in lower motor neuron disease, 93
  of obturator nerve post-calving cows, 91*b*
Parasites, acquired cardiovascular dysfunction and, 174
Parasitism, anemia and, 184*b*
Parasympathetic nervous system, 127, 286, 287*f*
  brainstem and sacral spinal cord, 128
  bronchoconstriction and, 527, 528*f*
  cardiovascular regulation by, 194–195, 194*f*–195*f*, 196*t*, 263–264, 264*t*
  salivary glands and, 308
Parathyroid hormone (PTH)
  in calcium/phosphorus homeostasis, 421–422, 422*f*–423*f*
  phosphate uptake and, 498
  renal calcium reabsorption and, 498–499
Paravertebral sympathetic ganglion chain, 127
Parenchyma, 459
Paresis, 93
Parietal cells, of gastric glands, 309–310
Parietal mucosa, 309
Parietal pleura, 525, 572–573
Parkinson's disease, 106
Parotid glands, 308
Paroxysmal atrial tachycardia, 221
Pars distalis, 389, 431
Pars intermedia, 389, 431
  embryonic development of, 385, 386*f*
Pars nervosa, 431
Pars tuberalis, embryonic development of, 385
Partial pressure of gases, 540
Partial thromboplastin time (PTT), 182
Particulates, inhalable, 568–569, 569*f*
Parturition, 450–457, 454*f*–455*f*
Passive electrotonic current, 62
Passive immunity, colostrum and, 461
Passive transport, 8, 10, 14
  of potassium, 13–14
Patent ductus arteriosus (PDA), 223, 223*f*, 226, 258*b*–260*b*, 581*b*–582*b*
Pectin, 341
Pendrin, 496
Penis, 430
Pentose-phosphate pathway, 362–363
Pepsin, 310, 321*t*
Pepsinogens, 310
Peptidases, 384
Peptide hormones, 380, 380*f*
  blood transport of, 381–382
Percent saturation, of hemoglobin, 553, 553*f*
Perfusion pressure, 175, 175*f*
  blood flow and, 232*f*
  of pulmonary circulation, 178–179
  of systemic circulation, 178–179, 230
  vascular resistance and, 232–233
Periaqueductal gray, 134, 134*f*
Peribronchial plexus, 537
Perilymph, 109, 166–167
Peripheral edema, 27, 27*b*
Peripheral nervous system (PNS). *See also* Autonomic nervous system.
  organization of, 51–52, 52*b*, 53*f*
Peripheral tissues, fuel mobilization in, 370
Peristalsis, 296
Peritubular capillary, 490–491, 491*f*
Permselectivity, of glomerular filtration, 483
Persistent luteal phase, in mares, 440*b*–441*b*
pH
  definition of, 584
  hydrogen ion concentration measurement as, 584–585, 586*f*–588*f*
  intracellular, 588–589
  normal range of, 584
  of urine, 513
Phagocytic cells, 611–613, 612*f*–613*f*
Phagolysosomes, 611
Pharynx, 295–296
Pheromones, reproductive cycles and, 447
Phosphate, proximal tubule uptake of, 498
Phosphatidylinositol (PI), 24–25, 24*f*
Phosphatidylinositol 4, 5-bisphosphate, 24–25, 24*f*
Phosphofructokinase, 369–370
Phospholipase, 24–25, 332
Phospholipids, 6, 7*f*
  bile acid formation from, 312–313, 313*f*
Phosphorus, in diabetic ketoacidosis, 416–417
Phosphorylation
  allosteric shape change of proteins and, 2, 3*f*
  of enzymes, 370*f*
  of tyrosine hydroxylase, 4
Photoperiod
  milk yield and, 467–468, 468*f*
  puberty and, 444
  reproductive cycles and, 446–447, 446*f*
Photopigments, 141
Photoreception, 141–143, 142*f*–143*f*, 143*t*
Photoreceptor, 140, 140*f*, 142*f*
Photorefractoriness, 446
Phrenic nerve, 523
Physiologic dead space, 520
Physiological regulation, 1–28, 27*b*
  biological membranes and, 6–7, 7*f*
  end-product inhibition of, 4
  enzymatic reactions and, 3–4, 3*f*
  information transmission and transduction, 15–26
    calcium ion transport in, 20–21, 22*f*
    cell signaling mechanisms, 15–16, 16*f*
    cyclic AMP in, 21–24, 23*f*
    G-protein-coupled receptors and, 17–18, 17*f*, 18*t*, 19*f*
    phosphatidylinositol 4, 5-bisphosphate in, 24–25, 24*f*
    steroid hormones in, 25–26, 26*f*
  of muscle contraction, 4–6
  practice questions on, 27–28
  proteins and, 2, 3*f*
  transport. *See* Transport.
Physiology, definition of, 1
Pia mater, 54
Pigs
  influenza in, 606*b*
  malignant hyperthermia in, 80*b*–81*b*
  salt toxicity in, 63*b*–64*b*
Pili, as antigens, 608
Piloerection, 600, 604
Pineal gland, 446
Pinocytosis, 243
Pitch, of sound waves, 166
Pituitary adenylate cyclase-activating peptide (PACAP), 284–286
Pituitary dwarfism, 391–392
Pituitary gland, 385–392, 431, 431*f*
  hormones of, 385–386, 386*f*, 389*t*. *See also specific hormones*.

Pituitary gland *(Continued)*
hypothalamic releasing hormones and, 390*f*
hypothalamic-hypophyseal portal system and, 178
Pituitary pars intermedia dysfunction (PPID), 392*b*–393*b*
PKA. *See* Protein kinase A.
PKC. *See* Protein kinase C.
Placenta
endocrine function of, 452–453, 453*f*
gas exchange at, 576, 577*t*
species variations in, 576–577, 577*f*
Placentitis, 581*b*–582*b*
Plants cell walls, 341–342
Plasma
carbon dioxide transport in, 555–556, 555*f*
components of, 10*t*, 180, 181*f*, 181*t*, 183*f*
oxygen transport in, 550, 551*f*
proteins in. *See* Plasma proteins.
Plasma clearance rate, 486–487
Plasma membranes
passive transport of potassium across, 13–14
of skeletal muscle, 73
Plasma oncotic pressure, 244
Plasma osmolality, antidiuretic hormone secretion and, 387–389, 388*f*
Plasma proteins, 180, 181*f*, 181*t*, 183*f*
hormones transported in, 396–397
in interstitial fluid, 245
Platelet aggregation, 182
Platelets, 181–182, 181*f*, 182*t*
Pleural cavity, 524*f*, 525
Pleural effusion, 248, 273–274
Pleural fluid, 572–573
Pleural space, 572–573
Pleuritis, 572–573, 573*b*–574*b*
Plicae circulares, 317
Pluripotent stem cells, 181–182
PMNs. *See* Polymorphonuclear neutrophils.
Pneumonia, right-to-left vascular shunts and, 546–547
Poikilotherms, 596–597
Poiseuille's law, 233
Polar molecules, definition of, 7
Polarization, 329–330
Polycythemia, 183, 552–553
Polymorphonuclear neutrophils (PMNs), 570–571
Pons, 53*f*, 54
Pontine micturition center (PMC), 134
Pontine respiratory group, 560, 562*f*
Population coding, by sensory receptors, 82
Portal systems, 178
hypothalamic-hypophyseal, 178
renal, 178
Portal vein, 178
Positive-feedback systems, 385, 432
Postabsorptive phase, 361–377
hepatic metabolism, 367*f*, 369–370, 370*f*
during nutrient utilization, 369–372, 369*f*
practice questions on, 376–377
Posterior chamber, 138
Postganglionic neurons, 126–127
Postreceptor cell responses, 383, 383*f*
Postrotatory nystagmus, 114–115
Postsynaptic potentials, 60–61, 61*f*–62*f*
Posture, standing, 277–278
Potassium
acid-base homeostasis and, 591–592
intestinal absorption, 328–329, 328*f*
passive transport of, 13–14
renal reabsorption of, 496–497, 497*f*
renal secretion of, 497
PP interval, 207, 207*f*
PR interval, 207, 207*f*
Preantral follicle, 434
Pre-Bötzinger complex, 560, 562*f*
Precapillary sphincters, 241
Precordial leads, 208
Preganglionic neurons, 135
Pregnancy, 450–457
corpus luteum in, 437–438, 451–452, 451*f*–452*f*
embryo development, 450–451, 453*f*
endocrine function of placenta in, 452–453
practice questions on, 456–457
prolonged, 456*b*
Pregnenolone, 380–381, 382*f*, 452–453
Premature contraction, 198
Premotor cortex, 103, 104*f*
Premotor neurons, 98, 98*f*
Preovulatory surge, of gonadotropins, 436–437, 438*f*
Preprohormones, 380, 380*f*
Prerenal azotemia, 499*b*–500*b*
Presbycusis, 171
Presbyopia, 140
Presynaptic terminals, of axons, 57–58, 58*f*
Primary hypoparathyroidism, 425
Primary motor cortex, 103, 104*f*
Primary sensory receptor, 82, 83*f*
Primary visual cortex, 144, 144*f*
Primate chorionic gonadotropin, 389
Primates
female libido in, 446
pregnancy in, 451–452, 452*f*
reproductive senescence in, 442, 445
Proestrus, 442, 443*f*
Progesterone
placental production of, 452
secretion of, 437–438
sexual behavior and, 445–446
Prolactin (PRL), 389–390, 389*t*, 433, 433*f*
gestational levels of, 453
in lactogenesis, 460–462, 461*f*, 464*f*
secretion of, 462
Prolactin-inhibiting hormone (PIH), 391*t*
Prolactin-releasing factor (PRF), 391*t*
Proliferation, cellular, cancer and, growth factor pathway and, 33–37
Proopiomelanocortin, 389, 389*f*
Propionic acid, 342
Proprioceptive positioning reaction, 104
Propulsive activity, 300
Prostacyclin, 256
Prostaglandin $F_{2\alpha}$
corpus luteum regression and, 438–439, 439*f*
ovulation and, 437
in parturition initiation, 453–455, 454*f*–455*f*
uterine synthesis of, 439
Protein hormones, 378
blood transport of, 381–382
postreceptor cell responses to, 383, 383*f*
receptors for, 383
synthesis of, 380–381, 380*f*
Protein kinase A (PKA), 21–23
Protein kinase C (PKC), 25
Proteins
binding sites of, 2, 3*f*
in colostrum, 461, 463*t*
digestion and, 311
functions of, 2–3, 3*f*
mediation of physiological change, 2
membrane
channel *vs.* carrier, 11, 12*f*
functions of, 10–11, 10*f*, 10*t*
in membrane structure, 6–7, 7*f*
metabolism of, insulin and, 415*t*
in milk, 462–463, 466
in muscle contraction, 4–6
in phospholipid bilayer structure, 6–7
plasma, 180, 181*f*, 181*t*, 183*f*
hormones transported in, 396–397
interstitial fluid and, 245
shapes of, function and, 2–3, 3*f*
surfactant, 524–525
Proteoglycan, 571
Prothrombin time (PT), 182
Proton pump, 309–310
Proto-oncogenes, 33
Protozoa and fermentative digestion, 340–341
Proventriculus, 303
Proximal muscles, 99–100
Proximal renal tubules
reabsorption by, 489–493, 490*f*–493*f*, 498, 513
secretion by, 493–496, 513
water reabsorption in, 501
Pseudopregnancy, 440*b*–441*b*
Psychogenic effects, in cardiovascular control, 268, 268*f*–269*f*
PT. *See* Prothrombin time.
PTH. *See* Parathyroid hormone.
PTT. *See* Partial thromboplastin time.
Puberty
female, 443–445, 444*f*
male, 477–478
mammary gland development and, 460–461, 461*f*–462*f*
Puerperal tetany, 468–469
Pulmonary arteries
blood pressure in, 178
smooth muscle, 534–535
species variation in, 532–533
vasoconstriction of, 535–536, 535*f*–536*f*
Pulmonary blood flow, 532–539
Pulmonary blood pressure, 178–179
Pulmonary capillaries, 572
Pulmonary circulation, 230–240, 532–537
alveolar hypoxia and, 535–536, 536*f*
arrangement of, 177–178, 177*f*
blood flow distribution in, 534, 535*f*
blood vessels of, 533, 533*f*
chemical mediators of, 534–535, 536*f*
exercise and, 537
fetal, 577–579, 578*f*
perfusion pressure of, 178–179, 234–236
practice questions on, 538–539
resistance in, 533–534, 534*f*
species variations in, 532–537
Pulmonary edema
formation of, 572
mitral regurgitation and, 225
Pulmonary embolism, heartworm disease with, 239*b*
Pulmonary fluid exchange, 571–573, 572*f*
Pulmonary microvasculature, 537
Pulmonary stretch receptors, 562
Pulmonary surfactant, 524, 525*f*, 580
Pulmonary vascular pressures, 534
Pulmonary vascular resistance, 534, 534*f*, 536*f*, 580–581
Pulmonary veins, 532–533
Pulmonary venous pressure, 534
Pulmonary wedge pressure, 534
Pulmonic stenosis, 222, 223*f*, 226*b*–227*b*
Pulse pressure, 236–238
Pump failure, 224
Pupil, 138, 145
Pupillary light reflex, 145, 146*f*
Purkinje cell layer, 119, 120*f*
Purkinje fibers, 188–189
Pyloric mucosa, 309
Pylorus, 297–298
Pyramidal cells, 158, 159*f*
Pyramidal decussation, 102–103
Pyramidal motor system, 98–99
Pyramidal tract. *See* Corticospinal tract.
Pyruvate, 362–363, 362*f*

Q

Q waves, 205, 206*f*, 207
QRS complex/waves, 206–207, 206*f*
QT intervals, 207, 207*f*

R

R waves, 205–206, 206*f*
Rabies, in horses, 304*b*–305*b*
Rabs protein, 18
Radiant heat exchange, 597
RAO. *See* Recurrent airway obstruction.
Ras protein, 18, 34
Rathke's pouch, 385, 386*f*
Reabsorption, 245, 276–277, 277*f*
Reaction coupling, proteins and, 2
Reactive hyperemia, 254–255, 255*f*
Receptor tyrosine kinases (RTKs), 16–17
  growth factor activation of, 33–34
Receptor/ligand complex, and specificity of information, 17
Receptor-mediated endocytosis, 14–15
Receptors. *See also specific receptors*.
  adrenergic, 263
  airway, 562
  in autonomic control of cardiovascular system, 263, 264*t*
  chemoreceptors. *See* Chemoreceptors.
  cholinergic. *See* Cholinergic receptors.
  classes, 16–17
  dendritic, 57
  downregulation of, 17
  G-protein-coupled, 17–18, 17*f*, 18*t*, 19*f*
  irritant, 562
  juxtacapillary, 562
  membrane, 7
  nuclear, 16–17, 25–26, 26*f*
  sensory. *See* Sensory receptors.
  in signaling pathways, 16–17
Recruitment
  capillary, 544
  of muscle motor units, 78
Rectal temperature, 597*t*, 601
Rectosphincteric reflex, 302–303, 303*f*
Recumbency
  air distribution in lung during, 529
  transition to standing posture, 277–278
Recurrent airway obstruction (RAO), 530*b*
Recurrent laryngeal neuropathy, 523
Red nucleus, 101
Reentrant arrhythmias, 199
Referred pain, 133
Reflex arcs, 82
  fundamental components of, 82–83, 83*f*
  intersegmental, 83–84, 84*f*
  in nervous system, 84
  segmental, 83, 84*f*
  suprasegmental, 83–84
Reflexes, 82–85, 85*b*. *See also* Reflex arcs; *specific reflexes*.
  cardiovascular, 263. *See also* Arterial baroreceptor reflex; Atrial volume receptor reflex.
  definition of, 82
  exercise, 278–279
  Hering-Breuer, 562
  intersegmental, 93–94
  knee jerk, 88–89, 88*f*
  muscle stretch (myotatic), 88–89, 88*f*
  myogenic, 484
  practice questions on, 85
  segmental
    in lower motor neuron disease, 93–94
    in upper motor neuron disease, 94–95
Refractory period, of cardiac muscle, 192, 192*f*
Relative humidity, 603–604
Relaxin
  in parturition, 454–455
  in pregnancy, 451, 453
Renal failure
  chronic, 487*b*–488*b*
  heart failure and, 274, 274*f*
Renal insufficiency, chronic, 507*b*–508*b*
Renal medulla
  hypotonicity in, 502
  osmolality adaptations of, 507
Renal portal system, 178
Renal tubules
  in acid-base balance, 588
  distal tubule. *See* Distal renal tubules.
  functional assessment of, 489
  practice questions on, 500
  proximal. *See* Proximal renal tubules.
  reabsorption of filtered substances, 489, 490*f*
Renin, 407–408, 484
Renin-angiotensin system, 408, 408*f*
Renin-angiotensin-aldosterone system, 274
  glomerular filtration rate and, 484
  volume receptor reflex activation of, 267
Repolarization
  atrial, 207
  ventricular, 206*f*, 207
Reproductive cycles, 443–444. *See also* Estrous cycle; Ovarian cycle.
  lactation and, 447
  pheromones and, 447
  photoperiod and, 446–447, 446*f*
  practice questions on, 449
  puberty, 443–445, 444*f*
  sexual behavior, 445–446, 445*f*
  types of, 442–443, 443*f*
Reproductive senescence, in primates, 445
Reproductive system
  development of, 428–435, 429*f*
  female. *See* Female reproductive system.
  male. *See* Male reproductive system.
  practice questions on, 435
Reservoir vessels, 241
Residual volume, 520
Respirable particulates, 568–569
Respiratory acidosis, 589
  with renal compensation, 516*b*
Respiratory alkalosis, 589–590, 590*f*
Respiratory burst, 611
Respiratory cycle, 520, 520*f*
Respiratory exchange ratio, 541, 541*b*
Respiratory function, overview of, 518–531, 519*f*. *See also* Ventilation.
Respiratory muscles, 523, 562–563
Respiratory pump, 279
Respiratory system. *See also* Lungs; Tracheobronchial tree; Upper airway; Ventilation.
  in acid-base balance, 591
  anatomy of, 519*f*
  dead space of, 520, 521*f*
  defense mechanisms of, 568–571, 569*f*–571*f*
  feedback control of, 560, 561*f*
  practice questions on, 575, 582–583
Resting membrane potential, 13, 60, 60*f*
Restorative nervous system, 131
Rete ovarii, 433
Reticular activating system, 160–163
Reticular formation, 54, 100
Reticular groove. *See* Esophageal groove.
Reticulorumen. *See also* Fermentative digestion.
  development of, 354–355
  dilution rate of, 339, 352
  feed intake rate, 351
  motility
    control of, 352–353
    maintenance of rumen environment and, 347–352
  physiological functions of, 347
  practice questions on, 359–360
  water movement, 351
Reticulospinal tracts, 98–99, 99*f*
Retina
  anatomy of, 139–140
  cell types of, 140, 140*f*
Retina *(Continued)*
  diagnostic testing of, 145
  fovea, 140–141, 141*f*–142*f*
  tapetum, 139*f*, 141
  visual image processing in, 143–144
Retinal ganglion cells, 144–145
Retinoblastoma, 37–38, 38*f*
Retino-geniculo-striate pathway, 144, 144*f*
Retino-hypothalamic pathway, 144
Retino-tectal pathway, 144
Retropulsion, in colon, 301
Rhodopsin, 142
Ribonucleic acid (RNA), 16–17
Right-to-left vascular shunts, 546–547, 547*f*
Rigor mortis, 5
Rods, 140, 140*f*
  photoreception by, 141–143, 142*f*, 143*t*
RR interval, 207, 207*f*
RTKs. *See* Receptor tyrosine kinases.
Rube Goldberg device, 31
Rubrospinal tracts, 98–99, 99*f*, 101
Rumen. *See* Reticulorumen.
Rumen fermentation, 347–351, 347*f*–350*f*
Ruminants
  gluconeogenesis in, 375, 375*f*
  saliva of, 308–309, 308*f*
Rumination (cud chewing), 351

S

S waves, 206, 206*f*
SA node. *See* Sinoatrial (SA) node.
Saccule, of vestibular system, 110–112, 113*f*–114*f*
Sacral spinal cord, PNS and, 128–129, 129*f*
Saliva
  enzymes of, 307–308
  ruminant, 308–309, 308*f*
Salivary amylase, 307
Salivary glands, 307–309
Salt toxicity, 63*b*–64*b*
Sarcolemma, 73
Sarcomas, 31
Sarcomeres, 73, 76–77, 77*f*, 186–187, 187*f*
Sarcoplasmic reticulum, 75–76, 75*f*–76*f*, 189
Scalae, 167, 168*f*
Schwann cells, 57
Sclera, 138
Scrotal circumference, 475, 475*f*–476*f*
Scrotum, 430, 471
Second messengers
  calcium ions as, 20–21, 22*f*
  cyclic adenosine monophosphate, 19, 23*f*
  definition of, 19
  diacylglycerol as, 19, 20*f*
  in hormone response, 383, 384*f*
  inositol 1, 4, 5-triphosphate as, 19, 24*f*
Secondary active transport, 324, 491
Secondary (or eructation) contractions, 350–351
Secondary sensory receptor, 82, 83*f*
Second-degree AV node block, 198, 211, 212*f*
Secretin, 289*t*, 298
  pancreatic secretions and, 311–312
Secretory diarrhea, 337
Segmental reflexes, 83, 84*f*, 93–94
Segments, in electrocardiograms, 207, 207*f*
Seizures
  in dogs, 48*b*–49*b*
  in foals, 155*b*–156*b*
Selective transport
  in biological membranes, 7
  membrane proteins and, 10–11, 10*f*, 10*t*
Semen, composition of, 472*t*
Semicircular ducts, 110, 112*f*
Senescence, 445
Sensorineural deafness, 171

Sensory neurons
in muscle innervation, 86
in reflex arcs, 82, 83*f*
of vestibular system, 109–110, 112*f*
visceral afferent, 52
Sensory receptors, 82, 83*f*
in joints, 90–91
in muscle, 90–91
primary *vs*. secondary, 82, 83*f*
of vestibular system, 109, 111*f*
Sensory systems, 105–106
Sensory-evoked potentials, 158–165, 164*b*–165*b*
Septic shock, 174
heart failure and, 274*f*, 275
Serial processing, 105
Serotonin, 573
platelets' release of, 182
Sertoli cells, 428–429
Serum, 182
Servo-assist function, 105
Sex hormones, 380–381
Sex pheromones, reproductive cycles and, 447
Sexual attractiveness, in spayed bitch, 448*b*
Sexual behavior, 445–446, 445*f*
Sexual receptivity, 445–446
Sherrington, Charles, 101
Shivering, 598, 604
Shock
cardiogenic, 174
endotoxic, 174, 184*b*
hemorrhagic, 174
septic, 174
Shock-lung syndrome, 174
Short-loop feedback systems, 391
Short-loop nephrons, 502–503, 503*f*
SIADH. *See* Syndrome of inappropriate antidiuretic hormone secretion.
Sick sinus syndrome, 197, 197*t*
Signaling, 2
calcium ion transport in, 22*f*
cellular mechanism of, 15–16, 16*f*
cyclic AMP in, 21–24, 23*f*
electrical, 13
G-protein-coupled receptors and, 17–18, 17*f*, 18*t*
phosphatidylinositol 4, 5-bisphosphate in, 24–25, 24*f*
steroid hormones in, 25–26, 26*f*
Sinoatrial (SA) node, 188, 188*f*, 190*f*, 194*f*
dysfunction of, 195, 195*f*, 200*f*
Sinus arrest, 197
Sinus bradycardia, 210–211, 211*f*
Sinus tachycardia, 198–199, 210–211, 211*f*
Skeletal muscle, 4, 73
action potentials and, 75–76, 76*f*
blood flow to, during exercise, 234, 278, 279*f*
body movement and, 73–80, 74*f*
contraction of, 5–6, 73–80, 74*f*, 76*f*
cardiac muscle contraction *vs*., 186–187, 187*f*, 187*t*
strength of, 77–78, 78*f*
innervation of, 78*f*
organization of, 73–75, 74*f*
oxygen delivery to, 176, 176*f*
practice questions on, 81
receptor organs, 86–92
clinical correlation, 91*b*
Golgi tendon organs, 86, 87*f*, 90, 90*f*
muscle spindles, 87, 87*f*–88*f*
Skin
in body defenses, 610
temperature-sensitive receptors in, 602, 602*f*
Slow waves, in gastrointestinal motility, 293–294, 294*f*
Slow-twitch muscle fibers, 77
Small intestine
motility of, 298–300, 299*f*
mucosa, 317–318, 318*f*
Smooth muscles, 4, 73
bronchial, air flow resistance and, 527, 528*f*
contraction of, 241
excitation-contraction coupling in, 79–80
gastrointestinal, 294–295, 295*f*–296*f*
multiunit, 79
structure of, 78–79, 79*f*
visceral, 79
Sodium
intestinal absorption, 327, 327*f*
mineralocorticoids and, 407, 407*f*
renal reabsorption of, 497, 503
Solutes
in interstitial fluid, exchange of, 241, 242*f*
reabsorption of, 489–500
Solvent drag, 490
Somatic cells, 30–31
Somatomedins, 390–391
Somatostatin, 420
Somatotopic map, 98*f*, 103, 105*f*
Somatotopic organization, in spinal cord, 98, 98*f*
Sound waves, 166, 167*f*
to the cochlea, 166, 167*f*
frequency, deciphering of, 170
transduction of, 167–170, 168*f*–169*f*
Spatial summation, of muscle motor units, 78
Sperm transport, 450
Spermatocytes, 473
Spermatogenesis, 473–475, 474*f*
Spermatogonia, 473
Spermatozoa, 450, 473, 473*t*, 475*f*
emission of, 473, 473*t*
production, testicular size and, 475, 476*t*
Spermiogenesis, 474
Sphincter of Oddi, 313
Sphincters, precapillary, 241
Spinal cord
anatomy and function of, 53–54, 53*f*
in control of movement, 98, 98*f*. *See also* Movement.
Spinal tap, 149, 151
Spinocerebellum, 121, 122*f*, 123
Spiral ganglion, 167, 168*f*
Splanchnic circulation, 178, 274, 274*f*
Splanchnic nerves, 127–128
Splenic contraction, after hemorrhage, 276
Split second heart sound, 217–218
Spontaneous nystagmus, 115
ST segment, abnormalities of, 209, 210*f*
Standing posture, 277–278
Starch, 320, 321*f*
digestion and, 311
Starling, Ernest Henry, 219
Starling equation, 241, 246–247, 571
Starling forces, 245*f*–246*f*, 247–248
balance of, 247
Starling's hypothesis, 9–10
Starling's mechanism, 272–273, 273*f*
Stellate cells, 119, 158
Stem cells, 43, 43*f*
pluripotent, 181–182
Steroid diabetes, 404
Steroid hormones. *See also specific hormones*.
adrenal cortex. *See* Adrenal cortex.
anabolic steroids, 478
metabolism of, 384
nuclear receptor interactions of, 25–26, 26*f*
postreceptor cell responses to, 383, 383*f*
receptors for, 383
sex hormones, 380–381
synthesis of, 380–381, 380*f*–381*f*
Stimulus-secretion coupling, 409
Stomach, motility of, 297
distal region in, 297–298, 298*f*
indigestible material and, 299
proximal region in, 297, 297*f*
Stomata, 572–573
Strangles, 615*b*
Stratum basale, 353
Stratum corneum, 353
Stratum granulosum, 353
Streptococcus equi, 615*b*
Stress, catecholamine release and, 412
Stretch receptors, pulmonary, 562
Stroke, 176–177
Stroke volume
cardiac output and, 218, 218*f*
definition of, 216
end-diastolic volume and, 218, 218*f*–219*f*, 221–222
heart rate and, 221–222, 221*f*
Strongylus vulgaris, 174
Subarachnoid space, 54, 151, 151*f*
Subepithelial irritant receptors, 570
Subepithelial vascular plexus, 537
Submucosal plexus, 284, 286*f*
Substance P, 284–286
Suckling, lactogenesis and, 464*f*–465*f*
Sucrase, 322
Sucrose, 466
Sudden cardiac death, 199
Sugars, 319–320
Sulfates, 397, 404
Superficial nephrons, 502, 503*f*
Superior olivary complex, 170
Supplementary motor cortex, 103, 104*f*
Suprasegmental reflex, 83–84
Supraventricular tachycardia, 198–199
Surface mucous cells, 309
Surface tension forces, 523–524
Surfactant, pulmonary, 524, 525*f*, 580
Surfactant proteins, 524–525
Suspensory ligaments
of eyes, 138
of udder, 459
Sustentacular cells, 563
Sweat glands, 601
Sympathetic cholinergic neurons, 264–265
Sympathetic nervous system, 127, 286, 287*f*
bronchoconstriction and, 527, 528*f*
cardiac regulation by, 194–195, 194*f*–195*f*, 196*t*
cardiovascular regulation by, 263
in heart failure compensation, 276
parasympathetic nervous system *vs*., 131, 132*t*–133*t*
thoracolumbar spinal cord and, 127–128, 128*f*–129*f*
Sympathetic trunk, 127
Synapses, 52, 66–72, 71*b*–72*b*
neuromuscular, 66–68, 67*f*
neuron-to-neuron, 70–71
practice questions on, 72
in reflex arcs, 82
Synaptic boutons, 66
Synaptic cleft, 57–58, 66, 67*f*
Synaptic vesicles, 57–58, 66, 69*f*
Synaptotagmin, 69*f*
Syncope, vasovagal, 263, 268, 269*f*
Syncytiotrophoblasts, 451–452
Syndrome of inappropriate antidiuretic hormone secretion (SIADH), 389
Synthesis phase, of cell cycle, 31, 31*f*
Systemic blood pressure, 178–179
Systemic circulation
arrangement of, 177–178, 177*f*
blood pressure in, 230, 231*f*
blood vessels in, 241, 242*f*
net resistance of, 234
organ blood flow regulation in, 234–236
perfusion pressure of, 178–179, 230
resistance to, 236
total peripheral resistance of, 234
vessel types in, 179*t*
Systemic edema, 248

Systole
  atrial, 217
  ventricular, 216–218, 217*f*, 223*f*
Systolic aortic pressure, 178
Systolic arterial pressure, 236
Systolic murmurs, 222, 222*t*, 223*f*

T

T cells, 617–619
  antigen-presenting cells and, 617, 619–621, 620*f*–621*f*
  cytokine secretion by, 618–619, 619*f*
  development of, 617–618, 618*f*
T wave, 206*f*, 207
Tachyarrhythmias, 198–199, 200*f*
Tachycardias, 198, 210–211, 211*f*–212*f*, 213*b*–214*b*, 214*f*
Tachypnea, 562
Tapetum, 139*f*, 141
Target organ, in reflex arcs, 82, 83*f*
TBG. *See* Thyroxine-binding globulin.
Tectorial membrane, 167
Tectospinal tracts, 98–99, 99*f*, 101
Teeth, phosphate and, 421
Telencephalon, 53*f*, 54
Telomerase, 42
Temperature, body
  core, 597*t*, 598, 601
  heat input/output and, 597, 597*f*
  practice questions on, 606–607
  rectal, 597*t*, 601
  regulation of. *See* Thermoregulation.
  tissue function and, 596
Temporal summation, of muscle motor units, 78
Testes
  anatomy and function of, 472
  embryonic development of, 429–430, 429*f*
  sperm spermatozoa and, 473
Testosterone
  in reproductive development, 429–430, 429*f*
  sexual behavior and, 446
Tetanus, 71–72
Tetraiodothyronine, 395
Tetralogy of Fallot, 546–547
Tetrodotoxin, 62
Thalamus, 54
Thecal layer, of follicle, 433–434
T-helper cells, 618–619, 619*f*–621*f*
Thermal gradient, 599–600
Thermogenesis, 598, 604
Thermoregulation, 596–607
  countercurrent mechanism in, 599, 600*f*
  fever and, 604–605, 604*f*–605*f*
  homeothermic *vs.* poikilothermic strategy for, 596–597
  hypothalamic set point and, 602, 603*f*
  peripheral vasoconstriction/vasodilation and, 603*f*, 604
  practice questions on, 606–607
  temperature-sensitive receptors and, 602
Thermoregulatory set point, 604–605
Thick filaments of muscle, 4, 5*f*, 186–187
Thin filaments of muscle, 4, 186–187
Third-degree atrioventricular node block, 198, 201*b*, 211, 212*f*
Thoracolumbar spinal cord, 127–128
Thoracolumbar system, 127
Threshold potential, 61
Thrombin, 182
Thromboxane A_2 (TXA_2), 256
Thrombus, 176–177
Thymocytes, 617–618
Thyroglobulin, 395, 396*f*
Thyroid gland, 395–401, 395*f*
Thyroid hormones, 395
  blood transport of, 382
  metabolism of, 397–399, 398*f*
  plasma transport of, 396–397
  release of, 396, 396*f*
  storage of, 396
  synthesis of, 395, 395*f*
Thyroid-stimulating hormone (TSH), 389–390, 389*t*, 397
Thyronine (T_4), 395, 395*f*, 400*t*
Thyroperoxidase, 395
Thyroprotein, 467
Thyrotropin-releasing hormone (TRH), 391*t*, 398–399, 399*f*, 433
Thyroxine, in thermogenesis, 604
Thyroxine-binding globulin (TBG), 396, 396*f*
Tidal volume, 520
Tight junctions, 14, 293–294, 317
Tissue
  body temperature and, 596
  gas exchange in, 544
Tissue pressure hypothesis, of autoregulation, 256
Titin, 73–75
TNF. *See* Tumor necrosis factor.
Toll-like receptors, 568
Tonotopic organization, 170
"Tonus" theory, 409
Total lung capacity, 520
Total peripheral resistance (TPR), 234, 263
  in blood pressure regulation, 234, 262, 266, 266*f*
Toxins, interference with action potential, 62
Trachea
  collapsing, 528
  smooth muscle contraction and, 527
Trachealis muscle, 527
Tracheobronchial tree
  anatomy of, 518
  dynamic compression of, 528
Transcalciferin, 423
Transcellular transport, 14
Transcortin, 382, 404
Transcription factors, 25–26
Transcytosis, 243
Transduction, light, 141–143, 142*f*–143*f*, 143*t*
Transit time, 293
Transitional period, 355
Transmural pressure, 528, 533–534, 534*f*
Transmural pressure difference, 175, 175*f*
Transport, 7–15
  active, 8, 10, 10*t*, 11*f*, 14, 175
  cotransport, 11
  epithelial, 14
  facilitated diffusion, 11, 12*f*
  membrane fusion and, 14–15, 15*f*
  membrane proteins and, 10–11, 10*f*, 10*t*
  osmosis and, 8, 8*f*
  passive, 8, 10, 14
    of potassium, 13–14
  proteins, 2
  in proximal renal tubule, 490, 491*f*
  selective
    biological membranes in, 7
    membrane proteins and, 10–11, 10*f*, 10*t*
  solute, 497
  spontaneous, 7–8
  Starling's hypothesis and, 9–10
  transcellular, 14
Transverse tubules, 75, 75*f*–76*f*
TRH. *See* Thyrotropin-releasing hormone.
Triacylglycerols, 363
Tricarboxylic acid cycle, 362
Triglycerides, 311, 332, 368–369
  accumulation in adipose tissue, 368–369
  in milk, 466
Triiodothyronine (T_3), 395, 395*f*, 397*f*, 400*t*
Trisaccharides, 319–320
Trophoblastin, 451
Tropomyosin, 5–6, 6*f*, 73–75
Troponin, 5–6, 6*f*, 73–75
Trypsin, 321–322, 321*t*
Trypsinogen, 321–322
TSH. *See* Thyroid-stimulating hormone.
Tubuloglomerular feedback (TGF), 485
Tumor necrosis factor (TNF), 571
Tumor suppressors, 32–33, 32*f*, 34*f*, 37–38, 38*f*
Turbulent airflow, 526
Tympanic membrane, 166
Tyrosine
  in catecholamine synthesis, 409–410, 410*f*–411*f*
  growth factors and, 33–34, 34*f*
  metabolism of, 3–4, 4*f*
  in thyroid hormone synthesis, 395
Tyrosine hydroxylase, phosphorylation, 4

U

Udders, 459, 460*f*, 469*b*. *See also* Mammary gland.
Ulcers
  corneal, colic from atropine administration to treat, 135*b*–136*b*
  gastric, 314*b*
Ultradian rhythms, 385
Umbilical arteries, 579
Umbilical blood vessels, 580–581
Umbilical veins, 579
Undernutrition, 372–373, 372*f*
Ungated leak channels, 12*f*
Uniporters, 11
Unstirred water layer, 318
Upper airway
  obstruction of, 592*b*–594*b*, 593*f*–594*f*
  structures of, 526
Upper motor neuron, 93–96, 94*f*
  disease of, 94–95
  practice questions on, 95–96
Urea, 346, 347*f*, 494, 503–504, 504*f*
  recycling
    in collecting ducts, 503–504, 504*f*
    in rumen, 347*f*, 355–356
Uremia, in heart failure, 274, 274*f*
Uric acid, 494
Urination, 134–135
Urine
  pH of, 513
  renal concentration/dilution of, 501, 502*f*
  storage reflexes, 134
Uterus
  prostaglandin $F_{2\alpha}$ production by, 451
  prostaglandin $F_{2\alpha}$ secretion by, 439, 439*f*
Utricle, 110–112, 113*f*–114*f*
Uveitis, 147*b*

V

Vagotomy, 562
Van't Hoff equation, 9
Varicosities, 126–127, 284, 286*f*
Vasa deferentia, 473
Vasa recta, countercurrent mechanism in, 505, 505*f*
Vascular resistance, 232–233, 232*f*
  blood flow and, 234–236, 235*f*
  local control of, 253
  pulmonary, 534, 534*f*
  total peripheral, 234, 235*f*
  vascular transmural pressure and, 533–534, 534*f*
Vascular shunts, right-to-left, 546–547, 547*f*
Vascular transmural pressure, 533–534, 534*f*
Vasculitis, 250*b*–251*b*
Vasoactive intestinal peptide (VIP)
  as gut neurocrine, 284–286, 433
  in lactogenesis, 462, 464*f*
  prolactin selection and, 462

Vasoconstriction, 233
  in heart failure, 274
  peripheral, in heat conservation, 604
  of pulmonary arteries, 235*f*, 535–536, 535*f*–536*f*
Vasodilation
  blood flow and, 233, 233*f*
  peripheral, 602–604, 603*f*
Vasopressin, 21, 385–386, 387*f*–388*f*
  effects of, 387
  secretion of, 387–389
Vasovagal syncope, 263, 268, 269*f*
VCR. *See* Vestibulocollic reflex.
Veins, 241
  as blood reservoir, 179, 179*t*, 263
  general layout of, 177–178, 177*f*
  portal, 178
  pulmonary, 532–533
  umbilical, 579
  walls of, 241
Velocity
  of airflow, 526
  of blood flow, 543
Venae cavae, 177–178, 179*t*
Venoconstriction, 263, 264*t*
Venous portal system, 431
Venous system, CSF absorption into, 153–154, 153*f*
Ventilation, 520–529
  in acid-base regulation, 588, 588*f*
  airway diameter and, 527, 528*f*
  alveolar, 520–521, 540–541, 544*f*–546*f*
    dead space *vs.*, 520, 547
  alveolar ventilation/perfusion ratio and, 544–546, 545*f*–546*f*
  collateral, 529
  control of, 560–567
    airway receptors, 562–563
    anesthesia and, 566*b*–567*b*
    central, 560–562, 561*f*
    chemoreceptors and, 560, 563–566, 563*f*–565*f*
    during exercise, 566
    gas exchange demands and, 560, 561*f*
    at high altitude, 565–566
    practice questions on, 567
    pulmonary receptors and, 562–563
  distribution of, 529, 529*f*
  energy requirements of, 521–523, 522*f*–523*f*
  frictional resistance to airflow and, 526–527, 526*f*–527*f*
  lung elasticity and, 523–524, 524*f*
  practice questions on, 531
  work of breathing, 527–528, 528*f*
Ventilation-perfusion matching, 236
Ventral corticospinal tract, 102–103
Ventricles (brain)
  anatomy of, 149, 150*f*
  cerebrospinal fluid and, 149–150, 150*f*–151*f*
Ventricles (cardiac), 216
  depolarization/repolarization of, on ECG, 205–207, 206*f*
  infarction of, 177
  pressure in, 218
Ventricular cells, 193
Ventricular compliance, 218–220, 219*f*–220*f*
Ventricular contractility
  depressed. *See* Heart failure.
  end-diastolic volume and, 216, 220–221, 220*f*
Ventricular contractions, premature, 210, 211*f*
Ventricular diastole, 216–218, 223*f*
Ventricular end-diastolic volume, 216
  determinants of, 218, 219*f*–220*f*
  stroke volume and, 218, 218*f*–219*f*, 221–222
  ventricular contractility and, 216, 220–221, 220*f*
Ventricular fibrillation, 199
Ventricular function curves, 218, 220*f*
  in heart failure, 273*f*, 275, 275*f*
Ventricular hypertrophy
  cardiac defects and, 224
  right, electrocardiography of, 208–209, 209*f*
Ventricular preload, 218–220, 219*f*
Ventricular repolarization, 205–207, 206*f*
Ventricular septal defect (VSD), 222, 223*f*, 226
Ventricular systole, 216–218, 217*f*, 223*f*
Ventricular tachycardia, 213*b*–214*b*, 214*f*
Ventriculus, 303
Very-low-density lipoproteins (VLDLs), 364–365, 365*f*, 368–369, 374
  hepatic, 365*f*, 374
Vestibular nuclear complex, 100
Vestibular reflexes, 114–115
Vestibular syndrome, 116*b*
Vestibular system, 109–117
  anatomy of, 109
  bilateral receptor system, 109, 110*f*
  practice questions on, 116–117
  regions of, 109–110, 112*f*
  saccule, 110–112, 113*f*–114*f*
  semicircular ducts, 110, 112*f*
  sensory information, 112–114
  utricle, 110–112, 113*f*–114*f*
  vestibular reflexes, 114–115, 115*f*
Vestibulocerebellum, 121, 122*f*, 123
Vestibulocollic reflex (VCR), 115
Vestibulospinal tracts, 98–99, 99*f*
VFAs. *See* Volatile fatty acids.
Villi, 317
VIP. *See* Vasoactive intestinal peptide.
Viruses, antigenic structures of, 608
Viscera, 126
Visceral afferent neurons, 52, 131–133
Visceral epithelium, 482, 483*f*
Visceral pleura, 525, 537
Visceral smooth muscle, 79
Viscosity, of blood, hematocrit and, 183, 183*f*
Visual cortex, 144–145, 144*f*
Visual photopigment, 141
Visual system, 138–148. *See also* Eyes; Retina.
  diagnostic testing of, 144
  photoreception and light transduction in, 141–143, 142*f*–143*f*
  practice questions on, 147–148
  visual pathways of, 144–145, 144*f*
Vital capacity, 520
Vitamin A, in colostrum, 461
Vitamin D
  in calcium/phosphorus homeostasis, 423–425, 424*f*–425*f*
  renal calcium reabsorption and, 498–499
Vitreous humor, 138
VLDLs. *See* Very-low-density lipoproteins.
Volatile fatty acids (VFAs)
  absorption of, 353–354, 354*f*
  production of, 342–343, 343*f*–344*f*
  rate of, 358, 358*f*
Voltage-gated channels, 12*f*, 13
Volume conduction, 158
Volume receptors, 267
Voluntary movement, 97
  brainstem motor pathways in, 98–99, 99*f*
    lateral, 99*f*, 101
    medial, 99–100, 99*f*
  corticospinal tract in, 97, 101–102, 102*f*
  motor cortices of, 103–104, 104*f*
  spinal cord in, 98
Vomiting
  brainstem and, 299–300
  metabolic alkalosis and, 591
VSD. *See* Ventricular septal defect.
Vulva, 430

W

Waste products
  cardiovascular transport of, 174–175
  plasma and, 181*t*
Water
  absorption of, 326–329
  cardiovascular transport of, 174–175
  intestinal secretion of, 329–330
  movement in reticulorumen, 339, 351–352
  in reticulorumen, 351–352
Water balance, 501–508
  antidiuretic hormone and, 506–507, 506*f*
  countercurrent mechanism and, 504–505, 504*f*–505*f*
  kidney and, 501
  practice questions on, 508
  proximal tubule reabsorption and, 501
  renal medullary osmolality adaptation and, 507
  sodium chloride reabsorption and, 505–506
  urea reabsorption/recycling and, 503–504
  urine concentrating ability and, 501, 502*f*
Water retention, in heart failure, 274*f*
Water vapor, partial pressure, 540
Weight loss, pain with, in horses, 314*b*
White blood cells. *See* Leukocytes.
White ramus, 127
Whitten effect, 447
Wolffian duct, 429–430, 429*f*
Work of breathing, 527–528, 528*f*

Y

Young-Helmholtz theory, 143

Z

Z disk, 73–75
Zinc fingers, 25
Zona fasciculata, 402–403, 402*f*
Zona glomerulosa, 402–403, 402*f*
Zona pellucida, 433–434
Zona reticularis, 402–403, 402*f*
Zonula occludens, 490, 491*f*
Zonular fibers, 138
Zymogen granules, 310
Zymogens, 310

This page intentionally left blank