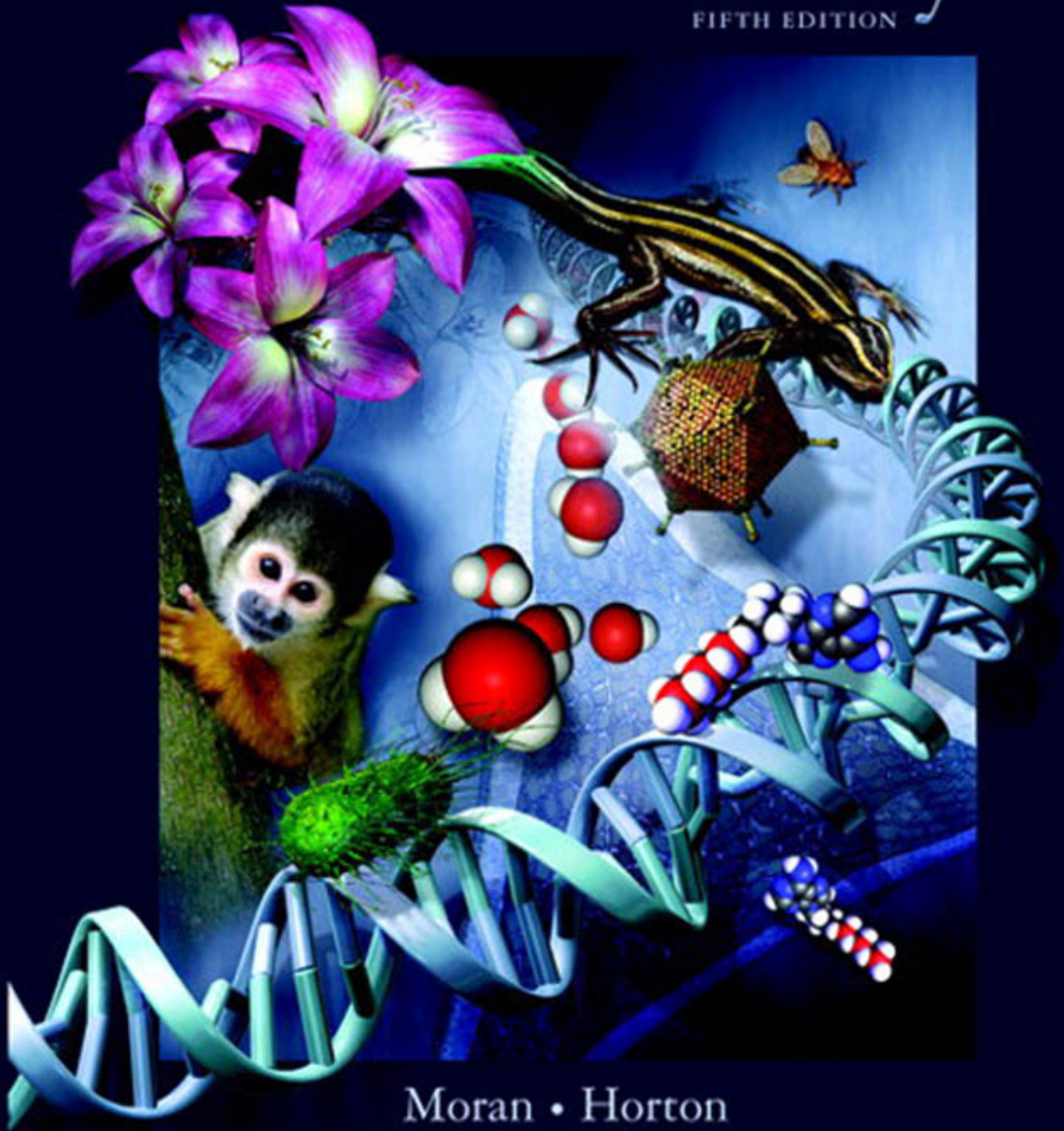
PRINCIPLES OF
Biochemistry
FIFTH EDITION



Moran • Horton
Scrimgeour • Perry

This page intentionally left blank

Principles of Biochemistry

This page intentionally left blank

Principles of Biochemistry

Fifth Edition



Laurence A. Moran
University of Toronto

H. Robert Horton
North Carolina State University

K. Gray Scrimgeour
University of Toronto

Marc D. Perry
University of Toronto

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreál Toronto
Delhi Mexico City Sao Pauló Sydney Hong Kong Seoul Singapore Taipei Tokyo

Downloaded From Vet-eBooks.com

Editor in Chief: Adam Jaworski
Executive Editor: Jeanne Zalesky
Marketing Manager: Erin Gardner
Project Editor: Jennifer Hart
Associate Editor: Jessica Neumann
Editorial Assistant: Lisa Tarabokjia
Marketing Assistant: Nicola Houston
Vice President, Executive Director of Development: Carol Truehart
Developmental Editor: Michael Sypes
Managing Editor, Chemistry and Geosciences: Gina M. Cheselka
Project Manager, Science: Wendy Perez
Senior Technical Art Specialist: Connie Long
Art Studios: Mark Landis Illustrations
/Jonathan Parrish
/2064 Design—Greg Gambino
Image Resource Manager: Maya Melenchuk
Photo Researcher: Eric Schrader
Art Manager: Marilyn Perry
Interior/Cover Designer: Tamara Newnam
Media Project Manager: Shannon Kong
Senior Manufacturing and Operations Manager: Nick Sklitsis
Operations Specialist: Maura Zaldivar
Composition/Full Service: Nesbitt Graphics, Inc.

Cover Illustration: Quade Paul, Echo Medical Media
Cover Image Credit: Monkey adapted from Simone van den Berg/Shutterstock

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on page 767.

Copyright ©2012, 2006, 2002, 1996 Pearson Education, Inc., All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, 1900 E. Lake Ave., Glenview, IL 60025. For information regarding permissions, call (847) 486-2635.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data
Principles of biochemistry / H. Robert Horton ... [et al]. — 5th ed.
p. cm.
ISBN 0-321-70733-8
1. Biochemistry. I. Horton, H. Robert, 1935-
QP514.2.P745 2012
612'.015—dc23
2011019987

ISBN 10: 0-321-70733-8
ISBN 13: 978-0-321-70733-8
1 2 3 4 5 6 7 8 9 10—DOW—16 15 14 13 12

PEARSON

www.pearsonhighered.com

Downloaded From Vet-eBooks.com

Science should be as simple as possible,
but not simpler.

– Albert Einstein

Downloaded From Vet-eBooks.com

This page intentionally left blank

Downloaded From Vet-eBooks.com

Brief Contents

Part One

Introduction

1 Introduction to Biochemistry **1**

2 Water **28**

Part Two

Structure and Function

3 Amino Acids and the Primary Structures of Proteins **55**

4 Proteins: Three-Dimensional Structure and Function **85**

5 Properties of Enzymes **134**

6 Mechanisms of Enzymes **162**

7 Coenzymes and Vitamins **196**

8 Carbohydrates **227**

9 Lipids and Membranes **256**

Part Three

Metabolism and Bioenergetics

10 Introduction to Metabolism **294**

11 Glycolysis **325**

12 Gluconeogenesis, the Pentose Phosphate Pathway, and Glycogen Metabolism **355**

13 The Citric Acid Cycle **385**

14 Electron Transport and ATP Synthesis **417**

15 Photosynthesis **443**

16 Lipid Metabolism **475**

17 Amino Acid Metabolism **514**

18 Nucleotide Metabolism **550**

Part Four

Biological Information Flow

19 Nucleic Acids **573**

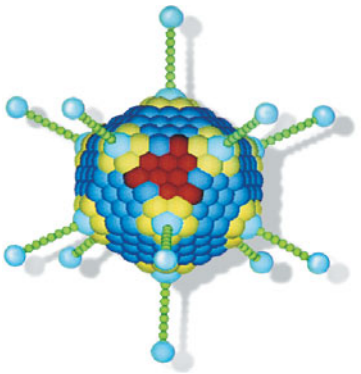
20 DNA Replication, Repair, and Recombination **601**

21 Transcription and RNA Processing **634**

22 Protein Synthesis **666**

Downloaded From Vet-eBooks.com

Contents



To the Student **xxiii**
Preface **xxv**
About the Authors **xxxiii**

Part One
Introduction

1 Introduction to Biochemistry **1**

1.1 Biochemistry Is a Modern Science **2**
1.2 The Chemical Elements of Life **3**
1.3 Many Important Macromolecules Are Polymers **4**
A. Proteins **6**
B. Polysaccharides **6**
C. Nucleic Acids **7**
D. Lipids and Membranes **9**
1.4 The Energetics of Life **10**
A. Reaction Rates and Equilibria **11**
B. Thermodynamics **12**
C. Equilibrium Constants and Standard Gibbs Free Energy Changes **13**
D. Gibbs Free Energy and Reaction Rates **14**
1.5 Biochemistry and Evolution **15**
1.6 The Cell Is the Basic Unit of Life **17**
1.7 Prokaryotic Cells: Structural Features **17**
1.8 Eukaryotic Cells: Structural Features **18**
A. The Nucleus **20**
B. The Endoplasmic Reticulum and Golgi Apparatus **20**
C. Mitochondria and Chloroplasts **21**
D. Specialized Vesicles **22**
E. The Cytoskeleton **23**
1.9 A Picture of the Living Cell **23**
1.10 Biochemistry Is Multidisciplinary **26**
Appendix: The Special Terminology of Biochemistry **26**
Selected Readings **27**



2 Water **28**

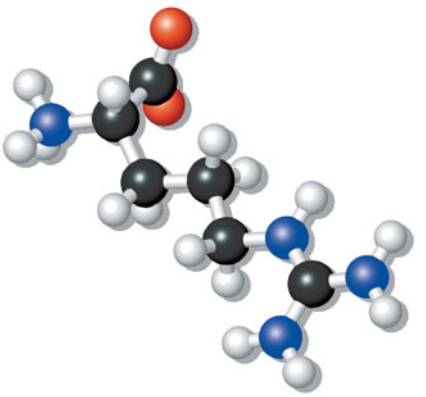
2.1 The Water Molecule Is Polar **29**
2.2 Hydrogen Bonding in Water **30**
Box 2.1 Extreme Thermophiles **32**
2.3 Water Is an Excellent Solvent **32**
A. Ionic and Polar Substances Dissolve in Water **32**
Box 2.2 Blood Plasma and Seawater **33**
B. Cellular Concentrations and Diffusion **34**
C. Osmotic Pressure **34**
2.4 Nonpolar Substances Are Insoluble in Water **35**

Downloaded From Vet-eBooks.com

2.5 Noncovalent Interactions 37
A. Charge–Charge Interactions 37
B. Hydrogen Bonds 37
C. Van der Waals Forces 38
D. Hydrophobic Interactions 39
2.6 Water Is Nucleophilic 39
Box 2.3 The Concentration of Water 41
2.7 Ionization of Water 41
2.8 The pH Scale 43
Box 2.4 The Little "p" in pH 44
2.9 Acid Dissociation Constants of Weak Acids 44
Sample Calculation 2.1 Calculating the pH of Weak Acid Solutions 49
2.10 Buffered Solutions Resist Changes in pH 50
Sample Calculation 2.2 Buffer Preparation 50
Summary 52
Problems 52
Selected Readings 54

PART TWO
Structure and Function

3 Amino Acids and the Primary Structures of Proteins 55



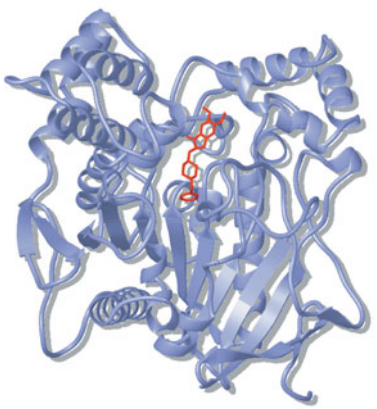
3.1 General Structure of Amino Acids 56
3.2 Structures of the 20 Common Amino Acids 58
Box 3.1 Fossil Dating by Amino Acid Racemization 58
A. Aliphatic R Groups 59
B. Aromatic R Groups 59
C. R Groups Containing Sulfur 60
D. Side Chains with Alcohol Groups 60
Box 3.2 An Alternative Nomenclature 61
E. Positively Charged R Groups 61
F. Negatively Charged R Groups and Their Amide Derivatives 62
G. The Hydrophobicity of Amino Acid Side Chains 62
3.3 Other Amino Acids and Amino Acid Derivatives 62
3.4 Ionization of Amino Acids 63
Box 3.3 Common Names of Amino Acids 64
3.5 Peptide Bonds Link Amino Acids in Proteins 67
3.6 Protein Purification Techniques 68
3.7 Analytical Techniques 70
3.8 Amino Acid Composition of Proteins 73
3.9 Determining the Sequence of Amino Acid Residues 74
3.10 Protein Sequencing Strategies 76
3.11 Comparisons of the Primary Structures of Proteins Reveal Evolutionary Relationships 79
Summary 82
Problems 82
Selected Readings 84



4 Proteins: Three-Dimensional Structure and Function 85
4.1 There Are Four Levels of Protein Structure 87
4.2 Methods for Determining Protein Structure 88

Downloaded From Vet-eBooks.com

4.3 The Conformation of the Peptide Group 91
Box 4.1 Flowering Is Controlled by *Cis/Trans* Switches 93
4.4 The α Helix 94
4.5 β Strands and β Sheets 97
4.6 Loops and Turns 98
4.7 Tertiary Structure of Proteins 99
A. Supersecondary Structures 100
B. Domains 101
C. Domain Structure, Function, and Evolution 102
D. Intrinsically Disordered Proteins 102
4.8 Quaternary Structure 103
4.9 Protein–Protein Interactions 109
4.10 Protein Denaturation and Renaturation 110
4.11 Protein Folding and Stability 114
A. The Hydrophobic Effect 114
B. Hydrogen Bonding 115
Box 4.2 CASP: The Protein Folding Game 116
C. Van der Waals Interactions and Charge–Charge Interactions 117
D. Protein Folding Is Assisted by Molecular Chaperones 117
4.12 Collagen, a Fibrous Protein 119
Box 4.3 Stronger Than Steel 121
4.13 Structure of Myoglobin and Hemoglobin 122
4.14 Oxygen Binding to Myoglobin and Hemoglobin 123
A. Oxygen Binds Reversibly to Heme 123
B. Oxygen-Binding Curves of Myoglobin and Hemoglobin 124
Box 4.4 Embryonic and Fetal Hemoglobins 126
C. Hemoglobin Is an Allosteric Protein 127
4.15 Antibodies Bind Specific Antigens 129
Summary 130
Problems 131
Selected Readings 133

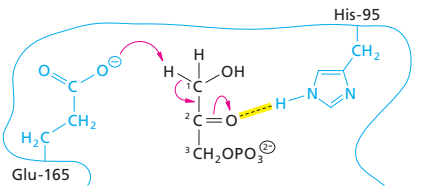


5 Properties of Enzymes 134
5.1 The Six Classes of Enzymes 136
Box 5.1 Enzyme Classification Numbers 137
5.2 Kinetic Experiments Reveal Enzyme Properties 138
A. Chemical Kinetics 138
B. Enzyme Kinetics 139
5.3 The Michaelis-Menten Equation 140
A. Derivation of the Michaelis-Menten Equation 141
B. The Calalytic Constant K_{cat} 143
C. The Meanings of K_m 144
5.4 Kinetic Constants Indicate Enzyme Activity and Catalytic Proficiency 144
5.5 Measurement of K_m and V_{max} 145
Box 5.2 Hyperbolas Versus Straight Lines 146
5.6 Kinetics of Multisubstrate Reactions 147
5.7 Reversible Enzyme Inhibition 148
A. Competitive Inhibition 149
B. Uncompetitive Inhibition 150

Downloaded From Vet-e-Books.com

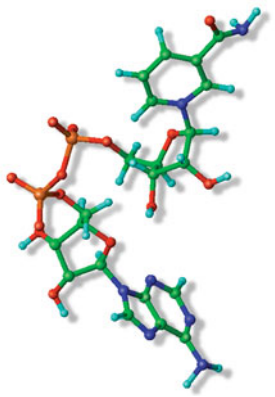
C. Noncompetitive Inhibition 150
D. Uses of Enzyme Inhibition 151
5.8 Irreversible Enzyme Inhibition 152
5.9 Regulation of Enzyme Activity 153
A. Phosphofructokinase Is an Allosteric Enzyme 154
B. General Properties of Allosteric Enzymes 155
C. Two Theories of Allosteric Regulation 156
D. Regulation by Covalent Modification 158
5.10 Multienzyme Complexes and Multifunctional Enzymes 158
Summary 159
Problems 159
Selected Readings 161

6 Mechanisms of Enzymes 162



6.1 The Terminology of Mechanistic Chemistry 162
A. Nucleophilic Substitutions 163
B. Cleavage Reactions 163
C. Oxidation–Reduction Reactions 164
6.2 Catalysts Stabilize Transition States 164
6.3 Chemical Modes of Enzymatic Catalysis 166
A. Polar Amino Acids Residues in Active Sites 166
Box 6.1 Site-Directed Mutagenesis Modifies Enzymes 167
B. Acid–Base Catalysis 168
C. Covalent Catalysis 169
D. pH Affects Enzymatic Rates 170
6.4 Diffusion-Controlled Reactions 171
A. Triose Phosphate Isomerase 172
Box 6.2 The "Perfect Enzyme"? 174
B. Superoxide Dismutase 175
6.5 Modes of Enzymatic Catalysis 175
A. The Proximity Effect 176
B. Weak Binding of Substrates to Enzymes 178
C. Induced Fit 179
D. Transition State Stabilization 180
6.6 Serine Proteases 183
A. Zymogens Are Inactive Enzyme Precursors 183
Box 6.3 Kornberg's Ten Commandments 183
B. Substrate Specificity of Serine Proteases 184
C. Serine Proteases Use Both the Chemical and the Binding Modes of Catalysis 185
Box 6.4 Clean Clothes 186
Box 6.5 Convergent Evolution 187
6.7 Lysozyme 187
6.8 Arginine Kinase 190
Summary 192
Problems 193
Selected Readings 194

Downloaded From Vet-eBooks.com



7 Coenzymes and Vitamins 196

7.1 Many Enzymes Require Inorganic Cations 197
7.2 Coenzyme Classification 197
7.3 ATP and Other Nucleotide Cosubstrates 198
Box 7.1 Missing Vitamins 200
7.4 NAD⊕ and NADP⊕ 200
Box 7.2 NAD Binding to Dehydrogenases 203
7.5 FAD and FMN 204
7.6 Coenzyme A and Acyl Carrier Protein 204
7.7 Thiamine Diphosphate 206
7.8 Pyridoxal Phosphate 207
7.9 Vitamin C 209
7.10 Biotin 211
Box 7.3 One Gene: One Enzyme 212
7.11 Tetrahydrofolate 213
7.12 Cobalamin 215
7.13 Lipoamide 216
7.14 Lipid Vitamins 217
A. Vitamin A 217
B. Vitamin D 218
C. Vitamin E 218
D. Vitamin K 218
7.15 Ubiquinone 219
Box 7.4 Rat Poison 220
7.16 Protein Coenzymes 221
7.17 Cytochromes 221
Box 7.5 Noble Prizes for Vitamins and Coenzymes 223
Summary 223
Problems 224
Selected Readings 226



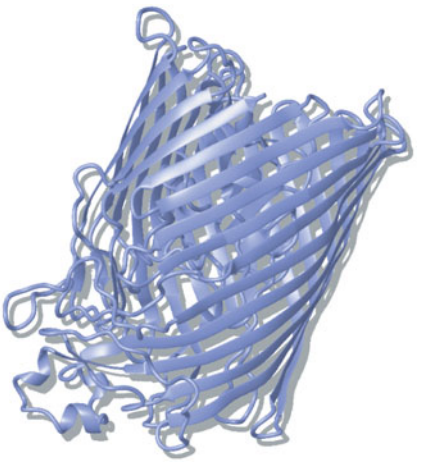
8 Carbohydrates 227

8.1 Most Monosaccharides Are Chiral Compounds 228
8.2 Cyclization of Aldoses and Ketoses 230
8.3 Conformations of Monosaccharides 234
8.4 Derivatives of Monosaccharides 235
A. Sugar Phosphates 235
B. Deoxy Sugars 235
C. Amino Sugars 235
D. Sugar Alcohols 236
E. Sugar Acids 236
8.5 Disaccharides and Other Glycosides 236
A. Structures of Disaccharides 237
B. Reducing and Nonreducing Sugars 238
C. Nucleosides and Other Glycosides 239
Box 8.1 The Problem with Cats 240
8.6 Polysaccharides 240
A. Starch and Glycogen 240
B. Cellulose 243

Downloaded From Vet-eBooks.com

C. Chitin 244
8.7 Glycoconjugates 244
A. Proteoglycans 244
Box 8.2 Nodulation Factors Are Lipo-Oligosaccharides 246
B. Peptidoglycans 246
C. Glycoproteins 248
Box 8.3 ABO Blood Group 250
Summary 252
Problems 253
Selected Readings 254

9 Lipids and Membranes 256
9.1 Structural and Functional Diversity of Lipids 256
9.2 Fatty Acids 256
Box 9.1 Common Names of Fatty Acids 258
Box 9.2 *Trans* Fatty Acids and Margarine 259
9.3 Triacylglycerols 261
9.4 Glycerophospholipids 262
9.5 Sphingolipids 263
9.6 Steroids 266
9.7 Other Biologically Important Lipids 268
9.8 Biological Membranes 269
A. Lipid Bilayers 269
Box 9.3 Gregor Mendel and Gibberellins 270
B. Three Classes of Membrane Proteins 270
Box 9.4 New Lipid Vesicles, or Liposomes 272
Box 9.5 Some Species Have Unusual Lipids in Their Membranes 274
C. The Fluid Mosaic Model of Biological Membranes 274
9.9 Membranes Are Dynamic Structures 275
9.10 Membrane Transport 277
A. Thermodynamics of Membrane Transport 278
B. Pores and Channels 279
C. Passive Transport and Facilitated Diffusion 280
D. Active Transport 282
E. Endocytosis and Exocytosis 283
9.11 Transduction of Extracellular Signals 283
A. Receptors 283
Box 9.6 The Hot Spice of Chili Peppers 284
B. Signal Transducers 285
C. The Adenylyl Cyclase Signaling Pathway 287
D. The Inositol–Phospholipid Signaling Pathway 287
Box 9.7 Bacterial Toxins and G Proteins 290
E. Receptor Tyrosine Kinases 290
Summary 291
Problems 292
Selected Readings 293



Downloaded From Vet-eBooks.com



PART THREE

Metabolism and Bioenergetics

10 Introduction to Metabolism 294

10.1 Metabolism Is a Network of Reactions 294
10.2 Metabolic Pathways 297
A. Pathways Are Sequences of Reactions 297
B. Metabolism Proceeds by Discrete Steps 297
C. Metabolic Pathways Are Regulated 297
D. Evolution of Metabolic Pathways 301
10.3 Major Pathways in Cells 302
10.4 Compartmentation and Interorgan Metabolism 304
10.5 Actual Gibbs Free Energy Change, Not Standard Free Energy Change, Determines the Direction of Metabolic Reactions 306
Sample Calculation 10.1 Calculating Standard Gibbs Free Energy Change from Energies of Formation 308
10.6 The Free Energy of ATP Hydrolysis 308
10.7 The Metabolic Roles of ATP 311
A. Phosphoryl Group Transfer 311
Sample Calculation 10.2 Gibbs Free Energy Change 312
Box 10.1 The Squiggle 312
B. Production of ATP by Phosphoryl Group Transfer 314
C. Nucleotidyl Group Transfer 315
10.8 Thioesters Have High Free Energies of Hydrolysis 316
10.9 Reduced Coenzymes Conserve Energy from Biological Oxidations 316
A. Gibbs Free Energy Change Is Related to Reduction Potential 317
B. Electron Transfer from NADH Provides Free Energy 319
Box 10.2 $NAD^{\oplus}$ and NADH Differ in Their Ultraviolet Absorption Spectra 321
10.10 Experimental Methods for Studying Metabolism 321
Summary 322
Problems 323
Selected Readings 324



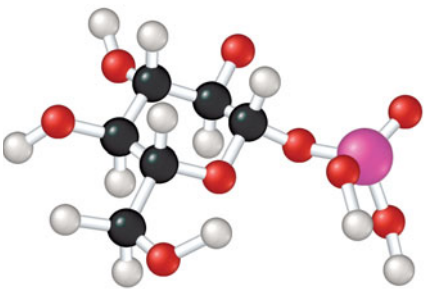
11 Glycolysis 325

11.1 The Enzymatic Reactions of Glycolysis 326
11.2 The Ten Steps of Glycolysis 326
1. Hexokinase 326
2. Glucose 6-Phosphate Isomerase 327
3. Phosphofructokinase-1 330
4. Aldolase 330
Box 11.1 A Brief History of the Glycolysis Pathway 331
5. Triose Phosphate Isomerase 332
6. Glyceraldehyde 3-Phosphate Dehydrogenase 333
7. Phosphoglycerate Kinase 335
Box 11.2 Formation of 2,3-*Bis*phosphoglycerate in Red Blood Cells 335
Box 11.3 Arsenate Poisoning 336
8. Phosphoglycerate Mutase 336
9. Enolase 338
10. Pryuvate Kinase 338

Downloaded From Vet-eBooks.com

11.3 The Fate of Pryuvate 338
A. Metabolism of Pryuvate to Ethanol 339
B. Reduction of Pyruvate to Lactate 340
Box 11.4 The Lactate of the Long-Distance Runner 341
11.4 Free Energy Changes in Glycolysis 341
11.5 Regulation of Glycolysis 343
A. Regulation of Hexose Transporters 344
B. Regulation of Hexokinase 344
Box 11.5 Glucose 6-Phosphate Has a Pivotal Metabolic Role in the Liver 345
C. Regulation of Phosphofructokinase-1 345
D. Regulation of Pyruvate Kinase 346
E. The Pasteur Effect 347
11.6 Other Sugars Can Enter Glycolysis 347
A. Sucrose Is Cleaved to Monosaccharides 348
B. Fructose Is Converted to Glyceraldehyde 3-Phosphate 348
C. Galactose Is Converted to Glucose 1-Phosphate 349
Box 11.6 A Secret Ingredient 349
D. Mannose Is Converted to Fructose 6-Phosphate 351
11.7 The Entner–Doudoroff Pathway in Bacteria 351
Summary 352
Problems 353
Selected Readings 354

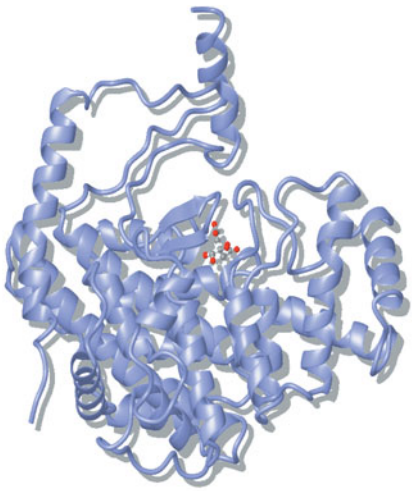
12 Gluconeogenesis, the Pentose Phosphate Pathway, and Glycogen Metabolism 355



12.1 Gluconeogenesis 356
A. Pyruvate Carboxylase 357
B. Phosphoenolpyruvate Carboxykinase 358
C. Fructose 1,6-*bis*phosphatase 358
Box 12.1 Supermouse 359
D. Glucose 6-Phosphatase 359
12.2 Precursors for Gluconeogenesis 360
A. Lactate 360
B. Amino Acids 360
C. Glycerol 361
D. Propionate and Lactate 361
E. Acetate 362
Box 12.2 Glucose Is Sometimes Converted to Sorbitol 362
12.3 Regulation of Gluconeogenesis 363
Box 12.3 The Evolution of a Complex Enzyme 364
12.4 The Pentose Phosphate Pathway 364
A. Oxidative Stage 366
B. Nonoxidative Stage 364
Box 12.4 Glucose 6-Phosphate Dehydrogenase Deficiency in Humans 367
C. Interconversions Catalyzed by Transketolase and Transaldolase 368
12.5 Glycogen Metabolism 368
A. Glycogen Synthesis 369
B. Glycogen Degradation 370
12.6 Regulation of Glycogen Metabolism in Mammals 372

Downloaded From Vet-eBooks.com

A. Regulation of Glycogen Phosphorylase 372
Box 12.5 Head Growth and Tail Growth 373
B. Hormones Regulate Glycogen Metabolism 375
C. Hormones Regulate Gluconeogenesis and Glycolysis 376
12.7 Maintenance of Glucose Levels in Mammals 378
12.8 Glycogen Storage Diseases 381
Summary 382
Problems 382
Selected Readings 383



13 The Citric Acid Cycle 385

Box 13.1 An Egregious Error 386
13.1 Conversion of Pyruvate to Acetyl CoA 387
Sample Calculation 13.1 390
13.2 The Citric Acid Cycle Oxidizes Acetyl CoA 391
Box 13.2 Where Do the Electrons Come From? 392
13.3 The Citric Acid Cycle Enzymes 394
1. Citrate Synthase 394
Box 13.3 Citric Acid 396
2. Aconitase 396
Box 13.4 Three-Point Attachment of Prochiral Substrates to Enzymes 397
3. Isocitrate Dehydrogenase 397
4. The α-Ketoglutarate Dehydrogenase Complex 398
5. Succinyl CoA Synthetase 398
6. Succinate Dehydrogenase Complex 399
Box 13.5 What's in a Name? 399
Box 13.6 On the Accuracy of the World Wide Web 401
7. Fumarase 401
8. Malate Deydrogenase 401
Box 13.7 Converting One Enzyme into Another 402
13.4 Entry of Pyruvate Into Mitochondria 402
13.5 Reduced Coenzymes Can Fuel the Production of ATP 405
13.6 Regulation of the Citric Acid Cycle 406
13.7 The Citric Acid Cycle Isn't Always a "Cycle" 407
Box 13.8 A Cheap Cancer Drug? 408
13.8 The Glyoxylate Pathway 409
13.9 Evolution of the Citric Acid Cycle 412
Summary 414
Problems 414
Selected Readings 416



14 Electron Transport and ATP Synthesis 417

14.1 Overview of Membrane-associated Electron Transport and ATP Synthesis 418
14.2 The Mitochondrion 418
Box 14.1 An Exception to Every Rule 420
14.3 The Chemiosmotic Theory and the Protonmotive Force 420
A. Historical Background: The Chemiosmotic Theory 420
B. The Protonmotive Force 421

Downloaded From Vet-eBooks.com

14.4 Electron Transport 423
A. Complexes I Through IV 423
B. Cofactors in Electron Transport 425
14.5 Complex I 426
14.6 Complex II 427
14.7 Complex III 428
14.8 Complex IV 431
14.9 Complex V: ATP Synthase 433
Box 14.2 Proton Leaks and Heat Production 435
14.10 Active Transport of ATP, ADP, and P_i Across the Mitochondrial Membrane 435
14.11 The P/O Ratio 436
14.12 NADH Shuttle Mechanisms in Eukaryotes 436
Box 14.3 The High Cost of Living 439
14.13 Other Terminal Electron Acceptors and Donors 439
14.14 Superoxide Anions 440
Summary 441
Problems 441
Selected Readings 442

15 Photosynthesis 443

15.1 Light-Gathering Pigments 444
A. The Structures of Chlorophylls 444
B. Light Energy 445
C. The Special Pair and Antenna Chlorophylls 446
Box 15.1 Mendel's Seed Color Mutant 447
D. Accessory Pigments 447
15.2 Bacterial Photosystems 448
A. Photosystem II 448
B. Photosystem I 450
C. Coupled Photosystems and Cytochrome *bf* 453
D. Reduction Potentials and Gibbs Free Energy in Photosynthesis 455
E. Photosynthesis Takes Place Within Internal Membranes 457
Box 15.2 Oxygen "Pollution" of Earth's Atmosphere 457
15.3 Plant Photosynthesis 458
A. Chloroplasts 458
B. Plant Photosystems 459
C. Organization of Cloroplast Photosystems 459
Box 15.3 Bacteriorhodopsin 461
15.4 Fixation of CO_2: The Calvin Cycle 461
A. The Calvin Cycle 462
B. Rubisco: Ribulose 1,5-*bis*phosphate Carboxylase-oxygenase 462
C. Oxygenation of Ribulose 1,5-*bis*phosphate 465
Box 15.4 Building a Better Rubisco 466
D. Calvin Cycle: Reduction and Regeneration Stages 466
15.5 Sucrose and Starch Metabolism in Plants 467
Box 15.5 Gregor Mendel's Wrinkled Peas 469
15.6 Additional Carbon Fixation Pathways 469
A. Compartmentalization in Bacteria 469



Downloaded From Vet-eBooks.com

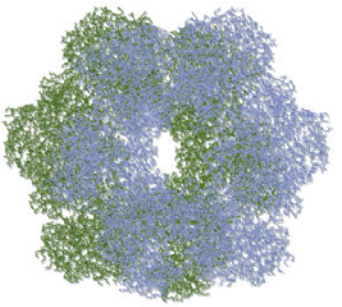
B. The C_4 Pathway 469
C. Crassulacean Acid Metabolism (CAM) 471
Summary 472
Problems 473
Selected Readings 474



16 Lipid Metabolism 475

16.1 Fatty Acid Synthesis 475
A. Synthesis of Malonyl ACP and Acetyl ACP 476
B. The Initiation Reaction of Fatty Acid Synthesis 477
C. The Elongation Reactions of Fatty Acid Synthesis 477
D. Activation of Fatty Acids 479
E. Fatty Acid Extension and Desaturation 479
16.2 Synthesis of Triacylglycerols and Glycerophospholipids 481
16.3 Synthesis of Eicosanoids 483
Box 16.1 *sn*-Glycerol 3-Phosphate 484
Box 16.2 The Search for a Replacement for Asprin 486
16.4 Synthesis of Ether Lipids 487
16.5 Synthesis of Sphingolipids 488
16.6 Synthesis of Cholesterol 488
A. Stage 1: Acetyl CoA to Isopentenyl Diphosphate 488
B. Stage 2: Isopentenyl Diphosphate to Squalene 488
C. Stage 3: Squalene to Cholesterol 490
D. Other Products of Isoprenoid Metabolism 490
Box 16.3 Lysosomal Storage Diseases 492
Box 16.4 Regulating Cholesterol Levels 493
16.7 Fatty Acid Oxidation 494
A. Activation of Fatty Acids 494
B. The Reactions of β-Oxidation 494
C. Fatty Acid Synthesis and β-Oxidation 497
D. Transport of Fatty Acyl CoA into Mitochondria 497
Box 16.5 A Trifunctional Enzyme for β-Oxidation 498
E. ATP Generation from Fatty Acid Oxidation 498
F. β-Oxidation of Odd-Chain and Unsaturated Fatty Acids 499
16.8 Eukaryotic Lipids Are Made at a Variety of Sites 501
16.9 Lipid Metabolism Is Regulated by Hormones in Mammals 502
16.10 Absorption and Mobilization of Fuel Lipids in Mammals 505
A. Absorption of Dietary Lipids 505
B. Lipoproteins 505
Box 16.6 Extra Virgin Olive Oil 506
Box 16.7 Lipoprotein Lipase and Coronary Heart Disease 507
C. Serum Albumin 508
16.11 Ketone Bodies Are Fuel Molecules 508
A. Ketone Bodies Are Synthesized in the Liver 509
B. Ketone Bodies Are Oxidized in Mitochondria 510
Box 16.8 Lipid Metabolism in Diabetes 511
Summary 511
Problems 511
Selected Readings 513

Downloaded From Vet-e-Books.com



17 Amino Acid Metabolism 514

17.1 The Nitrogen Cycle and Nitrogen Fixation 515

17.2 Assimilation of Ammonia 518

A. Ammonia Is Incorporated into Glutamate and Glutamine 518

B. Transamination Reactions 518

17.3 Synthesis of Amino Acids 520

A. Aspartate and Asparagine 520

B. Lysine, Methionine, Threonine 520

C. Alanine, Valine, Leucine, and Isoleucine 521

Box 17.1 Childhood Acute Lymphoblastic Leukemia Can Be Treated with Asparaginase 522

D. Glutamate, Glutamine, Arginine, and Proline 523

E. Serine, Glycine, and Cysteine 523

F. Phenylalanine, Tyrosine, and Tryptophan 523

G. Histidine 527

Box 17.2 Genetically Modified Food 528

Box 17.3 Essential and Nonessential Amino Acids in Animals 529

17.4 Amino Acids as Metabolic Precursors 529

A. Products Derived from Glutamate, Glutamine, and Aspartate 529

B. Products Derived from Serine and Glycine 529

C. Synthesis of Nitric Oxide from Arginine 530

D. Synthesis of Lignin from Phenylalanine 531

E. Melanin Is Made from Tyrosine 531

17.5 Protein Turnover 531

Box 17.4 Apoptosis–Programmed Cell Death 534

17.6 Amino Acid Catabolism 534

A. Alanine, Asparagine, Aspartate, Glutamate, and Glutamine 535

B. Arginine, Histidine, and Proline 535

C. Glycine and Serine 536

D. Threonine 537

E. The Branched Chain Amino Acids 537

F. Methionine 539

Box 17.5 Phenylketonuria, a Defect in Tyrosine Formation 540

G. Cysteine 540

H. Phenylalanine, Tryptophane, and Tyrosine 541

I. Lysine 542

17.7 The Urea Cycle Converts Ammonia into Urea 542

A. Synthesis of Carbamoyl Phosphate 543

B. The Reactions of the Urea Cycle 543

Box 17.6 Diseases of Amino Acid Metabolism 544

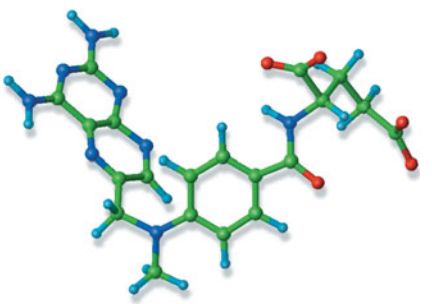
C. Ancillary Reactions of the Urea Cycle 547

17.8 Renal Glutamine Metabolism Produces Bicarbonate 547

Summary 548

Problems 548

Selected Readings 549



18 Nucleotide Metabolism 550

18.1 Synthesis of Purine Nucleotides 550

Box 18.1 Common Names of the Bases 552

18.2 Other Purine Nucleotides Are Synthesized from IMP 554

18.3 Synthesis of Pyrimidine Nucleotides 555

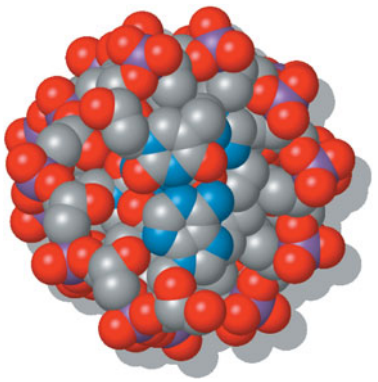
Downloaded From Vet-eBooks.com

A. The Pathway for Pyrimidine Synthesis 556
Box 18.2 How Some Enzymes Transfer Ammonia from Glutamate 558
B. Regulation of Pyrimidine Synthesis 559
18.4 CTP Is Synthesized from UMP 559
18.5 Reduction of Ribonucleotides to Deoxyribonucleotides 560
18.6 Methylation of dUMP Produces dTMP 560
Box 18.3 Free Radicals in the Reduction of Ribonucleotides 562
Box 18.4 Cancer Drugs Inhibit dTTP Synthesis 564
18.7 Modified Nucleotides 564
18.8 Salvage of Purines and Pyrimidines 564
18.9 Purine Catabolism 565
18.10 Pyrimidine Catabolism 568
Box 18.5 Lesch–Nyhan Syndrome and Gout 569
Summary 571
Problems 571
Selected Readings 572

PART FOUR
Biological Information Flow

19 Nucleic Acids 573

19.1 Nucleotides Are the Building Blocks of Nucleic Acids 574
A. Ribose and Deoxyribose 574
B. Purines and Pyrimidines 574
C. Nucleosides 575
D. Nucleotides 577
19.2 DNA Is Double-Stranded 579
A. Nucleotides Are Joined by 3′–5′ Phosphodiester Linkages 580
B. Two Antiparallel Strands Form a Double Helix 581
C. Weak Forces Stabilize the Double Helix 583
D. Conformations of Double-Stranded DNA 585
19.3 DNA Can Be Supercoiled 586
19.4 Cells Contain Several Kinds of RNA 587
Box 19.1 Pulling DNA 588
19.5 Nucleosomes and Chromatin 588
A. Nucleosomes 588
B. Higher Levels of Chromatin Structure 590
C. Bacterial DNA Packaging 590
19.6 Nucleases and Hydrolysis of Nucleic Acids 591
A. Alkaline Hydrolysis of RNA 591
B. Hydrolysis of RNA by Ribonuclease A 592
C. Restriction Endonucleases 593
D. *Eco*RI Binds Tightly to DNA 595
19.7 Uses of Restriction Endocucleases 596
A. Restriction Maps 596
B. DNA Fingerprints 596
C. Recombinant DNA 597
Summary 598
Problems 599
Selected Readings 599



Downloaded From Vet-eBooks.com

20 DNA Replication, Repair, and Recombination 601

20.1 Chromosomal DNA Replication Is Bidirectional 602

20.2 DNA Polymerase 603

A. Chain Elongation Is a Nucleotidyl-Group–Transfer Reaction 604

B. DNA Polymerase III Remains Bound to the Replication Fork 606

C. Proofreading Corrects Polymerization Errors 607

20.3 DNA Polymerase Synthesizes Two Strands Simultaneously 607

A. Lagging Strand Synthesis Is Discontinuous 608

B. Each Okazaki Fragment Begins with an RNA Primer 608

C. Okazaki Fragments Are Joined by the Action of DNA Polymerase I and DNA Ligase 609

20.4 Model of the Replisome 610

20.5 Initiation and Termination of DNA Replication 615

20.6 DNA Replication in Eukaryotes 615

A. The Polymerase Chain Reaction Uses DNA Polymerase to Amplify Selected DNA Sequences 615

B. Sequencing DNA Using Dideoxynucleotides 616

C. Massively Parallel DNA Sequencing by Synthesis 618

20.7 DNA Replication in Eukaryotes 619

20.8 Repair of Damaged DNA 622

A. Repair after Photodimerization: An Example of Direct Repair 622

B. Excision Repair 624

BOX 20.1 The Problem with Methylcytosine 626

20.9 Homologous Recombination 626

A. The Holliday Model of General Recombination 626

B. Recombination in *E. coli* 627

BOX 20.2 Molecular Links Between DNA Repair and Breast Cancer 630

C. Recombination Can Be a Form of Repair 631

Summary 631

Problems 632

Selected Readings 632



21 Transcription and RNA Processing 633

21.1 Types of RNA 634

21.2 RNA Polymerase 635

A. RNA Polymerase Is an Oligomeric Protein 635

B. The Chain Elongation Reaction 636

21.3 Transcription Initiation 638

A. Genes Have a 5′ → 3′ Orientation 638

B. The Transcription Complex Assembles at a Promoter 639

C. The σ sigma Subunit Recognizes the Promoter 640

D. RNA Polymerase Changes Conformation 641

21.4 Transcription Termination 643

21.5 Transcription in Eukaryotes 645

A. Eukaryotic RNA Polymerases 645

B. Eukaryotic Transcription Factors 647

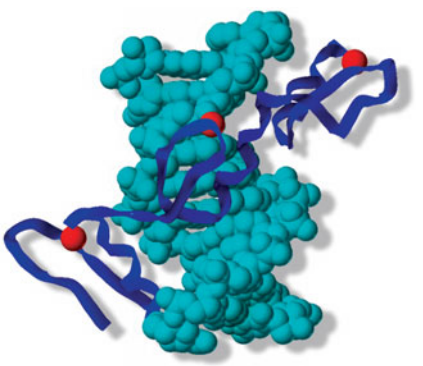
C. The Role of Chromatin in Eukaryotic Transcription 648

21.6 Transcription of Genes Is Regulated 648

21.7 The *lac* Operon, an Example of Negative and Positive Regulation 650

A. *lac* Repressor Blocks Transcription 650

B. The Structure of *lac* Repressor 651



Downloaded From Vet-e-Books.com



C. cAMP Regulatory Protein Activates Transcription **652**
21.8 Post-transcriptional Modification of RNA **654**
A. Transfer RNA Processing **654**
B. Ribosomal RNA Processing **655**
21.9 Eukaryotic mRNA Processing **655**
A. Eukaryotic mRNA Molecules Have Modified Ends **657**
B. Some Eukaryotic mRNA Precursors Are Spliced **657**
Summary **663**
Problems **663**
Selected Readings **664**

22 Protein Synthesis **665**

22.1 The Genetic Code **665**
22.2 Transfer RNA **668**
A. The Three-Dimensional Structure of tRNA **668**
B. tRNA Anticodons Base-Pair with mRNA Codons **669**
22.3 Aminoacyl-tRNA Synthetases **670**
A. The Aminoacyl-tRNA Synthetase Reaction **671**
B. Specificity of Aminoacyl-tRNA Synthetases **671**
C. Proofreading Activity of Aminoacyl-tRNA Synthetases **673**
22.4 Ribosomes **673**
A. Ribosomes Are Composed of Both Ribosomal RNA and Protein **674**
B. Ribosomes Contain Two Aminoacyl-tRNA Binding Sites **675**
22.5 Initiation of Translation **675**
A. Initiator tRNA **675**
B. Initiation Complexes Assemble Only at Initiation Codons **676**
C. Initiation Factors Help Form the Initiation Complex **677**
D. Translation Initiation in Eukaryotes **679**
22.6 Chain Elongation During Protein Synthesis Is a Three-Step Microcycle **679**
A. Elongation Factors Dock an Aminoacyl-tRNA in the A Site **680**
B. Peptidyl Transferase Catalyzes Peptide Bond Formation **681**
C. Translocation Moves the Ribosome by One Codon **682**
22.7 Termination of Translation **684**
22.8 Protein Synthesis Is Energetically Expensive **684**
22.9 Regulation of Protein Synthesis **685**
A. Ribosomal Protein Synthesis Is Coupled to Ribosome Assembly in *E. coli* **685**
Box 22.1 Some Antibiotics Inhibit Protein Synthesis **686**
B. Globin Synthesis Depends on Heme Availability **687**
C. The *E. coli trp* Operon Is Regulated by Repression and Attenuation **687**
22.10 Post-translational Processing **689**
A. The Signal Hypothesis **691**
B. Glycosylation of Proteins **694**
Summary **694**
Problems **695**
Selected Readings **696**
Solutions **697**
Glossary **751**
Illustration Credits **767**
Index **769**

Downloaded From Vet-eBooks.com

To the Student

Welcome to biochemistry—the study of life at the molecular level. As you venture into this exciting and dynamic discipline, you'll discover many new and wonderful things. You'll learn how some enzymes can catalyze chemical reactions at speeds close to theoretical limits—reactions that would otherwise occur only at imperceptibly low rates. You'll learn about the forces that maintain biomolecular structure and how even some of the weakest of those forces make life possible. You'll also learn how biochemistry has thousands of applications in day-to-day life—in medicine, drug design, nutrition, forensic science, agriculture, and manufacturing. In short, you'll begin a journey of discovery about how biochemistry makes life both possible and better.

Before we begin, we would like to offer a few words of advice:

Don't just memorize facts; instead, understand principles

In this book, we have tried to identify the most important principles of biochemistry. Because the knowledge base of biochemistry is continuously expanding, we must grasp the underlying themes of this science in order to understand it. This textbook is designed to expand on the foundation you have acquired in your chemistry and biology courses and to provide you with a biochemical framework that will allow you to understand new phenomena as you meet them.

Be prepared to learn a new vocabulary

An understanding of biochemical facts requires that you learn a biochemical vocabulary. This vocabulary includes the chemical structures of a number of key molecules. These molecules are grouped into families based on their structures and functions. You will also learn how to distinguish among members of each family and how small molecules combine to form macromolecules such as proteins and nucleic acids.

Test your understanding

True mastery of biochemistry lies with learning how to apply your knowledge and how to solve problems. Each chapter concludes with a set of carefully crafted problems that test your understanding of core principles. Many of these problems are mini case studies that present the problem within the context of a real biochemical puzzle.

For more practice, we are pleased to refer you to *The Study Guide for Principles of Biochemistry* by Scott Lefler and Allen Scism which presents a variety of supplementary questions that you may find helpful. You will also find additional problems on TheChemistryPlace® for *Principles of Biochemistry* (**http://www.chemplace.com**).

Learn to visualize in 3-D

Biochemicals are three-dimensional objects. Understanding what happens in a biochemical reaction at the molecular level requires that you be able to "see" what happens in three dimensions. We present the structures of simple molecules in several different ways in order to illustrate their three-dimensional conformation. In addition to the art in the book, you will find many animations and interactive molecular models on the website. We strongly suggest you look at these movies and do the exercises that accompany them as well as participate in the molecular visualization tutorials.

Feedback

Finally, please let us know of any errors or omissions you encounter as you use this text. Tell us what you would like to see in the next edition. With your help we will continue to evolve this work into an even more useful tool. Our e-mail addresses are at the end of the Preface. Good luck, and enjoy!

Downloaded From Vet-e-Books.com

This page intentionally left blank

Downloaded From Vet-eBooks.com

Preface

Given the breadth of coverage and diversity of ways to present topics in biochemistry, we have tried to make the text as modular as possible to allow for greater flexibility and organization. Each large topic resides in its own section. Reaction mechanisms are often separated from the main thread of the text and can be passed over by those who prefer not to cover this level of detail. The text is extensively cross-referenced to make it easier for you to reorganize the chapters and for students to see the interrelationships among various topics and to drill down to deeper levels of understanding.

We built the book explicitly for the beginning student taking a first course in biochemistry with the aim of encouraging students to think critically and to appreciate scientific knowledge for its own sake. Parts One and Two lay a solid foundation of chemical knowledge that will help students understand, rather than merely memorize, the dynamics of metabolic and genetic processes. These sections assume that students have taken prerequisite courses in general and organic chemistry and have acquired a rudimentary knowledge of the organic chemistry of carboxylic acids, amines, alcohols, and aldehydes. Even so, key functional groups and chemical properties of each type of biomolecule are carefully explained as their structures and functions are presented.

We also assume that students have previously taken a course in biology where they have learned about evolution, cell biology, genetics, and the diversity of life on this planet. We offer brief refreshers on these topics wherever possible.

New to this Edition

We are grateful for all the input we received on the first four editions of this text. You'll notice the following improvements in this fifth edition:

- **Key Concept** margin notes are provided throughout to highlight key concepts and principles that students must know.
- **Interest Boxes** have been updated and expanded, with 45% new to the fifth edition. We use interest boxes to explain some topics in more detail, to illustrate certain principles with specific examples, to stimulate students curiosity about science, to show applications of biochemistry, and to explain clinical relevance. We have also added a few interests boxes that warn students about misunderstanding and misapplications of biochemistry. Examples include Blood Plasma and Sea Water; Fossil Dating by Amino Acid Racemization; Embryonic and Fetal Hemoglobins; Clean Clothes; The Perfect Enzyme; Supermouse; The Evolution of a Complex Enzyme; An Egregious Error; Mendels Seed Color Mutant; Oxygen Pollution of Earth's Atmosphere; Extra Virgin Olive Oil; Missing Vitamins; Pulling DNA; and much more.
- New Material has been added throughout, including an improved explanation of early evolution (the Web of Life), more emphasis on protein protein interactions, a new section on intrinsically disordered proteins, and a better description of the distinction between Gibbs free energy changes and reaction rates. We have removed the final chapter on Recombinant DNA Technology and integrated much of that material into earlier chapters. We have added descriptions of a number of new protein structures and integrated them into two major themes: structure-function and multienzyme complexes. The best example is the fatty acid synthase complex in Chapter 16.

  In some cases new material was necessary because recent discoveries have changed our view of some reactions and processes. We now know, for example, that older versions of uric acid catabolism were incorrect, the correct pathway is shown in Figure 18.23.

Downloaded From Vet-e-Books.com

We have been careful not to add extra detail unless it supports and extends the basic concepts and principles that we have established over the past four editions. Similarly, we do not introduce new subjects unless they illustrate new concepts that were not covered in previous editions. The goal is to keep this textbook focused on the fundamentals that students need to know and prevent it from bloating up into an encyclopedia of mostly irrelevant information that detracts from the main pedagogical goals.

- **Selected Readings** after each chapter reflect the most current literature and these have been updated and extended where necessary. We have added over 120 new references and deleted many that are no longer appropriate. Although we have always included references to the pedagogical literature, you will note that we have added quite a few more references of this type. Students now have easy access to these papers and they are often more informative than advanced papers in the purely scientific literature.
- **Art** is an important component of a good textbook. Our art program has been extensively revised, with many new photos to illustrate concepts explained in the text; new and updated ribbon art, and improved versions of many figures. Many of the new photos are designed to attract and/or hold the students attention. They can be powerful memory aids and some of them are used to lighten up the subject in a way that is rarely seen in other textbooks (see page 204). We believe that the look and feel of the book has been much improved, making it more appealing to students without sacrificing any of the rigor and accuracy that has been a hallmark of previous editions.

A focus on principles

There are, in essence, two kinds of biochemistry textbooks: those for reference and those for teaching. It is difficult for one book to be both as it is those same thickets of detail sought by the professional that ensnare the struggling novice on his or her first trip through the forest. This text is unapologetically a text for teaching. It has been designed to foster student understanding and is not an encyclopedia of biochemistry. This book focuses unwaveringly on teaching basic principles and concepts, each principle supported by carefully chosen examples. We really do try to get students to see the forest and not the trees!

Because of this focus, the material in this book can be covered in a two-semester course without having to tell students to skip certain chapters or certain sections. The book is also suitable for a one-semester course that concentrates on certain aspects of biochemistry where some subjects are not covered. Instructors can be confident that the core principles and concepts are explained thoroughly and correctly.

A focus on chemistry

When we first wrote this text, we decided to take the time to explain in chemical terms the principles that we want to emphasize. In fact, one of these principles is to show students that life obeys the fundamental laws of physics and chemistry. To that end, we offer chemical explanations of most biochemical reactions, including mechanisms that tell students how and why things happen.

We are particularly proud of our explanations of oxidation-reduction reactions since these are extremely important in so many contexts. We describe electron movements in the early chapters, explain reduction potentials in Chapter 10 and use this understanding to teach about chemiosmotic theory and protonmotive force in Chapter 14 (Electron Transport and ATP Synthesis). The concept is reinforced in the chapter on photosynthesis.

A focus on biology

While we emphasize chemistry, we also stress the bio in biochemistry. We point out that biochemical systems evolve and that the reactions that occur in some species are variations on a larger theme. In this edition, we increase our emphasis on the similarities of

Downloaded From Vet-e-Books.com

prokaryotic and eukaryotic systems while we continue to avoid making generalizations about all organisms based on reactions that occur in a few.

The evolutionary, or comparative, approach to teaching biochemistry focuses attention on fundamental concepts. The evolutionary approach differs in many ways from other pedagogical methods such as an emphasis on fuel metabolism. The evolutionary approach usually begins with a description of simple fundamental principles or pathways or processes. These are often the pathways found in bacteria. As the lesson proceeds, the increasing complexity seen in some other species is explained. At the end of a chapter we are ready to describe the unique features of the process found in complex multicellular species, such as humans.

Our approach entails additional changes that distinguish us from other textbooks. When introducing a new chapter, such as lipid metabolism, amino acid metabolism, and nucleotide metabolism, most other textbooks begin by treating the molecules as potential food for humans. We start with the biosynthesis pathways since those are the ones fundamental to all organisms. Then we describe the degradation pathways and end with an explanation of how they realte to fuel metabolism. This biosynthesis first organization applies to all the major components of a cell (proteins, nucleotides, nucleic acids, lipids, amino acids) except carbohydrates where we continue to describe glycolysis ahead of gluconeogenesis. We do, however, emphasize that gluconeogenesis is the original, primitive pathway and glycolysis evolved later.

This has always been the way DNA replication, transcription, and translation have been taught. In this book we extend this successful strategy to all the other topics in biochemistry. The chapter on photosynthe sis is an excellent example of how it works in practice.

In some cases the emphasis on evolution can lead to a profound appreciation of how complex systems came to exist. Take the citric acid cycle as an example. Students are often told that such a process cannot be the product of evolution because all the parts are needed before the cycle can function. We explain in Section 13.9 how such a pathway can evolve in a stepwise manner.

A focus on accuracy

We are proud of the fact that this is the most scientifically accurate biochemistry textbook.We have gone to great lengths to ensure that our facts are correct and our explanations of basic concepts reflect the modern consensus among active researchers. Our success is due, in large part, to the dedication of our many reviewers and editors.

The emphasis on accuracy means that we check our reactions and our nomenclature against the IUPAC/IUBMB databases. The result is balanced reactions with correct products and substrates and correct chemical nomenclature. For example, we are one of the very few textbooks that show all of the citric acid cycle reactions correctly. Previous editions of this textbook have always scored highly on the Biochemical Howlers website [**bip.cnrs-mrs.fr/bip10/howler.htm**] and we feel confident that this edition will achieve a perfect score!

We take the time and effort to accurately describe some difficult concepts such as Gibbs free energy change in a steady-state situation where most reactions are near-equlibirium reactions ($\Delta G = 0$). We present correct definitions of the Central Dogma of Molecular Biology. We don't avoid genuine areas of scientific controversy such as the validity of the Three Domain Hypothesis or the mechanism of lysozyme.

A focus on structure-function

Biochemistry is a three-dimensional science. Our inclusion of the latest computer generated images is intended to clarify the shape and function of molecules and to leave students with an appreciation for the relationship between the structure and function. Many of the protein images in this edition are new; they have been skillfully prepared by Jonathan Parrish of the University of Alberta.

We offer a number of other opportunities. For those students with access to a computer, we have included Protein Data Bank (PDB) reference numbers for the coordinates

Downloaded From Vet-eBooks.com

from which all protein images were derived. This allows students to further explore the structures on their own. In addition, we have a gallery of prepared PDB files that students can view using Chime or any other molecular viewer; these are posted on the text's TheChemistryPlace® website [**chemplace.com**] as are animations of key dynamic processes as well as visualization tutorials using Chime.

The emphasis on protein/enzyme structure is a key part of the theme of structure-function that is one of the most important concepts in biochemistry. At various places in this new edition we have added material to emphasize this relationship and to develop it to a greater extent than we have in the past. Some of the most important reactions in the cell, such as the Q-cycle, cannot be properly understood without understanding the structure of the enzyme that catalyzes them. Similarly, understanding the properties of double-stranded DNA is essential to understanding how it serves as the storehouse of biological information.

Walkthrough of features with some visuals

Interests

Biochemistry is at the root of a number of related sciences, including medicine, forensic science, biotechnology, and bioengineering; there are many interesting stories to tell. Throughout the text, you will find boxes that relate biochemistry to other topics. Some of them are intended to be humorous and help students relate to the material.

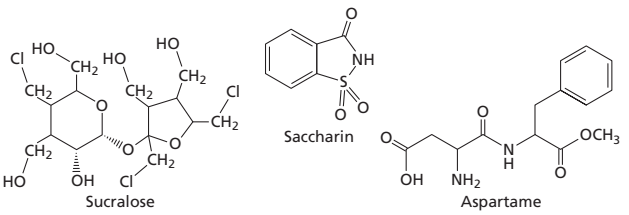
BOX 8.1 THE PROBLEM WITH CATS

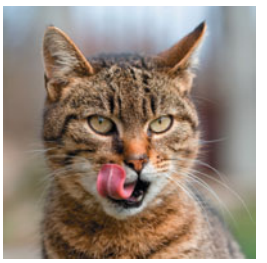
One of the characteristics of sugars is that they taste sweet. You certainly know the taste of sucrose and you probably know that fructose and lactose also taste sweet. So do many of the other sugars and their derivatives, although we don't recommend that you go into a biochemistry lab and start tasting all the carbohydrates in those white plastic bottles on the shelves.

Sweetness is not a physical property of molecules. It's a subjective interaction between a chemical and taste receptors in your mouth. There are five different kinds of taste receptors: sweet, sour, salty, bitter, and umami (umami is like the taste of glutamate in monosodium glutamate). In order to trigger the sweet taste, a molecule like sucrose has to bind to the receptor and initiate a response that eventually makes it to your brain. Sucrose elicits a moderately strong response that serves as the standard for sweetness. The response to fructose is almost twice as strong and the response to lactose is only about one-fifth as strong as that of sucrose. Artificial sweeteners such as saccharin (Sweet'N Low®), sucralose (Splenda®), and aspartame (NutraSweet®) bind to the sweetness receptor and cause the sensation of sweetness. They are hundreds of times more sweet than sucrose.

The sweetness receptor is encoded by two genes called *Tas1r2* and *Tas1r3*. We don't know how sucrose and the other ligands bind to this receptor even though this is a very active area of research. In the case of sucrose and the artifical sweeteners, how can such different molecules elicit the taste of sweet?

Cats, including lions, tigers and cheetahs, do not have a functional *Tas1r2* gene. It has been converted to a pseudogene because of a 247 bp deletion in exon 3. It's very likely that your pet cat has never experienced the taste of sweetness. That explains a lot about cats.





▲ **Cats are carnivores.** They probably can't taste sweetness.

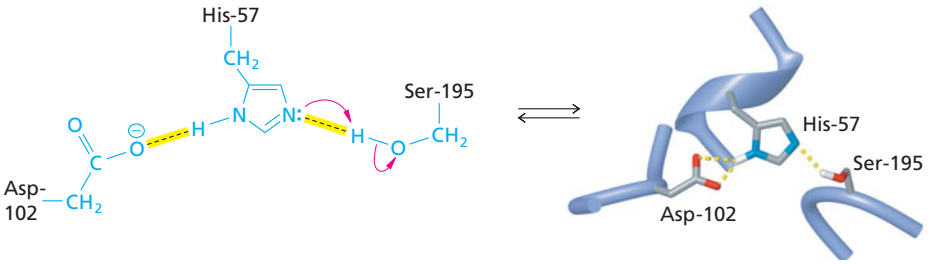
Downloaded From Vet-eBooks.com

Key Concepts

To help guide students to the information important in each concept, Key Concept notes have been provided in the margin highlighting this information.

Complete Explanations of the Chemistry

There are thousands of metabolic reactions in a typical organism. You might try to memorize them all but eventually you will run out of memory. What's more, memorization will not help you if you encounter something you haven't seen before. In this book, we show you some of the basic mechanisms of enzyme-catalyzed reactions—an extension of what you learned in organic chemistry. If you understand the mechanism, you'll understand the chemistry. You'll have less to memorize, and you'll retain the information more effectively.



KEY CONCEPT

The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the direction of a reaction when the concentrations of all products and reactants are at 1 M concentration. These conditions will never occur in living cells. Biochemists are only interested in actual Gibbs free energy changes (ΔG), which are usually close to zero. The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the relative concentrations of reactants and products when the reaction reaches equilibrium.

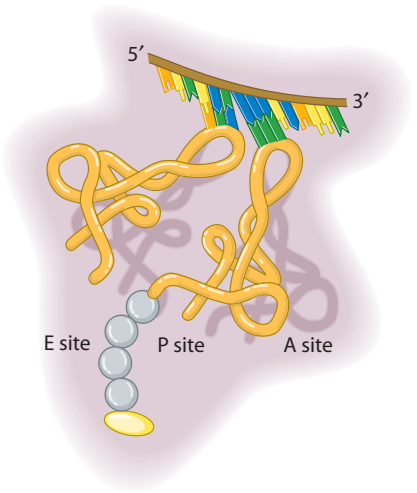
Margin Notes

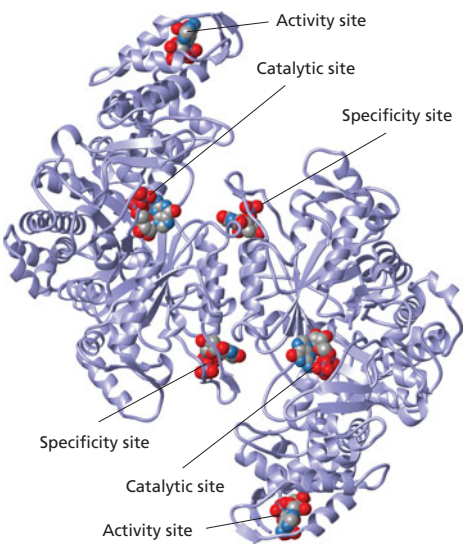
There is a great deal of detail in biochemistry but we want you to see both the forest and the trees. When we need to cross-reference something discussed earlier in the book, or something that we will come back to later, we put it in the margin. Backward references offer a review of concepts you may have forgotten. Forward references will help you see the big picture.

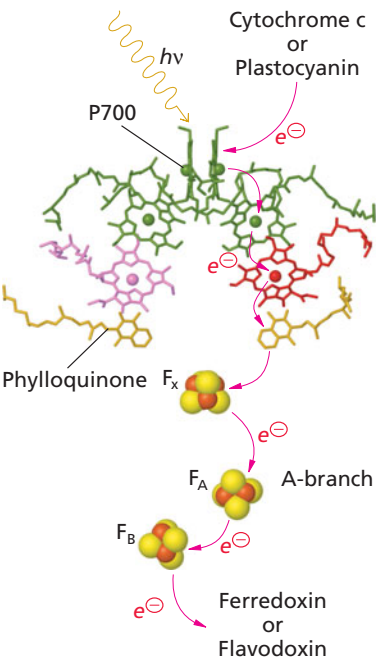
The distinction between the normal flow of information and the Central Dogma of Molecular Biology is explained in Section 1.1 and the introduction to Chapter 21.

Art

Biochemistry is a three-dimensional science and we have placed a great emphasis on helping you visualize abstract concepts and molecules too small to see. We have tried to make illustrative figures both informative and beautiful.







Downloaded From Vet-eBooks.com

Sample Calculations

Sample Calculations are included throughout the text to provide a problem solving model and illustrate required calculations.

SAMPLE CALCULATION 10.2 Gibbs Free Energy Change

Q: In a rat hepatocyte, the concentrations of ATP, ADP, and P_i are 3.4 mM, 1.3 mM, and 4.8 mM, respectively. Calculate the Gibbs free energy change for hydrolysis of ATP in this cell. How does this compare to the standard free energy change?

A: The actual Gibbs free energy change is calculated according to Equation 10.10.

$$\Delta G_{reaction} = \Delta G^{o\prime}_{reaction} + RT \ln \frac{[ADP][P_i]}{[ATP]} = \Delta G^{\circ}_{reaction} + 2.303\ RT \log \frac{[ADP][P_i]}{[ATP]}$$

When known values and constants are substituted (with concentrations expressed as molar values), assuming pH7.0 and 25°C.

$$\Delta G = -32000\ \text{J mol}^{-1} + (8.31\ \text{JK}^{-1}\text{mol}^{-1})(298\ \text{K})\left[2.303 \log \frac{(1.3 \times 10^{-3})(4.8 \times 10^{-3})}{(3.4 \times 10^{-3})}\right]$$

$$\Delta G = -32000\ \text{J mol}^{-1} + (2480\ \text{J mol}^{-1})[2.303\ \ \log(1.8 \times 10^{-3})]$$

$$\Delta G = -32000\ \text{J mol}^{-1} - 16000\ \text{J mol}^{-1}$$

$$\Delta G = -48000\ \text{J mol}^{-1} = -48\ \text{kJ mol}^{-1}$$

The actual free energy change is about $1^1/_2$ times the standard free energy change.

The Organization

We adopt the metabolism-first strategy of organizing the topics in this book. This means we begin with proteins and enzymes then describe carbohydrates and lipids. This is followed by a description of intermediary metabolism and bioenergetics. The structure of nucleic acids follows the chapter on nucleotide metabolism and the information flow chapters are at the back of the book.

While we believe there are significant advantages to teaching the subjects in this order, we recognize that some instructors prefer to teach information flow earlier in the course. We have tried to make the last four chapters on nucleic acids, DNA replication, transcription, and translation less dependant on the earlier chapters but they do discuss aspects of enzymes that rely on Chapters 4, 5 and 6. Instructors may choose to introduce these last four chapters after a description of enzymes if they wish.

This book has a chapter on coenzymes unlike most other biochemistry textbooks. We believe that it is important to put more emphasis on the role of coenzymes (and vitamins) and that's why we have placed this chapter right after the two chapters on enzymes. We know that most instructors prefer to teach the individual coenzymes when specific examples come up in other contexts. We do that as well. This organization allows instructors to refer back to chapter 7 at whatever point they wish.

Student Supplements

The Study Guide for Principles of Biochemistry

by Scott Lefler
(Arizona State University) and
Allen J. Scism
(Central Missouri State University)

No student should be without this helpful resource. Contents include the following:

- carefully constructed drill problems for each chapter, including short-answer, multiple-choice, and challenge problems
- comprehensive, step-by-step solutions and explanations for all problems
- a remedial chapter that reviews the general and organic chemistry that students require for biochemistry—topics are ingeniously presented in the context of a metabolic pathway
- tables of essential data

Downloaded From Vet-e-Books.com

Chemistry Place for *Principles of Biochemistry*

An online student tool that includes 3-D modules to help visualize biochemistry and MediaLabs to investigate important issues related to its particular chapter. Please visit the site at **http://www.chemplace.com.**

Acknowledgments

We are grateful to our many talented and thoughtful reviewers who have helped shape this book.

Reviewers who helped in the Fifth Edition:

Accuracy Reviewers

Barry Ganong, Mansfield University
Scott Lefler, Arizona State
Kathleen Nolta, University of Michigan

Content Reviewers

Michelle Chang, University of California, Berkeley
Kathleen Comely, Providence College
Ricky Cox, Murray State University
Michel Goldschmidt-Clermont, University of Geneva
Phil Klebba, University of Oklahoma, Norman
Kristi McQuade, Bradley University
Liz Roberts-Kirchoff, University of Detroit, Mercy
Ashley Spies, University of Illinois
Dylan Taatjes, University of Colorado, Boulder
David Tu, Pennsylvania State University
Jeff Wilkinson, Mississippi State University
Lauren Zapanta, University of Pittsburgh

Reviewers who helped in the Fourth Edition:

Accuracy Reviewers

Neil Haave, University of Alberta
David Watt, University of Kentucky

Content Reviewers

Consuelo Alvarez, Longwood University
Marilee Benore Parsons, University of Michigan
Gary J. Blomquist, University of Nevada, Reno
Albert M. Bobst, University of Cincinnati
Kelly Drew, University of Alaska, Fairbanks
Andrew Feig, Indiana University
Giovanni Gadda, Georgia State University
Donna L. Gosnell, Valdosta State University
Charles Hardin, North Carolina State University
Jane E. Hobson, Kwantlen University College
Ramji L. Khandelwal, University of Saskatchewan
Scott Lefler, Arizona State
Kathleen Nolta, University of Michigan
Jeffrey Schineller, Humboldt State University
Richard Shingles, Johns Hopkins University
Michael A. Sypes, Pennsylvania State University
Martin T. Tuck, Ohio University
Julio F. Turrens, University of South Alabama
David Watt, University of Kentucky
James Zimmerman, Clemson University

Thank you to J. David Rawn who's work laid the foundation for this text. We would also like to thank our colleagues who have previously contributed material for particular chapters and whose careful work still inhabits this book:

Roy Baker, University of Toronto
Roger W. Brownsey, University of British Columbia
Willy Kalt, Agriculture Canada
Robert K. Murray, University of Toronto
Ray Ochs, St. John's University
Morgan Ryan, American Scientist
Frances Sharom, University of Guelph
Malcolm Watford, Rutgers, The State University of New Jersey

Putting this book together was a collaborative effort, and we would like to thank various members of the team who have helped give this project life: Jonathan Parrish, Jay McElroy, Lisa Shoemaker, and the artists of Prentice Hall; Lisa Tarabokjia, Editorial Assistant, Jessica Neumann, Associate Editor, Lisa Pierce, Assistant Editor in charge of supplements, Lauren Layn, Media Editor, Erin Gardner, Marketing Manager; and Wendy Perez, Production Editor. We would also like to thank Jeanne Zalesky, our Executive Editor at Prentice Hall.

Finally, we close with an invitation for feedback. Despite our best efforts (and a terrific track record in the previous editions), there are bound to be mistakes in a work of this size. We are committed to making this the best biochemistry text available; please know that all comments are welcome.

Laurence A. Moran
l.moran@utoronto.ca
Marc D. Perry
marc.perry@utoronto.ca

Downloaded From Vet-e-Books.com

This page intentionally left blank

Downloaded From Vet-eBooks.com



About the Authors

Laurence A. Moran

After earning his Ph.D. from Princeton University in 1974, Professor Moran spent four years at the Université de Genève in Switzerland. He has been a member of the Department of Biochemistry at the University of Toronto since 1978, specializing in molecular biology and molecular evolution. His research findings on heat-shock genes have been published in many scholarly journals. (l.moran@utoronto.ca)

H. Robert Horton

Dr. Horton, who received his Ph.D. from the University of Missouri in 1962, is William Neal Reynolds Professor Emeritus and Alumni Distinguished Professor Emeritus in the Department of Biochemistry at North Carolina State University, where he served on the faculty for over 30 years. Most of Professor Horton's research was in protein and enzyme mechanisms.

K. Gray Scrimgeour

Professor Scrimgeour received his doctorate from the University of Washington in 1961 and was a faculty member at the University of Toronto for over 30 years. He is the author of *The Chemistry and Control of Enzymatic Reactions* (1977, Academic Press), and his work on enzymatic systems has been published in more than 50 professional journal articles during the past 40 years. From 1984 to 1992, he was editor of the journal *Biochemistry and Cell Biology*. (gray@scrimgeour.ca)

Marc D. Perry

After earning his Ph.D. from the University of Toronto in 1988, Dr. Perry trained at the University of Colorado, where he studied sex determination in the nematode *C. elegans*. In 1994 he returned to the University of Toronto as a faculty member in the Department of Molecular and Medical Genetics. His research has focused on developmental genetics, meiosis, and bioinformatics. In 2008 he joined the Ontario Institute for Cancer Research. (marc.perry@utoronto.ca)

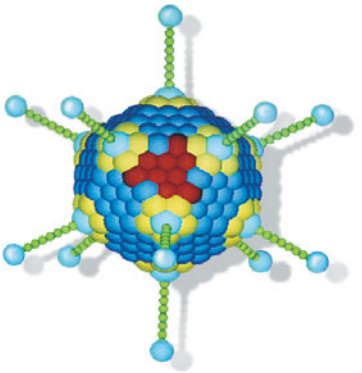
New problems and solutions for the fifth edition were created by Laurence A. Moran, University of Toronto. The remaining problems were created by Drs. Robert N. Lindquist, San Francisco State University, Marc Perry, and Diane M. De Abreu of the University of Toronto.

Downloaded From Vet-eBooks.com

This page intentionally left blank

Downloaded From Vet-eBooks.com





Introduction to Biochemistry

Anything found to be true of E. coli must also be true of elephants.

—Jacques Monod

Biochemistry is the discipline that uses the principles and language of chemistry to explain biology. Over the past 100 years biochemists have discovered that the same chemical compounds and the same central metabolic processes are found in organisms as distantly related as bacteria, plants, and humans. It is now known that the basic principles of biochemistry are common to all living organisms. Although scientists usually concentrate their research efforts on particular organisms, their results can be applied to many other species.

This book is called *Principles of Biochemistry* because we will focus on the most important and fundamental concepts of biochemistry—those that are common to most species. Where appropriate, we will point out features that distinguish particular groups of organisms.

Many students and researchers are primarily interested in the biochemistry of humans. The causes of disease and the importance of proper nutrition, for example, are fascinating topics in biochemistry. We share these interests and that's why we include many references to human biochemistry in this textbook. However, we will also try to interest you in the biochemistry of other species. As it turns out, it is often easier to understand basic principles of biochemistry by studying many different species in order to recognize common themes and patterns but a knowledge and appreciation of other species will do more than help you learn biochemistry. It will also help you recognize the fundamental nature of life at the molecular level and the ways in which species are related through evolution from a common ancestor. Perhaps future editions of this book will include chapters on the biochemistry of life on other planets. Until then, we will have to be satisfied with learning about the diverse life on our own planet.

We begin this introductory chapter with a few highlights of the history of biochemistry, followed by short descriptions of the chemical groups and molecules you will encounter throughout this book. The second half of the chapter is an overview of cell structure in preparation for your study of biochemistry.

Top: Adenovirus. Viruses consist of a nucleic acid molecule surrounded by a protein coat.

Downloaded From Vet-e-Books.com



▲ **Friedrich Wöhler (1800–1882).** Wöhler was one of the founders of biochemistry. By synthesizing urea, Wöhler showed that compounds found in living organisms could be made in the laboratory from inorganic substances.



▲ **Some of the apparatus used by Louis Pasteur in his Paris laboratory.**



▲ **Eduard Buchner (1860–1917).** Buchner was awarded the Nobel Prize in Chemistry in 1907 "for his biochemical researches and his discovery of cell-free fermentation."

1.1 Biochemistry Is a Modern Science

Biochemistry has emerged as an independent science only within the past 100 years but the groundwork for the emergence of biochemistry as a modern science was prepared in earlier centuries. The period before 1900 saw rapid advances in the understanding of basic chemical principles such as reaction kinetics and the atomic composition of molecules. Many chemicals produced in living organisms had been identified by the end of the 19th century. Since then, biochemistry has become an organized discipline and biochemists have elucidated many of the chemical processes of life. The growth of biochemistry and its influence on other disciplines will continue in the 21st century.

In 1828, Friedrich Wöhler synthesized the organic compound urea by heating the inorganic compound ammonium cyanate.

$$NH_4(OCN) \xrightarrow{\text{Heat}} H_2N-\overset{\displaystyle O}{\overset{\|}{C}}-NH_2 \qquad (1.1)$$

This experiment showed for the first time that compounds found exclusively in living organisms could be synthesized from common inorganic substances. Today we understand that the synthesis and degradation of biological substances obey the same chemical and physical laws as those that predominate outside of biology. No special or "vitalistic" processes are required to explain life at the molecular level. Many scientists date the beginnings of biochemistry to Wöhler's synthesis of urea, although it would be another 75 years before the first biochemistry departments were established at universities.

Louis Pasteur (1822–1895) is best known as the founder of microbiology and an active promoter of germ theory. But Pasteur also made many contributions to biochemistry including the discovery of stereoisomers.

Two major breakthroughs in the history of biochemistry are especially notable—the discovery of the roles of enzymes as catalysts and the role of nucleic acids as information-carrying molecules. The very large size of proteins and nucleic acids made their initial characterization difficult using the techniques available in the early part of the 20th century. With the development of modern technology we now know a great deal about how the structures of proteins and nucleic acids are related to their biological functions.

The first breakthrough—identification of enzymes as the catalysts of biological reactions—resulted in part from the research of Eduard Buchner. In 1897 Buchner showed that extracts of yeast cells could catalyze the fermentation of the sugar glucose to alcohol and carbon dioxide. Previously, scientists believed that only living cells could catalyze such complex biological reactions.

The nature of biological catalysts was explored by Buchner's contemporary, Emil Fischer. Fischer studied the catalytic effect of yeast enzymes on the hydrolysis (breakdown by water) of sucrose (table sugar). He proposed that during catalysis an enzyme and its reactant, or substrate, combine to form an intermediate compound. He also proposed that only a molecule with a suitable structure can serve as a substrate for a given enzyme. Fischer described enzymes as rigid templates, or locks, and substrates as matching keys. Researchers soon realized that almost all the reactions of life are catalyzed by enzymes and a modified lock-and-key theory of enzyme action remains a central tenet of modern biochemistry.

Another key property of enzyme catalysis is that biological reactions occur much faster than they would without a catalyst. In addition to speeding up the rates of reactions, enzyme catalysts produce very high yields with few, if any, by-products. In contrast, many catalyzed reactions in organic chemistry are considered acceptable with yields of 50% to 60%. Biochemical reactions must be more efficient because by-products can be toxic to cells and their formation would waste precious energy. The mechanisms of catalysis are described in Chapter 5.

The last half of the 20th century saw tremendous advances in the area of structural biology, especially the structure of proteins. The first protein structures were solved in the 1950s and 1960s by scientists at Cambridge University (United Kingdom) led by

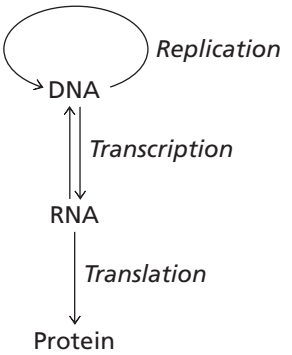
Downloaded From Vet-eBooks.com

John C. Kendrew and Max Perutz. Since then, the three-dimensional structures of several thousand different proteins have been determined and our understanding of the complex biochemistry of proteins has increased enormously. These rapid advances were made possible by the availability of larger and faster computers and new software that could carry out the many calculations that used to be done by hand using simple calculators. Much of modern biochemistry relies on computers.

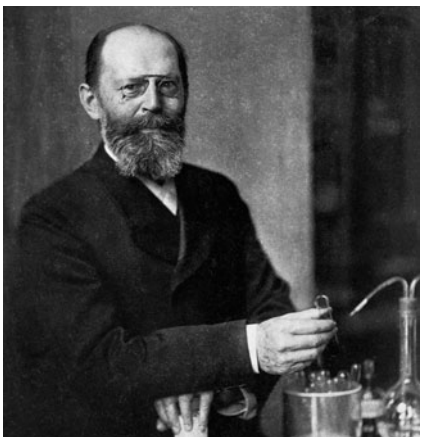
The second major breakthrough in the history of biochemistry—identification of nucleic acids as information molecules—came a half-century after Buchner's and Fischer's experiments. In 1944 Oswald Avery, Colin MacLeod, and Maclyn McCarty extracted deoxyribonucleic acid (DNA) from a pathogenic strain of the bacterium *Streptococcus pneumoniae* and mixed the DNA with a nonpathogenic strain of the same organism. The nonpathogenic strain was permanently transformed into a pathogenic strain. This experiment provided the first conclusive evidence that DNA is the genetic material. In 1953 James D. Watson and Francis H. C. Crick deduced the three-dimensional structure of DNA. The structure of DNA immediately suggested to Watson and Crick a method whereby DNA could reproduce itself, or replicate, and thus transmit biological information to succeeding generations. Subsequent research showed that information encoded in DNA can be transcribed to ribonucleic acid (RNA) and then translated into protein.

The study of genetics at the level of nucleic acid molecules is part of the discipline of molecular biology and molecular biology is part of the discipline of biochemistry. In order to understand how nucleic acids store and transmit genetic information, you must understand the structure of nucleic acids and their role in information flow. You will find that much of your study of biochemistry is devoted to considering how enzymes and nucleic acids are central to the chemistry of life.

As Crick predicted in 1958, the normal flow of information from nucleic acid to protein is not reversible. He referred to this unidirectional information flow from nucleic acid to protein as the *Central Dogma of Molecular Biology*. The term "Central Dogma" is often misunderstood. Strictly speaking, it does not refer to the overall flow of information shown in the figure. Instead, it refers to the fact that once information in nucleic acids is transferred to protein it cannot flow backwards from protein to nucleic acids.



▲ **Information flow in molecular biology.** The flow of information is normally from DNA to RNA. Some RNAs (messenger RNAs) are translated. Some RNA can be reverse transcribed back to DNA but according Crick's Central Dogma of Molecular Biology the transfer of information from nucleic acid (e.g., mRNA) to protein is irreversible.



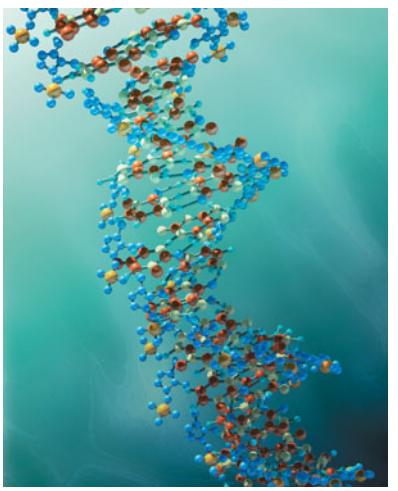
▲ **Emil Fischer (1852–1919).** Fischer made many contributions to our understanding of the structures and functions of biological molecules. He received the Nobel Prize in Chemistry in 1902 "in recognition of the extraordinary services he has rendered by his work on sugar and purine synthesis."

1.2 The Chemical Elements of Life

Six nonmetallic elements—carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur—account for more than 97% of the weight of most organisms. All these elements can form stable covalent bonds. The relative amounts of these six elements vary among organisms. Water is a major component of cells and accounts for the high percentage (by weight) of oxygen. Carbon is much more abundant in living organisms than in the rest of the universe. On the other hand, some elements, such as silicon, aluminum, and iron, are very common in the Earth's crust but are present only in trace amounts in cells. In addition to the standard six elements (CHNOPS), there are 23 other elements commonly found in living organisms (Figure 1.1). These include five ions that are essential in all species: calcium (Ca^{2+}), potassium (K^{+}), sodium (Na^{+}), magnesium (Mg^{2+}), and chloride (Cl^{-}) Note that the additional 23 elements account for only 3% of the weight of living organisms.

Most of the solid material of cells consists of carbon-containing compounds. The study of such compounds falls into the domain of organic chemistry. A course in organic chemistry is helpful in understanding biochemistry because there is considerable overlap between the two disciplines. Organic chemists are more interested in reactions that take place in the laboratory, whereas biochemists would like to understand how reactions occur in living cells.

Figure 1.2a shows the basic types of organic compounds commonly encountered in biochemistry. Make sure you are familiar with these terms because we will be using them repeatedly in the rest of this book.



▲ **DNA encodes most of the information required in living cells.**

Downloaded From Vet-e-Books.com

IA												IIIA	IVA	VA	VIA	VIIA	0
1 H 1.008	IIA																2 He 4.003
3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
11 Na 22.99	12 Mg 24.31	IIIB	IVB	VB	VIB	VIIB	VIIIB			IB	IIB	13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.07	17 Cl 35.45	18 Ar 39.95
19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.87	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.93	28 Ni 58.69	29 Cu 63.55	30 Zn 65.39	31 Ga 69.72	32 Ge 72.61	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc (98)	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.6	53 I 126.9	54 Xe 131.3
55 Cs 132.9	56 Ba 137.3	57 * La 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.8	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (209)	85 At (210)	86 Rn (222)
87 Fr (223)	88 Ra (226)	89** Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (264)	108 Hs (265)	109 Mt (268)	110 (269)	111 (272)	112 (277)	113	114 (285)	115	116 (289)	117	118 (293)

58* Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (145)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.0	71 Lu 175.0
90** Th 232.0	91 Pa 231	92 U 238.0	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (262)

▲ **Figure 1.1**
Periodic Table of the Elements. The important elements found in living cells are shown in color. The red elements (CHNOPS) are the six abundant elements. The five essential ions are purple. The trace elements are shown in dark blue (more common) and light blue (less common).

The synthesis of RNA (transcription) and protein (translation) are described in Chapters 21 and 22, respectively.

KEY CONCEPT

More than 97% of the weight of most organisms is made up of only six elements: carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur (CHNOPS).

KEY CONCEPT

Living things obey the standard laws of physics and chemistry. No "vitalistic" force is required to explain life at the molecular level.

Biochemical reactions involve specific chemical bonds or parts of molecules called functional groups (Figure 1.2b). We will encounter several common linkages in biochemistry (Figure 1.2c). Note that all these linkages consist of several different atoms and individual bonds between atoms. We will learn more about these compounds, functional groups, and linkages throughout this book. Ester and ether linkages are common in fatty acids and lipids. Amide linkages are found in proteins. Phosphate ester and phosphoanhydride linkages occur in nucleotides.

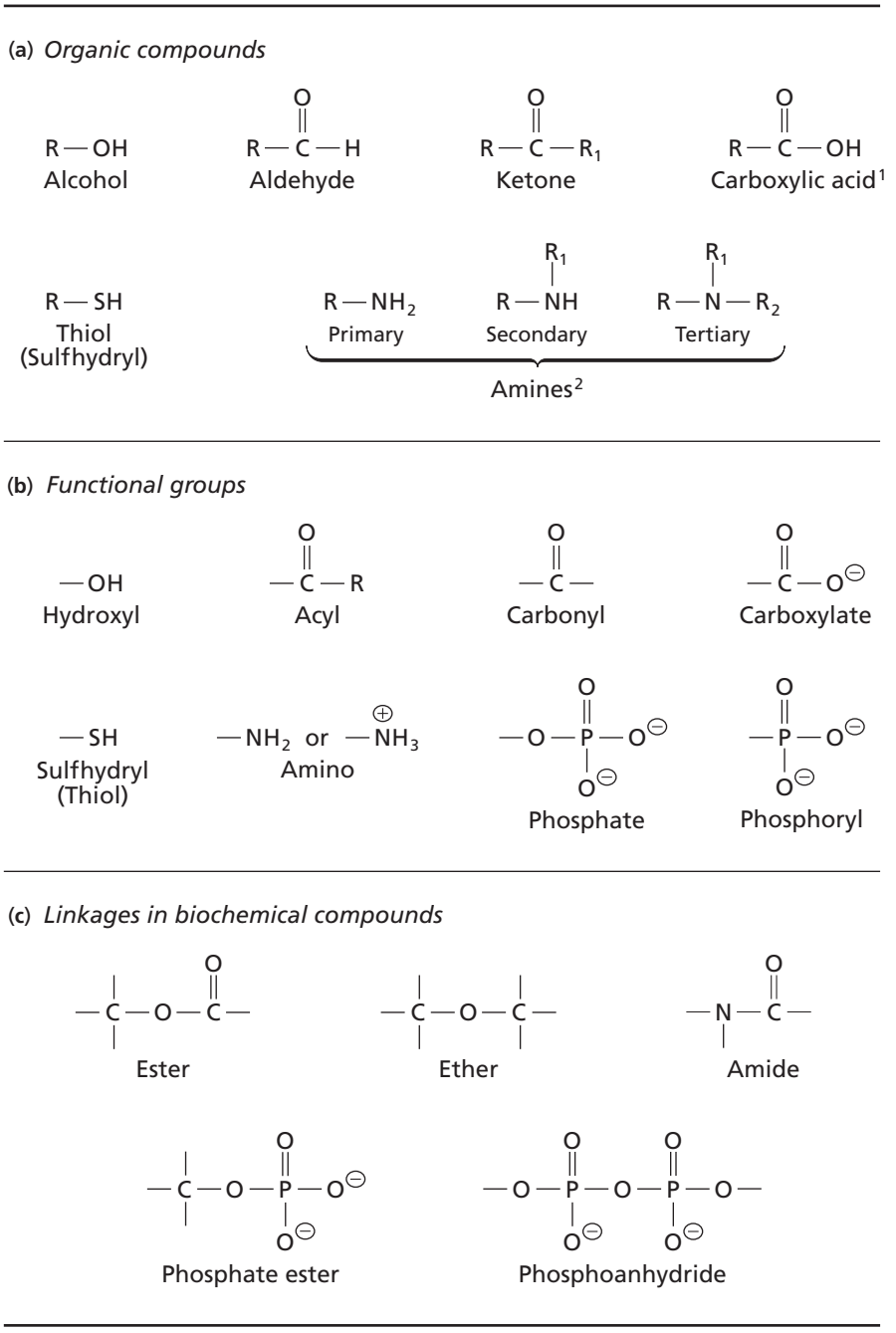
An important theme of biochemistry is that the chemical reactions occurring inside cells are the same kinds of reactions that take place in a chemistry laboratory. The most important difference is that almost all reactions in living cells are catalyzed by enzymes and thus proceed at very high rates. One of the main goals of this textbook is to explain how enzymes speed up reactions without violating the fundamental reaction mechanisms of organic chemistry.

The catalytic efficiency of enzymes can be observed even when the enzymes and reactants are isolated in a test tube. Researchers often find it useful to distinguish between biochemical reactions that take place in an organism (*in vivo*) and those that occur under laboratory conditions (*in vitro*).

1.3 Many Important Macromolecules Are Polymers

In addition to numerous small molecules, much of biochemistry deals with very large molecules that we refer to as macromolecules. Biological macromolecules are usually a form of polymer created by joining many smaller organic molecules, or monomers, via condensation (removal of the elements of water). In some cases, such as certain carbohydrates, a single monomer is repeated many times; in other cases, such as proteins and nucleic acids, a variety of different monomers is connected in a particular order. Each monomer of a given polymer is added by repeating the same enzyme-catalyzed reaction.

Downloaded From Vet-eBooks.com



◄ **Figure 1.2**
General formulas of **(a)** organic compounds, **(b)** functional groups, and **(c)** linkages common in biochemistry. R represents an alkyl group ($CH_3—(CH_2)_n—$).

[1] Under most biological conditions, carboxylic acids exist as carboxylate anions: $R—C(=O)—O^{\ominus}$

[2] Under most biological conditions, amines exist as ammonium ions: $R—NH_3^{\oplus}$, $R—NH_2^{\oplus}(R_1)$ and $R—NH^{\oplus}(R_1)—R_2$

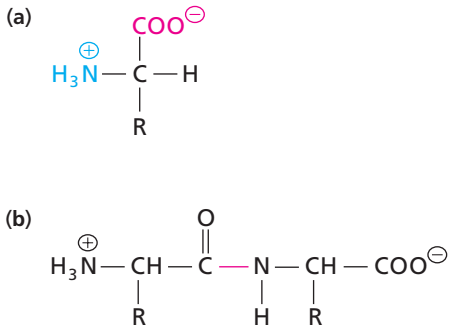
Thus, all of the monomers, or residues, in a macromolecule are aligned in the same direction and the ends of the macromolecule are chemically distinct.

Macromolecules have properties that are very different from those of their constituent monomers. For example, starch is a polymer of the sugar glucose but it is not soluble in water and does not taste sweet. Observations such as this have led to the general principle of the hierarchical organization of life. Each new level of organization results in properties that cannot be predicted solely from those of the previous level. The levels of complexity, in increasing order, are: atoms, molecules, macromolecules, organelles, cells, tissues, organs, and whole organisms. (Note that many species lack one or more of these levels of complexity. Single-celled organisms, for example, do not have tissues and organs.) The following sections briefly describe the principal types of macromolecules and how their sequences of residues or three-dimensional shapes grant them unique properties.

Downloaded From Vet-e-Books.com

In discussing molecules and macromolecules we will often refer to the **molecular weight** of a compound. A more precise term for molecular weight is **relative molecular mass** (abbreviated M_r). It is the mass of a molecule relative to one-twelfth (1/12) the mass of an atom of the carbon isotope ^{12}C. (The atomic weight of this isotope has been defined as exactly 12 atomic mass units. Note that the atomic weight of carbon shown in the Periodic Table represents the average of several different isotopes, including ^{13}C and ^{14}C.) Because M_r is a relative quantity, it is dimensionless and has no units associated with its value. The relative molecular mass of a typical protein, for example, is 38,000 (M_r = 38,000). The absolute molecular mass of a compound has the same magnitude as the molecular weight except that it is expressed in units called daltons (1 dalton = 1 atomic mass unit). The molecular mass is also called the molar mass because it represents the mass (measured in grams) of 1 mole, or 6.022×10^{23} molecules. The molecular mass of a typical protein is 38,000 daltons, which means that 1 mole weighs 38 kilograms. The main source of confusion is that the term "molecular weight" has become common jargon in biochemistry although it refers to relative molecular mass and not to weight. It is a common error to give a molecular weight in daltons when it should be dimensionless. In most cases, this isn't a very important mistake but you should know the correct terminology.

The relative molecular mass (M_r) of a molecule is a dimensionless quantity referring to the mass of a molecule relative to one-twelfth (1/12) the mass of an atom of the carbon isotope ^{12}C. Molecular weight (M.W.) is another term for relative molecular mass.



▲ **Figure 1.3**
Structure of an amino acid and a dipeptide. **(a)** Amino acids contain an amino group (blue) and a carboxylate group (red). Different amino acids contain different side chains (designated —R). **(b)** A dipeptide is produced when the amino group of one amino acid reacts with the carboxylate group of another to form a peptide bond (red).

KEY CONCEPT

Biochemical molecules are three-dimensional objects.

A. Proteins

Twenty common amino acids are incorporated into proteins in all cells. Each amino acid contains an amino group and a carboxylate group, as well as a side chain (R group) that is unique to each amino acid (Figure 1.3a). The amino group of one amino acid and the carboxylate group of another are condensed during protein synthesis to form an amide linkage, as shown in Figure 1.3b. The bond between the carbon atom of one amino acid residue and the nitrogen atom of the next residue is called a peptide bond. The end-to-end joining of many amino acids forms a linear polypeptide that may contain hundreds of amino acid residues. A functional protein can be a single polypeptide or it can consist of several distinct polypeptide chains that are tightly bound to form a more complex structure.

Many proteins function as enzymes. Others are structural components of cells and organisms. Linear polypeptides fold into a distinct three-dimensional shape. This shape is determined largely by the sequence of its amino acid residues. This sequence information is encoded in the gene for the protein. The function of a protein depends on its three-dimensional structure, or conformation.

The structures of many proteins have been determined and several principles governing the relationship between structure and function have become clear. For example, many enzymes contain a cleft, or groove, that binds the substrates of a reaction. This cavity contains the active site of the enzyme—the region where the chemical reaction takes place. Figure 1.4a shows the structure of the enzyme lysozyme that catalyzes the hydrolysis of specific carbohydrate polymers. Figure 1.4b shows the structure of the enzyme with the substrate bound in the cleft. We will discuss the relationship between protein structure and function in Chapters 4 and 6.

There are many ways of representing the three-dimensional structures of biopolymers such as proteins. The lysozyme molecule in Figure 1.4 is shown as a cartoon where the conformation of the polypeptide chain is represented as a combination of wires, helical ribbons, and broad arrows. Other kinds of representations in the following chapters include images that show the position of every atom. Computer programs that create these images are freely available on the Internet and the structural data for proteins can be retrieved from a number of database sites. With a little practice, any student can view these molecules on a computer monitor.

B. Polysaccharides

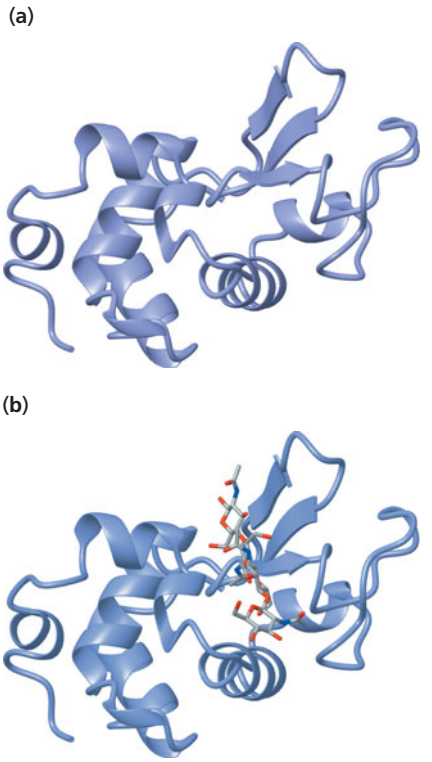
Carbohydrates, or saccharides, are composed primarily of carbon, oxygen, and hydrogen. This group of compounds includes simple sugars (monosaccharides) as well as their polymers (polysaccharides). All monosaccharides and all residues of polysaccharides contain several hydroxyl groups and are therefore polyalcohols. The most common monosaccharides contain either five or six carbon atoms.

Downloaded From Vet-e-Books.com

Sugar structures can be represented in several ways. For example, ribose (the most common five-carbon sugar) can be shown as a linear molecule containing four hydroxyl groups and one aldehyde group (Figure 1.5a). This linear representation is called a Fischer projection (after Emil Fischer). In its usual biochemical form, however, the structure of ribose is a ring with a covalent bond between the carbon of the aldehyde group (C-1) and the oxygen of the C-4 hydroxyl group, as shown in Figure 1.5b. The ring form is most commonly shown as a Haworth projection (Figure 1.5c). This representation is a more accurate way of depicting the actual structure of ribose. The Haworth projection is rotated 90° with respect to the Fischer projection and portrays the carbohydrate ring as a plane with one edge projecting out of the page (represented by the thick lines). However, the ring is not actually planar. It can adopt numerous conformations in which certain ring atoms are out-of-plane. In Figure 1.5d, for example, the C-2 atom of ribose lies above the plane formed by the rest of the ring atoms.

Some conformations are more stable than others so the majority of ribose molecules can be represented by one or two of the many possible conformations. Nevertheless, it's important to note that most biochemical molecules exist as a collection of structures with different conformations. The change from one conformation to another does not require the breaking of any covalent bonds. In contrast, the two basic forms of carbohydrate structures, linear and ring forms, do require the breaking and forming of covalent bonds.

Glucose is the most abundant six-carbon sugar (Figure 1.6a on page 8). It is the monomeric unit of cellulose, a structural polysaccharide, and of glycogen and starch, which are storage polysaccharides. In these polysaccharides, each glucose residue is joined covalently to the next by a covalent bond between C-1 of one glucose molecule and one of the hydroxyl groups of another. This bond is called a glycosidic bond. In cellulose, C-1 of each glucose residue is joined to the C-4 hydroxyl group of the next residue (Figure 1.6b). The hydroxyl groups on adjacent chains of cellulose interact noncovalently creating strong, insoluble fibers. Cellulose is probably the most abundant biopolymer on Earth because it is a major component of flowering plant stems including tree trunks. We will discuss carbohydrates further in Chapter 8.



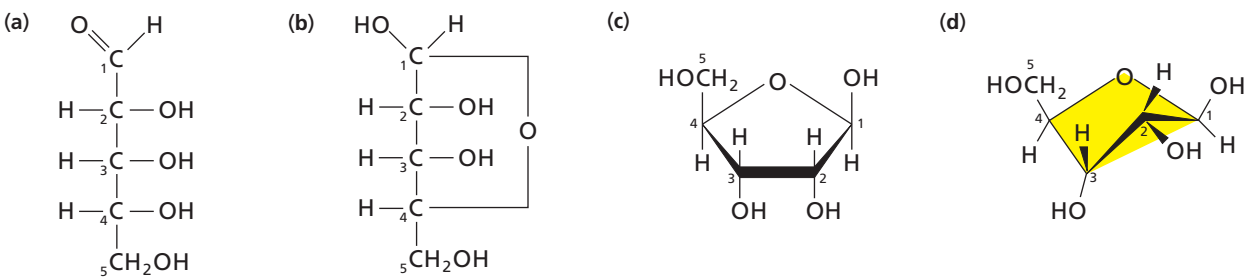
▲ **Figure 1.4 Chicken (*Gallus gallus*) eggwhite lysozyme. (a)** Free lysozyme. Note the characteristic cleft that includes the active site of the enzyme. **(b)** Lysozyme with bound substrate. [PDB 1LZC].

C. Nucleic Acids

Nucleic acids are large macromolecules composed of monomers called nucleotides. The term *polynucleotide* is a more accurate description of a single molecule of nucleic acid, just as *polypeptide* is a more accurate term than *protein* for single molecules composed of amino acid residues. The term *nucleic acid* refers to the fact that these polynucleotides were first detected as acidic molecules in the nucleus of eukaryotic cells. We

The rules for drawing a molecule as a Fischer projection are described in Section 8.1.

Conformations of monosaccharides are described in more detail in Section 8.3.

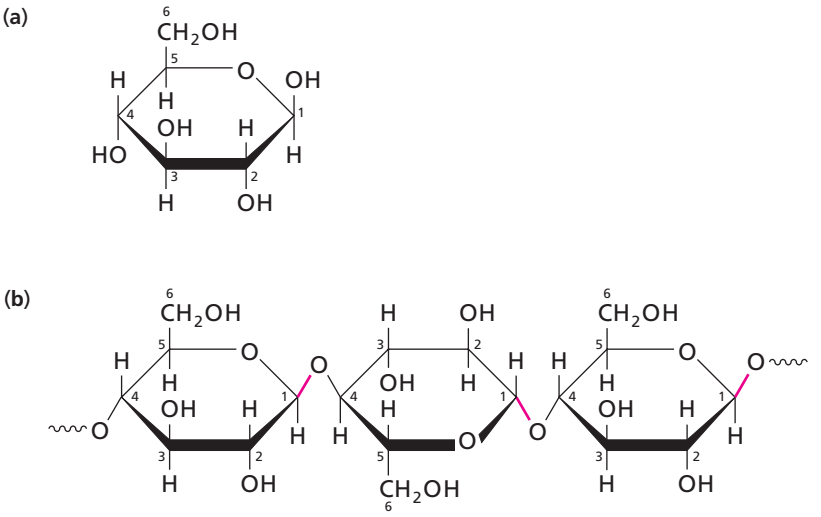


▲ **Figure 1.5**

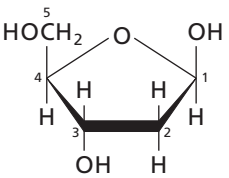
Representations of the structure of ribose. (a) In the Fischer projection, ribose is drawn as a linear molecule. **(b)** In its usual biochemical form, the ribose molecule is in a ring, shown here as a Fischer projection. **(c)** In a Haworth projection, the ring is depicted as lying perpendicular to the page (as indicated by the thick lines, which represent the bonds closest to the viewer). **(d)** The ring of ribose is not actually planar but can adopt 20 possible conformations in which certain ring atoms are out-of-plane. In the conformation shown, C-2 lies above the plane formed by the rest of the ring atoms.

Downloaded From Vet-e-Books.com

Figure 1.6 ►
Glucose and cellulose. (a) Haworth projection of glucose. **(b)** Cellulose, a linear polymer of glucose residues. Each residue is joined to the next by a glycosidic bond (red).



The structures of nucleic acids are described in Chapter 19.



▲ Figure 1.7
Deoxyribose, the sugar found in deoxyribonucleotides. Deoxyribose lacks a hydroxyl group at C-2.

now know that nucleic acids are not confined to the eukaryotic nucleus but are abundant in the cytoplasm and in prokaryotes that don't have a nucleus.

Nucleotides consist of a five-carbon sugar, a heterocyclic nitrogenous base, and at least one phosphate group. In ribonucleotides, the sugar is ribose; in deoxyribonucleotides, it is the derivative deoxyribose (Figure 1.7). The nitrogenous bases of nucleotides belong to two families known as purines and pyrimidines. The major purines are adenine (A) and guanine (G); the major pyrimidines are cytosine (C), thymine (T), and uracil (U). In a nucleotide, the base is joined to C-1 of the sugar, and the phosphate group is attached to one of the other sugar carbons (usually C-5).

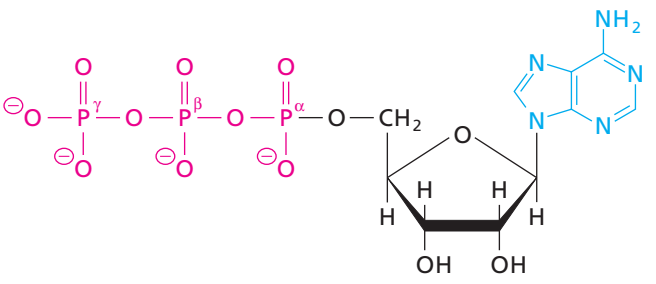
The structure of the nucleotide adenosine triphosphate (ATP) is shown in Figure 1.8. ATP consists of an adenine moiety linked to ribose by a glycosidic bond. There are three phosphoryl groups (designated α, β, and γ) esterified to the C-5 hydroxyl group of the ribose. The linkage between ribose and the α-phosphoryl group is a phosphoester linkage because it includes a carbon and a phosphorus atom, whereas the β- and γ-phosphoryl groups in ATP are connected by phosphoanhydride linkages that don't involve carbon atoms (see Figure 1.2). All phosphoanhydrides possess considerable chemical potential energy and ATP is no exception. It is the central carrier of energy in living cells. The potential energy associated with the hydrolysis of ATP can be used directly in biochemical reactions or coupled to a reaction in a less obvious way.

The role of ATP in biochemical reactions is described in Section 10.7.

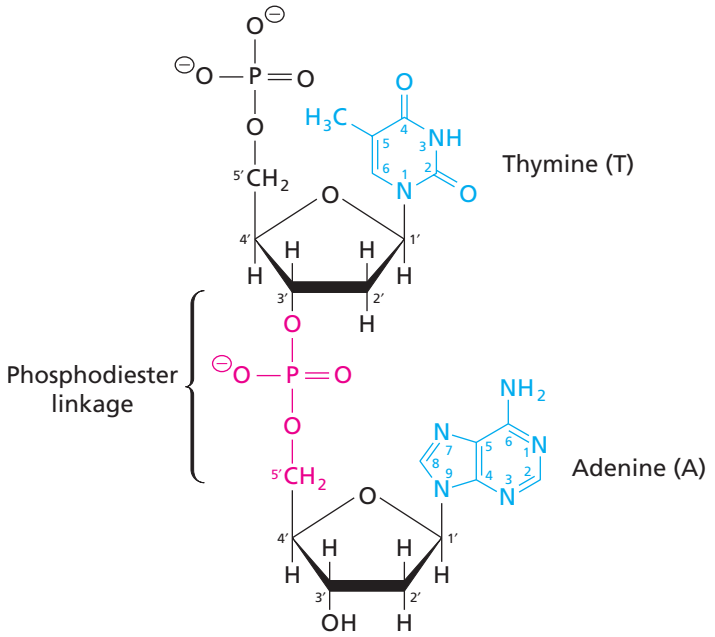
In polynucleotides, the phosphate group of one nucleotide is covalently linked to the C-3 oxygen atom of the sugar of another nucleotide creating a second phosphoester linkage. The entire linkage between the carbons of adjacent nucleotides is called a phosphodiester linkage because it contains two phosphoester linkages (Figure 1.9). Nucleic acids contain many nucleotide residues and are characterized by a backbone consisting of alternating sugars and phosphates. In DNA, the bases of two different polynucleotide strands interact to form a helical structure.

There are several ways of depicting nucleic acid structures depending on which features are being described. The ball-and-stick model shown in Figure 1.10 is ideal for showing the individual atoms and the ring structure of the sugars and the bases. In this case, the

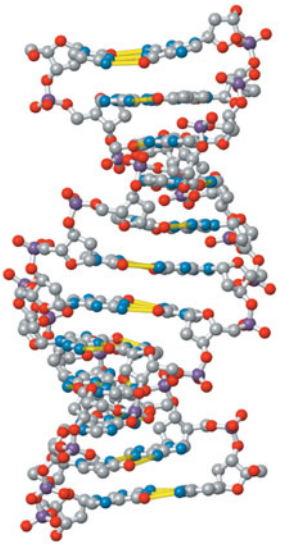
Figure 1.8 ►
Structure of adenosine triphosphate (ATP). The nitrogenous base adenine (blue) is attached to ribose (black). Three phosphoryl groups (red) are also bound to the ribose.



Downloaded From Vet-eBooks.com



◀ **Figure 1.9**
Structure of a dinucleotide. One deoxyribonucleotide residue contains the pyrimidine thymine (top), and the other contains the purine adenine (bottom). The residues are joined by a phosphodiester linkage between the two deoxyribose moieties. (The carbon atoms of deoxyribose are numbered with primes to distinguish them from the atoms of the bases thymine and adenine.)



▲ **Figure 1.10**
Short segment of a DNA molecule. Two different polynucleotides associate to form a double helix. The sequence of base pairs on the inside of the helix carries genetic information.

two helices can be traced by following the sugar–phosphate backbone emphasized by the presence of the purple phosphorus atoms surrounded by four red oxygen atoms. The individual base pairs are viewed edge-on in the interior of the molecule. We will see several other DNA models in Chapter 19.

RNA contains ribose rather than deoxyribose and it is usually a single-stranded polynucleotide. There are four different kinds of RNA molecules. Messenger RNA (mRNA) is involved directly in the transfer of information from DNA to protein. Transfer RNA (tRNA) is a smaller molecule required for protein synthesis. Ribosomal RNA (rRNA) is the major component of ribosomes. Cells also contain a heterogeneous class of small RNAs that carry out a variety of different functions. In Chapters 19 to 22, we will see how these RNA molecules differ and how their structures reflect their biological roles.

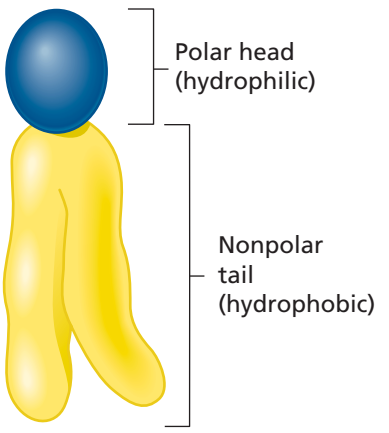
D. Lipids and Membranes

The term "lipid" refers to a diverse class of molecules that are rich in carbon and hydrogen but contain relatively few oxygen atoms. Most lipids are not soluble in water but they do dissolve in some organic solvents. Lipids often have a polar, hydrophilic (water-loving) head and a nonpolar, hydrophobic (water-fearing) tail (Figure 1.11). In an aqueous environment, the hydrophobic tails of such lipids associate while the hydrophobic heads are exposed to water, producing a sheet called a lipid bilayer. Lipid bilayers form the structural basis of all biological membranes. Membranes separate cells or compartments within cells from their environments by acting as barriers that are impermeable to most water-soluble compounds. Membranes are flexible because lipid bilayers are stabilized by noncovalent forces.

The simplest lipids are fatty acids—these are long-chain hydrocarbons with a carboxylate group at one end. Fatty acids are commonly found as part of larger molecules called glycerophospholipids consisting of glycerol 3-phosphate and two fatty acyl groups (Figure 1.12 on the next page). Glycerophospholipids are major components of biological membranes.

Other kinds of lipids include steroids and waxes. Steroids are molecules like cholesterol and many sex hormones. Waxes are common in plants and animals but perhaps the most familiar examples are beeswax and the wax that forms in your ears.

Membranes are among the largest and most complex cellular structures. Strictly speaking, membranes are aggregates, not polymers. However, the association of lipid molecules with each other creates structures that exhibit properties not shown by individual component molecules. Their insolubility in water and the flexibility of lipid aggregates give biological membranes many of their characteristics.



▲ **Figure 1.11**
Model of a membrane lipid. The molecule consists of a polar head (blue) and a nonpolar tail (yellow).

Hydrophobic interactions are discussed in Chapter 2.

Downloaded From Vet-eBooks.com

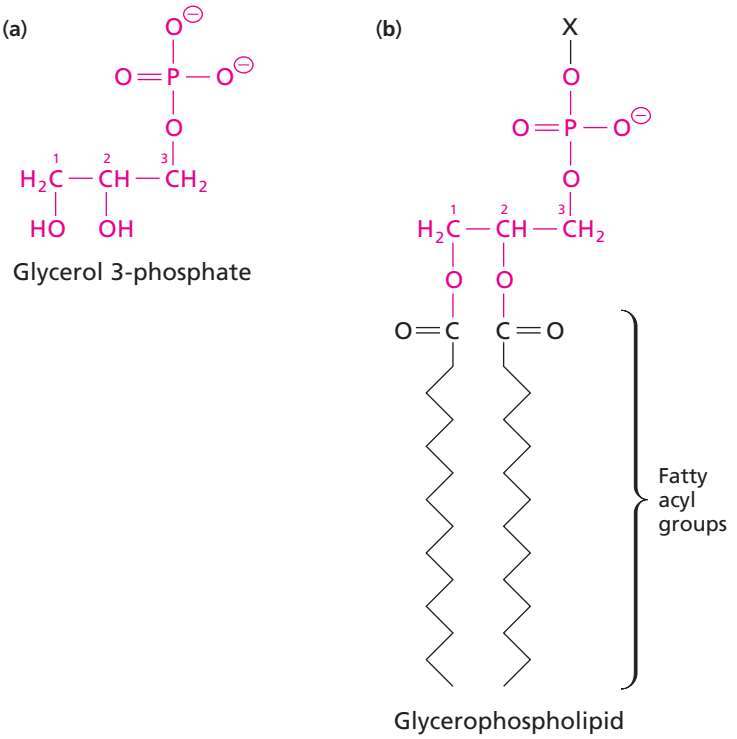


Figure 1.12 ►
Structures of glycerol 3-phosphate and a glycerophospholipid. (a) The phosphate group of glycerol 3-phosphate is polar. **(b)** In a glycerophospholipid, two nonpolar fatty acid chains are bound to glycerol 3-phosphate through ester linkages. X represents a substituent of the phosphate group.

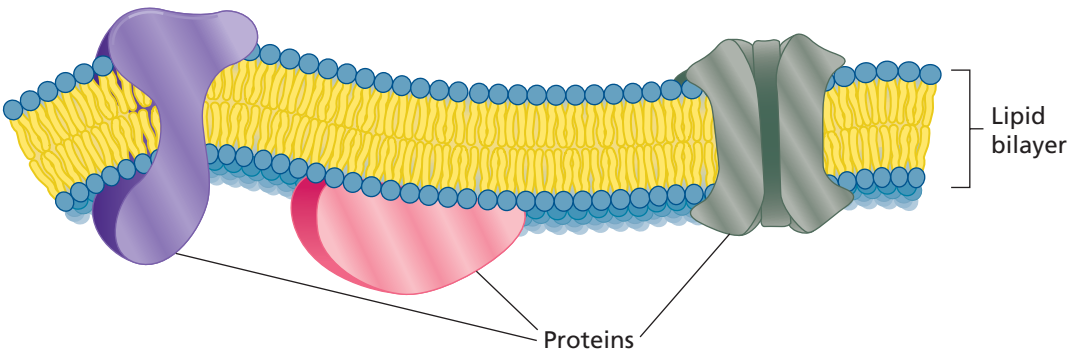
KEY CONCEPT

Most of the energy required for life is supplied by light from the sun.

Biological membranes also contain proteins as shown in Figure 1.13. Some of these membrane proteins serve as channels for the entry of nutrients and the exit of wastes. Other proteins catalyze reactions that occur specifically at the membrane surface. They are the sites of many important biochemical reactions. We will discuss lipids and biological membranes in greater detail in Chapter 9.

1.4 The Energetics of Life

The activities of living organisms do not depend solely on the biomolecules described in the preceding section and on the multitude of smaller molecules and ions found in cells. Life also requires the input of energy. Living organisms are constantly transforming energy into useful work to sustain themselves, to grow, and to reproduce. Almost all this energy is ultimately supplied by the sun.



▲ **Figure 1.13**
General structure of a biological membrane. Biological membranes consist of a lipid bilayer with associated proteins. The hydrophobic tails of individual lipid molecules associate to form the core of the membrane. The hydrophilic heads are in contact with the aqueous medium on either side of the membrane. Most membrane proteins span the lipid bilayer; others are attached to the membrane surface in various ways.

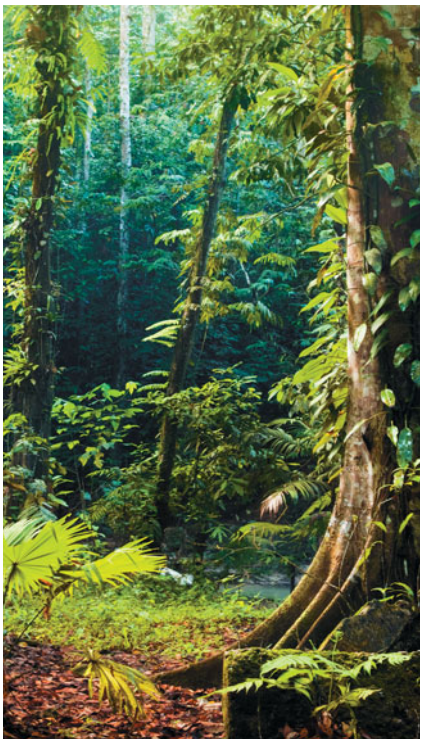
Downloaded From Vet-e-Books.com

Sunlight is captured by plants, algae, and photosynthetic bacteria and used for the synthesis of biological compounds. Photosynthetic organisms can be ingested as food and their component molecules used by organisms such as protozoa, fungi, nonphotosynthetic bacteria, and animals. These organisms cannot directly convert sunlight into useful biochemical energy. The breakdown of organic compounds in both photosynthetic and nonphotosynthetic organisms releases energy that can be used to drive the synthesis of new molecules and macromolecules.

Photosynthesis is one of the key biochemical processes that are essential for life, even though many species, including animals, benefit only indirectly. One of the byproducts of photosynthesis is oxygen. It is likely that Earth's atmosphere was transformed by oxygen-producing photosynthetic bacteria during the first several billion years of its history (a natural example of terraforming). In Chapter 15, we will discuss the amazing set of reactions that capture sunlight and use it to synthesize biopolymers.

The term *metabolism* describes the myriad reactions in which organic compounds are synthesized and degraded and useful energy is extracted, stored, and used. The study of the changes in energy during metabolic reactions is called *bioenergetics*. Bioenergetics is part of the field of thermodynamics, a branch of physical science that deals with energy changes. Biochemists have discovered that the basic thermodynamic principles that apply to energy flow in nonliving systems also apply to the chemistry of life.

Thermodynamics is a complex and highly sophisticated subject but we don't need to master all of its complexities and subtleties in order to understand how it can contribute to an understanding of biochemistry. We will avoid some of the complications of thermodynamics in this book and concentrate instead on using it to describe some biochemical principles (discussed in Chapter 10).



▲ **Sunlight on a tropical rain forest.** Plants convert sunlight and inorganic nutrients into organic compounds.

A. Reaction Rates and Equilibria

The rate, or speed, of a chemical reaction depends on the concentration of the reactants. Consider a simple chemical reaction where molecule A collides with molecule B and undergoes a reaction that produces products C and D.

$$A + B \longrightarrow C + D \qquad \textbf{(1.2)}$$

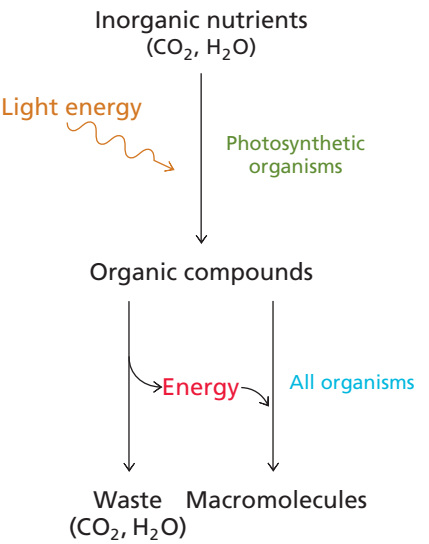
The rate of this reaction is determined by the concentrations of A and B. At high concentrations, these reactants are more likely to collide with each other; at low concentrations, the reaction might take a long time. We indicate the concentration of a reacting molecule by enclosing its symbol in square brackets. Thus, [A] means "the concentration of A"—usually expressed in moles per liter (M). The rate of the reaction is directly proportional to the product of the concentrations of A and B. This rate can be described by a proportionality constant, k, that is more commonly called a rate constant.

$$\text{rate} \propto [A][B] \qquad \text{rate} = k[A][B] \qquad \textbf{(1.3)}$$

Almost all biochemical reactions are reversible. This means that C and D can collide and undergo a chemical reaction to produce A and B. The rate of the reverse reaction will depend on the concentrations of C and D and that rate can be described by a different rate constant. By convention, the forward rate constant is k_1 and the reverse rate constant is k_{-1}. Reaction 1.4 is a more accurate way of depicting the reaction shown in Reaction 1.2.

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C + D \qquad \textbf{(1.4)}$$

If we begin a test tube reaction by mixing high concentrations of A and B, then the initial concentrations of C and D will be zero and the reaction will only proceed from left to right. The rate of the initial reaction will depend on the beginning concentrations of A and B and the rate constant k_1. As the reaction proceeds, the amount of A and B will decrease and the amount of C and D will increase. The reverse reaction will start to become significant as the products accumulate. The speed of the reverse reaction will depend on the concentrations of C and D and the rate constant k_{-1}.



▲ **Energy flow.** Photosynthetic organisms capture the energy of sunlight and use it to synthesize organic compounds. The breakdown of these compounds in both photosynthetic and nonphotosynthetic organisms generates energy needed for the synthesis of macromolecules and for other cellular requirements.

Downloaded From Vet-e-Books.com

KEY CONCEPT

The rate of a chemical reaction depends on the concentrations of the reactants. The higher the concentration, the faster the reaction.

KEY CONCEPT

Almost all biochemical reactions are reversible. When the forward and reverse reactions are equal, the reaction is at equilibrium.

At some point, the rates of the forward and reverse reactions will be equal and there will be no further change in the concentrations of A, B, C, and D. In other words, the reaction will have reached equilibrium. At equilibrium,

$$k_1[A][B] = k_{-1}[C][D] \quad (1.5)$$

In many cases we are interested in the final concentrations of the reactants and products once the reaction has reached equilibrium. The ratio of product concentrations to reactant concentrations defines the equilibrium constant, K_{eq}. The equilibrium constant is also equal to the ratio of the forward and reverse rate constants and since k_1 and k_{-1} are constants, so is K_{eq}. Rearranging Equation 1.5 gives,

$$\frac{k_1}{k_{-1}} = \frac{[C][D]}{[A][B]} = K_{eq} \quad (1.6)$$

In theory, the concentrations of products and reactants could be identical once the reaction reaches equilibrium. In that case, $K_{eq} = 1$ and the forward and reverse rate constants have the same values. In most cases the value of the equilibrium constant ranges from 10^{-3} to 10^{3} meaning that the rate of one of the reactions is much faster than the other. If $K_{eq} = 10^{3}$ then the reaction will proceed mostly to the right and the final concentrations of C and D will be much higher than the concentrations of A and B. In this case, the forward rate constant (k_1) will be 1000 times greater than the reverse rate constant (k_{-1}). This means that collisions between C and D are much less likely to produce a chemical reaction than collisions between A and B.

B. Thermodynamics

If we know the energy changes associated with a reaction or process, we can predict the equilibrium concentrations. We can also predict the direction of a reaction provided we know the initial concentrations of reactants and products. The thermodynamic quantity that provides this information is the Gibbs free energy (G), named after J. Willard Gibbs who first described this quantity in 1878.



▲ **Josiah Willard Gibbs (1839–1903).** Gibbs was one of the greatest American scientists of the 19th century. He founded the modern field of chemical thermodynamics.

It turns out that molecules in solution have a certain energy that depends on temperature, pressure, concentration, and other states. The Gibbs free energy change (ΔG) for a reaction is the difference between the free energy of the products and the free energy of the reactants. The overall Gibbs free energy change has two components known as the enthalpy change (ΔH, the change in heat content) and the entropy change (ΔS, the change in randomness). A biochemical process may generate heat or absorb it from the surroundings. Similarly, a process may occur with an increase or a decrease in the degree of disorder, or randomness, of the reactants.

Starting with an initial solution of reactants and products, if the reaction proceeds to produce more products, then ΔG must be less than zero ($\Delta G < 0$). In chemistry terms, we say that the reaction is spontaneous and energy is released. When ΔG is greater than zero ($\Delta G > 0$), the reaction requires external energy to proceed and it will not yield more products. In fact, more reactants will accumulate as the reverse reaction is favored. When ΔG equals zero ($\Delta G = 0$), the reaction is at equilibrium; the rates of the forward and reverse reactions are identical and the concentrations of the products and reactants no longer change.

We are mostly interested the overall Gibbs free energy change, expressed as

$$\Delta G = \Delta H - T\Delta S \quad (1.7)$$

KEY CONCEPT

The Gibbs free energy change (ΔG) is the difference between the free energy of the products of a reaction and that of the reactants (substrates).

where T is the temperature in Kelvin.

A series of linked processes, such as the reactions of a metabolic pathway in a cell, usually proceeds only when associated with an overall negative Gibbs free energy change. Biochemical reactions or processes are more likely to occur, both to a greater extent and more rapidly, when they are associated with an increase in entropy and a decrease in enthalpy.

Downloaded From Vet-eBooks.com

If we knew the Gibbs free energy of every product and every reactant, it would be a simple matter to calculate the Gibbs free energy change for a reaction by using Equation 1.8.

$$\Delta G_{\text{reaction}} = \Delta G_{\text{products}} - \Delta G_{\text{reactants}} \quad (1.8)$$

Unfortunately, we don't often know the absolute Gibbs free energies of every biochemical molecule. What we do know are the thermodynamic parameters associated with the *synthesis* of these molecules from simple precursors. For example, glucose can be formed from water and carbon dioxide. We don't need to know the absolute values of the Gibbs free energy of water and carbon dioxide in order to calculate the amount of enthalpy and entropy that are required to bring them together to make glucose. In fact, the heat released by the reverse reaction (breakdown of glucose to carbon dioxide and water) can be measured using a calorimeter. This gives us a value for the change in enthalpy of synthesis of glucose (ΔH). The entropy change (ΔS) for this reaction can also be determined. We can use these quantities to determine the Gibbs free energy of the reaction. The true Gibbs free energy of formation $\Delta_f G$ is the difference between the absolute free energy of glucose and that of the elements carbon, oxygen and hydrogen.

There are tables giving these Gibbs free energy values for the formation of most biological molecules. They can be used to calculate the Gibbs free energy change for a reaction in the same way that we might use absolute values as in Equation 1.9.

$$\Delta G_{\text{reaction}} = \Delta_f G_{\text{products}} - \Delta_f G_{\text{reactants}} \quad (1.9)$$

In this textbook we will often refer to the $\Delta_f G$ value as the Gibbs free energy of a compound since it can be easily used in calculations as though it were an absolute value. It can also be called just "Gibbs energy" by dropping the word "free."

There's an additional complication that hasn't been mentioned. For any reaction, including the degradation of glucose, the actual free energy change depends on the concentrations of reactants and products. Let's consider the hypothetical reaction in Equation 1.2. If we begin with a certain amount of A and B and none of the products C and D, then it's obvious that the reaction can only go in one direction, at least initially. In thermodynamic terms, $\Delta G_{\text{reaction}}$ is favorable under these conditions. The higher the concentrations of A and B, the more likely the reaction will occur. This is an important point that we will return to many times as we learn about biochemistry—the actual Gibbs free energy change in a reaction depends on the concentrations of the reactants and products.

What we need are some standard values of ΔG that can be adjusted for concentration. These standard values are the Gibbs free energy changes measured under certain conditions. By convention, the standard conditions are 25°C (298 K), 1 atm standard pressure, and 1.0 M concentration of all products and reactants. In most biochemical reactions, the concentration of $H^{\oplus}$ is important, and this is indicated by the pH, as will be described in the next chapter. The standard condition for biochemistry reactions is pH = 7.0, which corresponds to 10^{-7} M $H^{\oplus}$ (rather than 1.0 M as for other reactants and products). The Gibbs free energy change under these standard conditions is indicated by the symbol $\Delta G^{\circ\prime}$.

The actual Gibbs free energy is related to its standard free energy by

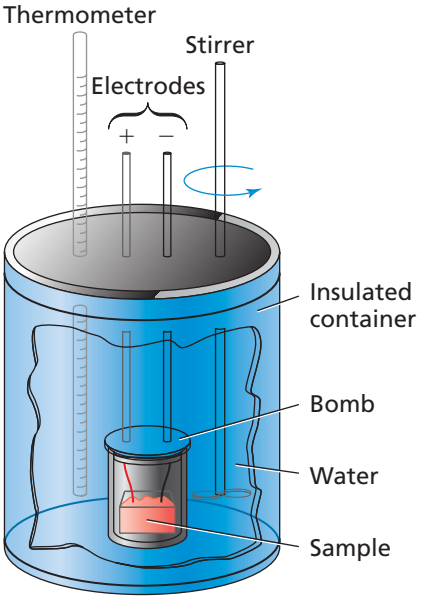
$$\Delta G_A = \Delta G^{\circ\prime}_A + RT \ln[A] \quad (1.10)$$

where R is the universal gas constant ($8.315\ \text{kJ}^{-1}\ \text{mol}^{-1}$) and T is the temperature in Kelvin. Gibbs free energy is expressed in units of kJ mol^{-1}. (An older unit is kcal mol^{-1}, which equals 4.184 kJ mol^{-1}.) The term $RT \ln[A]$ is sometimes given as $2.303\ RT \log[A]$.

C. Equilibrium Constants and Standard Gibbs Free Energy Changes

For a given reaction, such as that in Reaction 1.2, the actual Gibbs free energy change is related to the standard free energy change by

$$\Delta G_{\text{reaction}} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln \frac{[C][D]}{[A][B]} \quad (1.11)$$



▲ The heat given off during a reaction can be determined by carrying out the reaction in a sensitive calorimeter.

The importance of the relationship between ΔG and concentration is explained in Section 10.5.

KEY CONCEPT

The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the direction of a reaction when the concentrations of all products and reactants are at 1 M concentration. These conditions will never occur in living cells. Biochemists are only interested in actual Gibbs free energy changes (ΔG), which are usually close to zero. The standard Gibbs free energy change ($\Delta G^{\circ\prime}$) tells us the relative concentrations of reactants and products when the reaction reaches equilibrium.

Downloaded From Vet-eBooks.com

KEY CONCEPT

$$\Delta G = \Delta G^{\circ\prime} + RT \ln \frac{[C][D]}{[A][B]}$$

at equilibrium $\Delta G^{\circ\prime} + RT \ln K_{eq} = 0$

If the reaction has reached equilibrium, the ratio of concentrations in the last term of Equation 1.11 is, by definition, the equilibrium constant (K_{eq}). When the reaction is at equilibrium there is no net change in the concentrations of reactants and products, so the actual Gibbs free energy change is zero ($\Delta G_{reaction} = 0$). This allows us to write an equation relating the standard Gibbs free energy change and the equilibrium constant. Thus, at equilibrium,

$$\Delta G^{o\prime}_{reaction} = -RT \ln K_{eq} = -2.303\ RT \log K_{eq} \qquad (1.12)$$

This important equation relates thermodynamics and reaction equilibria. Note that it is the equilibrium constant that is related to the Gibbs free energy change and not the individual rate constants described in Equations 1.6 and 1.7. It is the *ratio* of those individual rate constants that is important and not their absolute values. The forward and reverse rates might both be very slow or very fast and still give the same ratio.

D. Gibbs Free Energy and Reaction Rates

KEY CONCEPT

The rate of a reaction is not determined by the Gibbs free energy change.

Thermodynamic considerations can tell us if a reaction is favored but do not tell how quickly a reaction will occur. We know, for example, that iron rusts and copper turns green, but these reactions may take only a few seconds or many years. That's because, the rate of a reaction depends on other factors, such as the activation energy.

Activation energies are usually depicted as a hump, or barrier, in diagrams that show the progress of a reaction from left to right. In Figure 1.14, we plot the Gibbs free energy at different stages of a reaction as it goes from reactants to products. This progress is called the reaction coordinate.

The overall change in free energy (ΔG) can be negative, as shown on the left, or positive, as shown on the right. In either case, there's an excess of energy required in order for the reaction to proceed. The difference between the top of the energy peak and the energy of the product or reactant with the highest Gibbs free energy is known as the activation energy ($\Delta G‡$).

The *rate* of this reaction depends on the nature of the reaction. Using our example from Equation 1.2, if every collision between A and B is effective, then the rate is likely to be fast. On the other hand, if the orientation of individual molecules has to be exactly right for a reaction to occur then many collisions will be nonproductive and the rate will be slower. In addition to orientation, the rate depends on the kinetic energy of the individual molecules. At any given temperature some will be moving slowly when they collide and they will not have enough energy to react. Others will be moving rapidly and will carry a lot of kinetic energy.

The activation energy is meant to reflect these parameters. It is a measure of the probability that a reaction will occur. The activation energy depends on the temperature—it is lower at higher temperatures. It also depends on the concentration of reactants—at high concentrations there will be more collisions and the rate of the reaction will be faster.

The important point is that the rate of a reaction is not predictable from the overall Gibbs free energy change. Some reactions, such as the oxidation of iron or copper, will proceed very slowly because their activation energies are high.

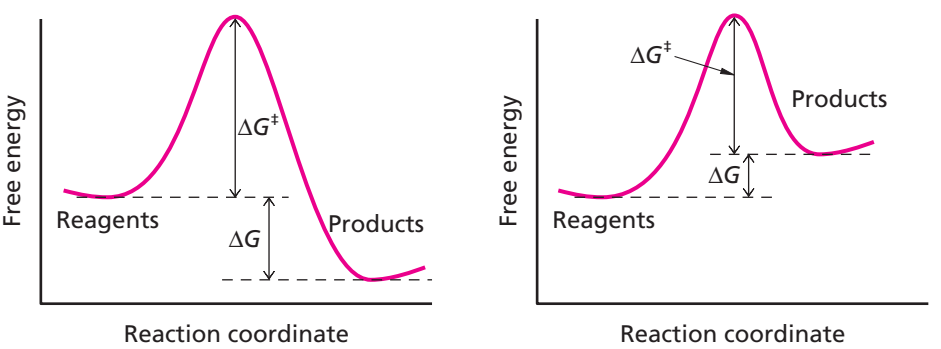


Figure 1.14 ▶
The progress of a reaction is depicted from left (reactants) to right (products). In the first diagram, the overall Gibbs free energy change is negative since the Gibbs free energy of the products is lower than that of the reactants. In order for the reaction to proceed, the reactants have to overcome an activation energy barrier ($\Delta G‡$). In the second diagram, the overall Gibbs free energy change for the reaction is positive and the minimum activation energy is smaller. This means that the reverse reaction will proceed faster than the forward reaction.

Downloaded From Vet-eBooks.com

Most of the reactions that take place inside a cell are very slow in the test tube even though they are thermodynamically favored. Inside a cell the rates of the normally slow reactions are accelerated by enzymes. The rates of enzyme-catalyzed reactions can be 10^{20} times greater than the rates of the corresponding uncatalyzed reactions. We will spend some time describing how enzymes work—it is one of the most fascinating topics in biochemistry.

1.5 Biochemistry and Evolution

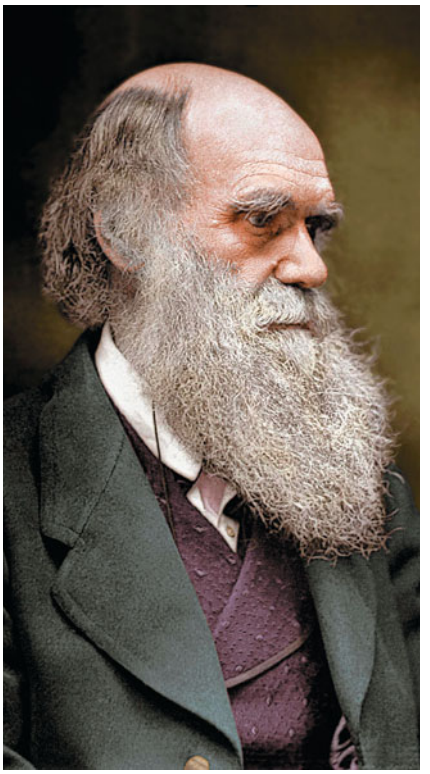
A famous geneticist, Theodosius Dobzhansky, once said, "Nothing in biology makes sense except in the light of evolution." This is also true of biochemistry. Biochemists and molecular biologists have made major contributions to our understanding of evolution at the molecular level and the evidence they have uncovered confirms and extends the data from comparative anatomy, population genetics, and paleontology. We've come a long way from the original evidence of evolution first summarized by Charles Darwin in the middle of the 19th century.

We now have a very reliable outline of the history of life and the relationships of the many diverse species in existence today. The first organisms were single cells that we would probably classify today as prokaryotes. Prokaryotes, or bacteria, do not have a membrane-bounded nucleus. Fossils of primitive bacteria-like organisms have been found in geological formations that are at least 3 billion years old. The modern species of bacteria belong to such diverse groups as the cyanobacteria, which are capable of photosynthesis, and the thermophiles, which inhabit hostile environments such as thermal hot springs.

Eukaryotes have cells that possess complex internal architecture, including a prominent nucleus. In general, eukaryotic cells are more complex and much larger than prokaryotic cells. A typical eukaryotic tissue cell has a diameter of about 25 μm (25,000 nm), whereas prokaryotic cells are typically about 1/10 that size. However, evolution has produced tremendous diversity and extreme deviations from typical sizes are common. For example, some eukaryotic unicellular organisms are large enough to be visible to the naked eye and some nerve cells in the spinal columns of vertebrates can be several feet long. There are also megabacteria that are larger than most eukaryotic cells.

All cells on Earth (prokaryotes and eukaryotes) appear to have evolved from a common ancestor that existed more than 3 billion years ago. The evidence for common ancestry includes the presence in all living organisms of common biochemical building blocks, the same general patterns of metabolism, and a common genetic code (with rare, slight variations). We will see many examples of this evidence throughout this book. The basic plan of the primitive cell has been elaborated on with spectacular inventiveness through billions of years of evolution.

The importance of evolution for a thorough understanding of biochemistry cannot be overestimated. We will encounter many pathways and processes that only make sense

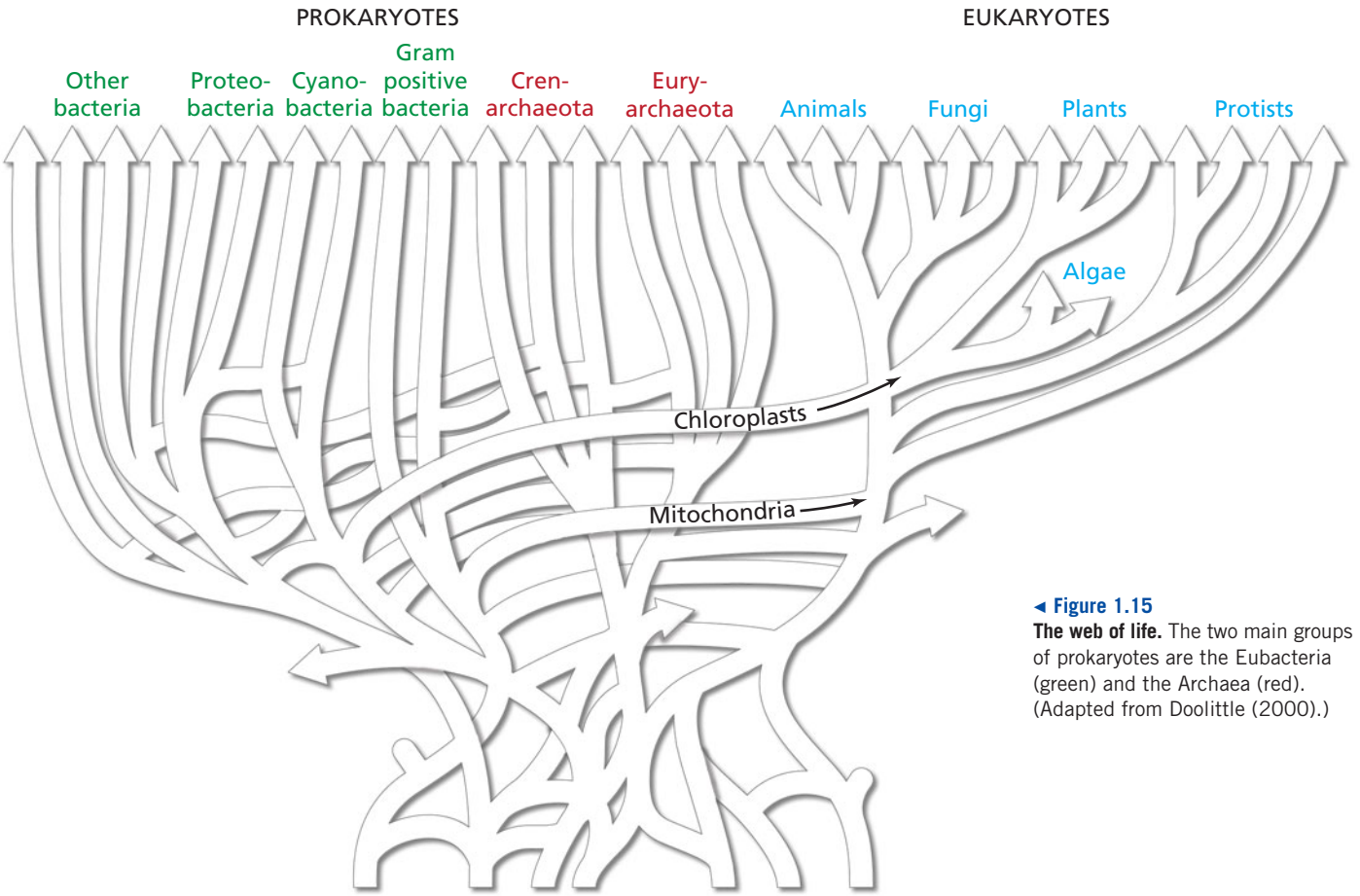


▲ **Charles Darwin (1809–1882).** Darwin published *The Origin of Species* in 1859. His theory of evolution by natural selection explains adaptive evolution.



◀ **Burgess Shale animals.** Many transitional fossils support the basic history of life that has been worked out over the past few centuries. *Pikia,* (left) is a primitive chordate from the time of the Cambrian explosion about 530 million years ago. These primitive chordates are the ancestors of all modern chordates, including humans. On the right is *Opabinia,* a primitive invertebrate.

Downloaded From Vet-eBooks.com



◀ **Figure 1.15**
The web of life. The two main groups of prokaryotes are the Eubacteria (green) and the Archaea (red). (Adapted from Doolittle (2000).)

when we appreciate that they have evolved from more primitive precursors. The evidence for evolution at the molecular level is preserved in the sequences of the genes and proteins that we will study as we learn about biochemistry. In order to fully understand the fundamental principles of biochemistry we will need to examine pathways and processes in a variety of different species including bacteria and a host of eukaryotic model organisms such as yeast, fruit flies, flowering plants, mice, and humans. The importance of comparative biochemistry has been recognized for over 100 years but its value has increased enormously in the last decade with the publication of complete genome sequences. We are now able to compare the complete biochemical pathways of many different species.

The relationship of the earliest forms of life can be determined by comparing the sequences of genes and proteins in modern species. The latest evidence shows that the early forms of unicellular life exchanged genes frequently giving rise to a complicated network of genetic relationships. Eventually, the various lineages of bacteria and archaebacteria emerged, along with primitive eukaryotes. Further evolution of eukaryotes occurred when they formed a symbiotic union with bacteria, giving rise to mitochondria and chloroplasts.

The new "web of life" view of evolution (Figure 1.15) replaces a more traditional view that separated prokaryotes into two entirely separate *domains* called Eubacteria and Archaea. That distinction is not supported by the data from hundreds of sequenced genomes so we now see prokaryotes as a single large group with many diverse subgroups, some of which are shown in the figure. It is also clear that eukaryotes contain many genes that are more closely related to the old eubacterial groups as well as a minority of genes that are closer to the old achaeal groups. The early history of life seems to be dominated by rampant gene exchange between species and this has led to a web of life rather than a tree of life.

Many students are interested in human biochemistry, particularly those aspects of biochemistry that relate to health and disease. That is an exciting part of biochemistry but in order to obtain a deep understanding of who we are, we need to know where we came from. An evolutionary perspective helps explain why we can't make some vitamins

Downloaded From Vet-e-Books.com

and amino acids and why we have different blood types and different tolerances for milk products. Evolution also explains the unique physiology of animals, which have adapted to using other organisms as a source of metabolic fuel.

1.6 The Cell Is the Basic Unit of Life

Every organism is either a single cell or is composed of many cells. Cells exist in a remarkable variety of sizes and shapes but they can usually be classified as either eukaryotic or prokaryotic, although some taxonomists continue to split prokaryotes into two groups: Eubacteria and Archaea.

A simple cell can be pictured as a droplet of water surrounded by a plasma membrane. The water droplet contains dissolved and suspended material including proteins, polysaccharides, and nucleic acids. The high lipid content of membranes makes them flexible and self-sealing. Membranes present impermeable barriers to large molecules and charged species. This property of membranes allows for much higher concentrations of biomolecules within cells than in the surrounding medium.

The material enclosed by the plasma membrane of a cell is called the cytoplasm. The cytoplasm may contain large macromolecular structures and subcellular membrane-bound organelles. The aqueous portion of the cytoplasm minus the subcellular structures is called the cytosol. Eukaryotic cells contain a nucleus and other internal membrane-bound organelles within the cytoplasm.

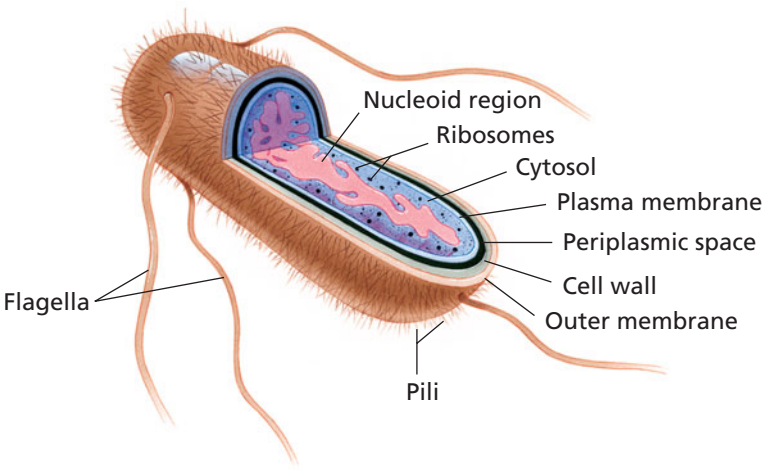
Viruses are subcellular infectious particles. They consist of a nucleic acid molecule surrounded by a protein coat and, in some cases, a membrane. Virus nucleic acid can contain as few as three genes or as many as several hundred. Despite their biological importance, viruses are not truly cells because they cannot carry out independent metabolic reactions. They propagate by hijacking the reproductive machinery of a host cell and diverting it to the formation of new viruses. In a sense, viruses are genetic parasites.

There are thousands of different viruses. Those that infect prokaryotic cells are usually called bacteriophages, or phages. Much of what we know about biochemistry is derived from the study of viruses and bacteriophages and their interaction with the cells they infect. For example, introns were first discovered in a human adenovirus like the one shown on the first page of this chapter and the detailed mapping of genes was first carried out with bacteriophage T4.

In the following two sections we will explore the structural features of typical prokaryotic and eukaryotic cells.

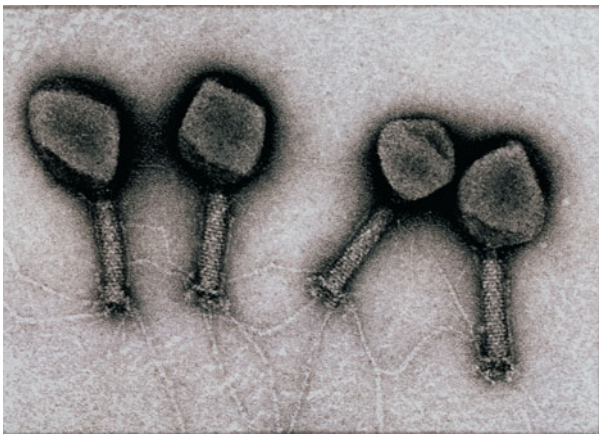
1.7 Prokaryotic Cells: Structural Features

Prokaryotes are usually single-celled organisms. The best studied of all living organisms is the bacterium *Escherichia coli* (Figure 1.16). This organism has served for half a century as a model biological system and many of the biochemical reactions described later in this book were first discovered in *E. coli*. *E. coli* is a fairly typical species of bacteria but some bacteria are as different from *E. coli* as we are from diatoms, daffodils and dragonflies.



◂ Figure 1.16
Escherichia coli. An *E. coli* cell is about 0.5 μm in diameter and 1.5 μm long. Proteinaceous fibers called flagella rotate to propel the cell. The shorter pili aid in sexual conjugation and may help *E. coli* cells adhere to surfaces. The periplasmic space is an aqueous compartment separating the plasma membrane and the outer membrane.

Downloaded From Vet-eBooks.com



► **Bacteriophage T4.** Much of our current understanding of biochemistry comes from studies of bacterial viruses such as bacteriophage T4.

Much of this diversity is apparent only at the molecular level. (See Figure 1.15 for the names of some major groups of prokaryotes.)

Prokaryotes have been found in almost every conceivable environment on Earth, from hot sulfur springs to beneath the ocean floor to the insides of larger cells. They account for a significant amount of the biomass on Earth.

Prokaryotes share a number of features in spite of their differences. They lack a nucleus—their DNA is packed in a region of the cytoplasm called the nucleoid region. Many bacterial species have only 1000 genes. From a biochemist's perspective one of the most fascinating things about bacteria is that, although their chromosomes contain a relatively small number of genes, they carry out most of the fundamental biochemical reactions found in all cells, including our own. Hundreds of bacterial genomes have been completely sequenced and it is now possible to begin to define the minimum number of enzymes that are consistent with life.

Most bacteria have no internal membrane compartments, although there are many exceptions. The plasma membrane is usually surrounded by a cell wall made of a rigid network of covalently linked carbohydrate and peptide chains. This cell wall confers the characteristic shape of an individual species of bacteria. Despite its mechanical strength, the cell wall is porous. In addition to the cell wall most bacteria, including *E. coli*, possess an outer membrane consisting of lipids, proteins, and lipids linked to polysaccharides. The space between the inner plasma membrane and the outer membrane is called the periplasmic space. It is the major membrane-bound compartment in bacteria and plays a crucial role in some important biochemical processes.

Many bacteria have protein fibers, called pili, on their outer surface. The pili serve as attachment sites for cell-cell interactions. Many species have one or more flagella. These are long, whip-like structures that can be rotated like the propeller on a boat thus driving the bacterium through its aqueous environment.

The small size of prokaryotes provides a high ratio of surface area to volume. Simple diffusion is therefore an adequate means for distributing nutrients throughout the cytoplasm. One of the prominent macromolecular structures in the cytoplasm is the ribosome—a large RNA-protein complex required for protein synthesis. All living cells have ribosomes but we will see later that bacterial ribosomes differ from eukaryotic ribosomes in significant details.



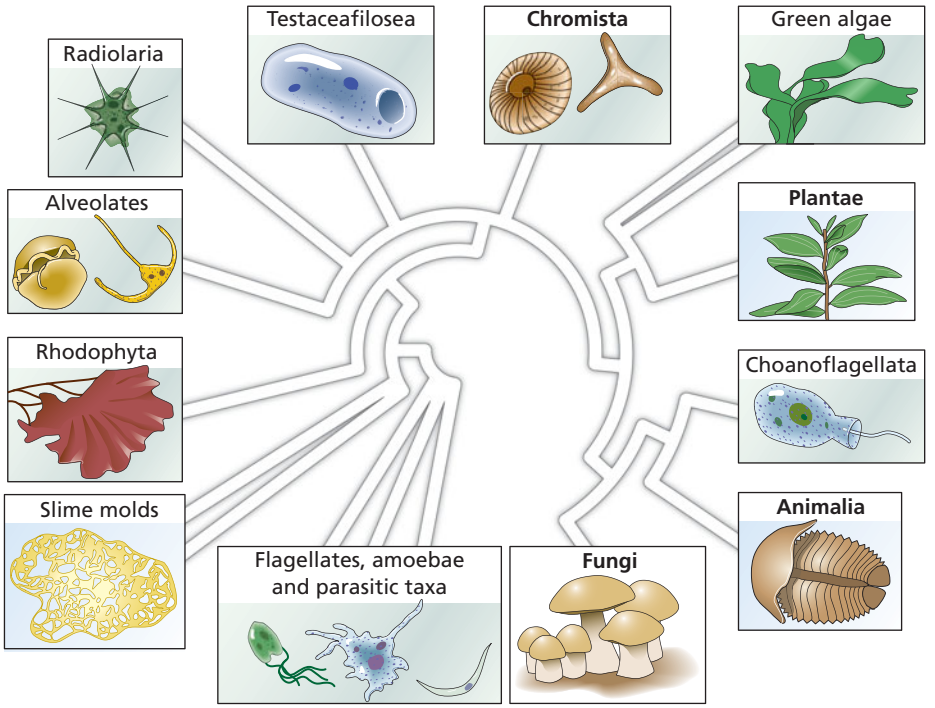
▲ **Max Delbruck and Salvatore Luria.** Max Delbruck (seated) and Salvatore Luria at the Cold Spring Harbor Laboratories in 1953. Delbruck and Luria founded the "phage group," a group of scientists who worked on the genetics and biochemistry of bacteria and bacteriophage in the 1940s, 1950s, and 1960s.

1.8 Eukaryotic Cells: Structural Features

Eukaryotes include plants, animals, fungi, and protists. Protists are mostly small, single-celled organisms that don't fit into one of the other classes. Along with bacteria these four groups make up the five kingdoms of life according to one popular classification scheme. (Older schemes retain the four eukaryotic kingdoms but divide the bacteria into Eubacteria and Archaea.)

As members of the animal kingdom we are mostly aware of other animals. As relatively large organisms we tend to focus on the large scale. Hence, we know about plants and mushrooms but not microscopic species.

Downloaded From Vet-e-Books.com



◀ **Figure 1.17**
The eukaryotic tree of life. The traditional Plantae, Animalia, and Fungi kingdoms are branches within the much larger "kingdom" of Protists.

The latest trees of eukaryotes help us understand the diversity of the protist kingdom. As shown in Figure 1.17, the animal, plant, and fungal "kingdoms" occupy relatively small branches on the eukaryotic tree of life.

Eukaryotic cells are surrounded by a single plasma membrane unlike bacteria, which usually have a double membrane. The most obvious feature that distinguishes eukaryotes from prokaryotes is the presence of a membrane-bound nucleus in eukaryotes. In fact, eukaryotes are defined by the presence of a nucleus (from the Greek: *eu*-, "true" and *karuon*, "nut" or "kernel.").

As mentioned earlier, eukaryotic cells are almost always larger than bacterial cells, commonly 1000-fold greater in volume. Because of their large size complex internal structures and mechanisms are required for rapid transport and communication both inside the cell and to and from the external medium. A mesh of protein fibers called the cytoskeleton extends throughout the cell contributing to cell shape and to the management of intracellular traffic.

Almost all eukaryotic cells contain additional internal membrane-bound compartments called organelles. The specific functions of organelles are often closely tied to their physical properties and structures. Nevertheless, a significant number of specific biochemical processes occur in the cytosol and the cytosol, like organelles, is highly organized.

The interior of a eukaryotic cell contains an intracellular membrane network. Independent organelles, including the nucleus, mitochondria, and chloroplasts, are embedded in this membrane system that pervades the entire cell. Materials flow within paths defined by membrane walls and tubules. The intracellular traffic of materials between compartments is rapid, highly selective, and closely regulated.

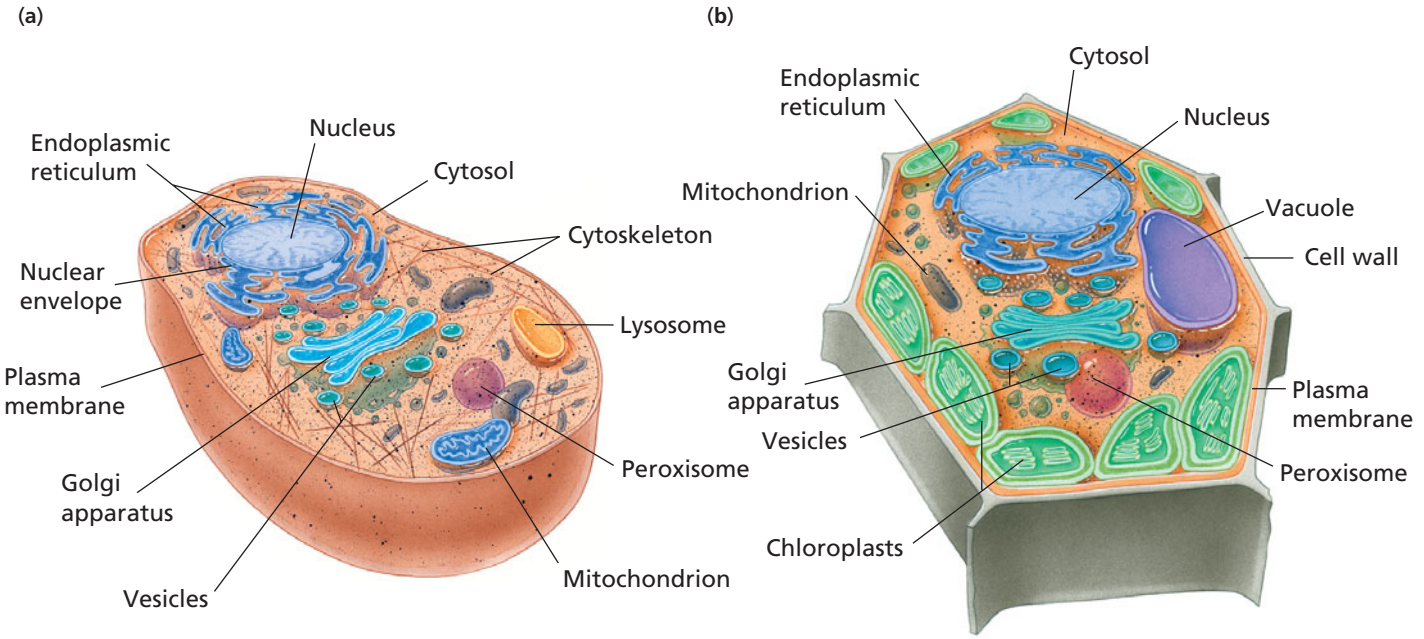
Figure 1.18 on the next page shows typical animal and plant cells. Both types have a nucleus, mitochondria, and a cytoskeleton. Plant cells also contain chloroplasts and vacuoles and are often surrounded by a rigid cell wall. Chloroplasts, also found in algae and some other protists, are the sites of photosynthesis. Plant cell walls are mostly composed of cellulose, one of the polysaccharides described in Section 1.3B.

Most multicellular eukaryotes contain tissues. Groups of similarly specialized cells within tissues are surrounded by an extracellular matrix containing proteins and polysaccharides. The matrix physically supports the tissue and in some cases directs cell growth and movement.

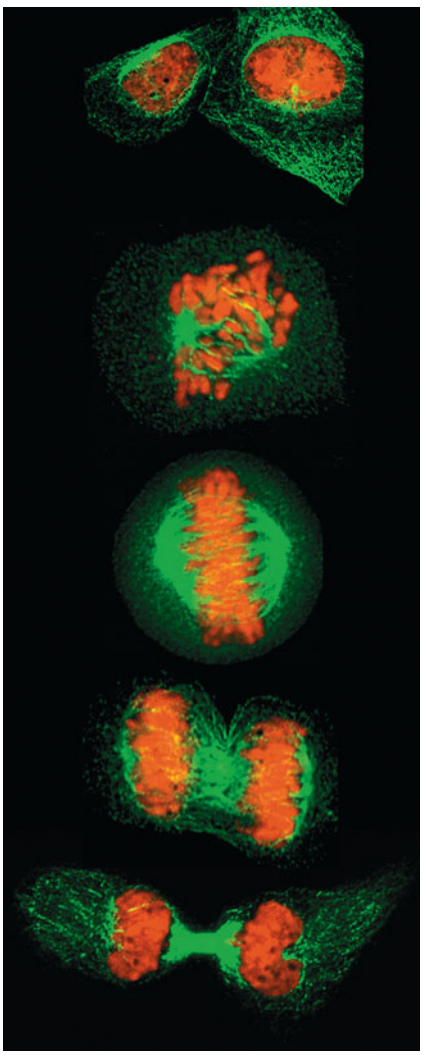
KEY CONCEPT

Animals are a relatively small, highly specialized, branch on the tree of life.

Downloaded From Vet-e-Books.com



▲ **Figure 1.18**
Eukaryotic cells. (a) Composite animal cell. Animal cells are typical eukaryotic cells containing organelles and structures also found in protists, fungi, and plants. **(b)** Composite plant cell. Most plant cells contain chloroplasts, the sites of photosynthesis in plants and algae; vacuoles, large, fluid-filled organelles containing solutes and cellular wastes; and rigid cell walls composed mostly of cellulose.



A. The Nucleus

The nucleus is usually the most obvious structure in a eukaryotic cell. It is structurally defined by the nuclear envelope, a membrane with two layers that join at protein-lined nuclear pores. The nuclear envelope is connected to the endoplasmic reticulum (see below). The nucleus is the control center of the cell containing 95% of its DNA, which is tightly packed with positively charged proteins called histones and coiled into a dense mass called chromatin. Replication of DNA and transcription of DNA into RNA occur in the nucleus. Many eukaryotes have a dense mass in the nucleus called the nucleolus. The nucleolus is a major site of RNA synthesis and the site of assembly of ribosomes.

Most eukaryotes contain far more DNA than do prokaryotes. Whereas the genetic material, or genome, of prokaryotes is usually a single circular molecule of DNA, the eukaryotic genome is organized as multiple linear chromosomes. In eukaryotes new DNA and histones are synthesized in preparation for cell division and the chromosomal material condenses and separates into two identical sets of chromosomes. This process is called mitosis (Figure 1.19). The cell is then pinched in two to complete cell division.

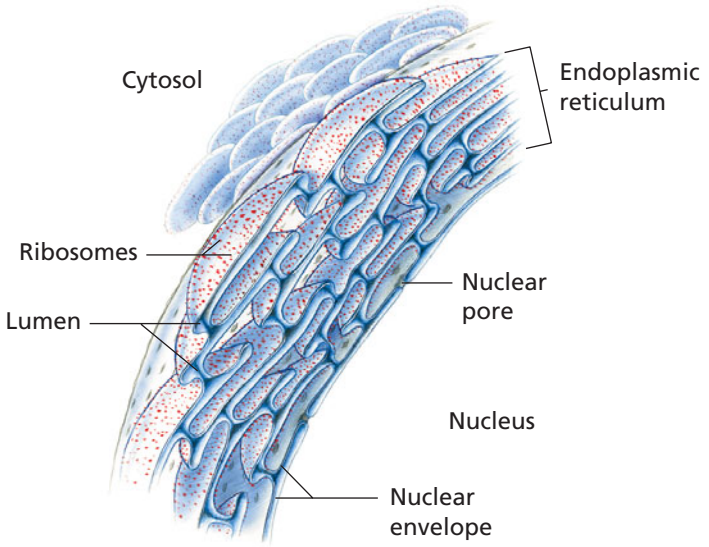
Most eukaryotes are diploid—they contain two complete sets of chromosomes. From time to time eukaryotic cells undergo meiosis resulting in the production of four haploid cells each with a single set of chromosomes. Two haploid cells—eggs and sperm, for example—can then fuse to regenerate a typical diploid cell. This process is one of the key features of sexual reproduction in eukaryotes.

B. The Endoplasmic Reticulum and Golgi Apparatus

A network of membrane sheets and tubules called the endoplasmic reticulum (ER) extends from the outer membrane of the nucleus. The aqueous region enclosed within the endoplasmic reticulum is called the lumen. In many cells part of the surface of the endoplasmic reticulum is coated with ribosomes that are actively synthesizing proteins.

◀ **Figure 1.19**
Mitosis. The five stages of mitosis are shown. Chromosomes (red) condense and line up in the center of the cell. Spindle fibers (green) are responsible for separating the recently duplicated chromosomes.

Downloaded From Vet-e-Books.com

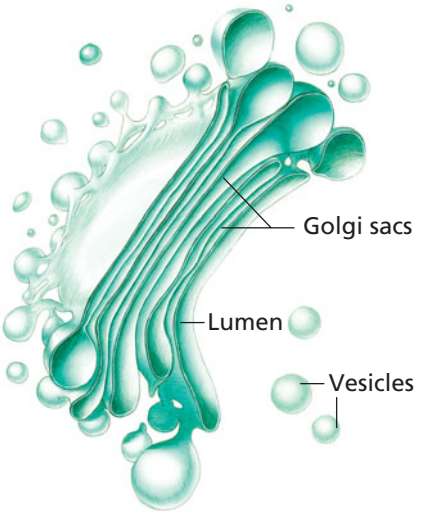


◀ **Nuclear envelope and endoplasmic reticulum (ER) of a eukaryotic cell.**

Protein synthesis, sorting, and secretion are described in Chapter 22.

As synthesis continues the protein is translocated through the membrane into the lumen. Proteins destined for export from the cell are completely extruded through the membrane into the lumen where they are packaged in membranous vesicles. These vesicles travel through the cell and fuse with the plasma membrane releasing their contents into the extracellular space. The synthesis of proteins destined to remain in the cytosol occurs at ribosomes that are not bound to the endoplasmic reticulum.

A complex of flattened, fluid-filled, membranous sacs called the Golgi apparatus is often found close to the endoplasmic reticulum and the nucleus. Vesicles that bud off from the endoplasmic reticulum fuse with the Golgi apparatus. The proteins carried by the vesicles may be chemically modified as they pass through the layers of the Golgi apparatus. The modified proteins are then sorted, packaged in new vesicles, and transported to specific destinations inside or outside the cell. The Golgi apparatus was discovered by Camillo Golgi in the 19th century (Nobel Laureate, 1906), although it wasn't until many decades later that its role in protein secretion was established.



▲ **Golgi apparatus.** The Golgi apparatus is responsible for the modification and sorting of proteins that have been transported to the Golgi apparatus by vesicles from the ER. Vesicles budding off the Golgi apparatus carry modified material to destinations inside and outside the cell.

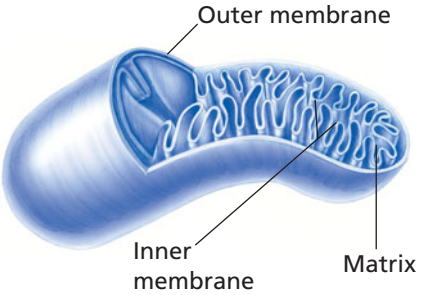
C. Mitochondria and Chloroplasts

Mitochondria and chloroplasts have central roles in energy transduction. Mitochondria are the main sites of oxidative energy metabolism. They are found in almost all eukaryotic cells. Chloroplasts are the sites of photosynthesis in plants and algae.

The mitochondrion has an inner and an outer membrane. The inner membrane is highly folded, resulting in a surface area three to five times that of the outer membrane. It is impermeable to ions and most metabolites. The aqueous phase enclosed by the inner membrane is called the mitochondrial matrix. Many of the enzymes involved in aerobic energy metabolism are found in the inner membrane and the matrix.

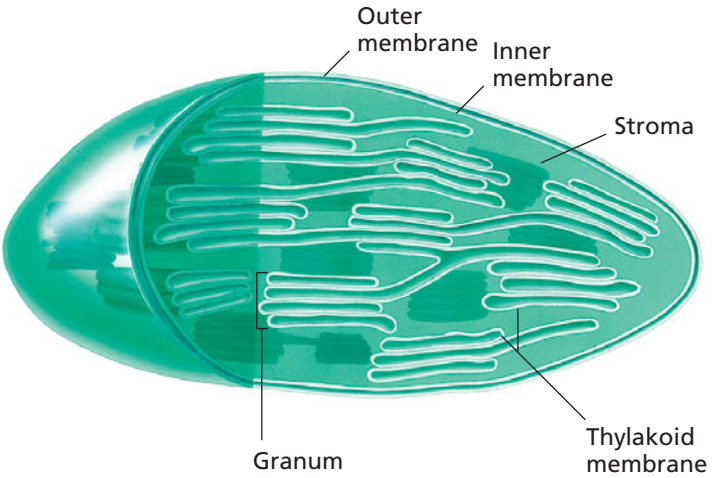
Mitochondria come in many sizes and shapes. The standard jellybean-shaped mitochondrion shown here is found in many cell types but some mitochondria are spherical or have irregular shapes.

The most important role of the mitochondrion is to oxidize organic acids, fatty acids, and amino acids to carbon dioxide and water. Much of the released energy is conserved in the form of a proton concentration gradient across the inner mitochondrial membrane. This stored energy is used to drive the conversion of adenosine diphosphate (ADP) and inorganic phosphate (P_i) to the energy-rich molecule ATP in a phosphorylation process that will be described in detail in Chapter 14. ATP is then used by the cell for such energy-requiring processes as biosynthesis, transport of certain molecules and ions against concentration and charge gradients, and generation of mechanical force for such purposes as locomotion and muscle contraction. The number of mitochondria found in cells varies widely. Some eukaryotic cells contain only a few mitochondria whereas others have thousands.



▲ **Mitochondrion.** Mitochondria are the main sites of energy transduction in aerobic eukaryotic cells. Carbohydrates, fatty acids, and amino acids are metabolized in this organelle.

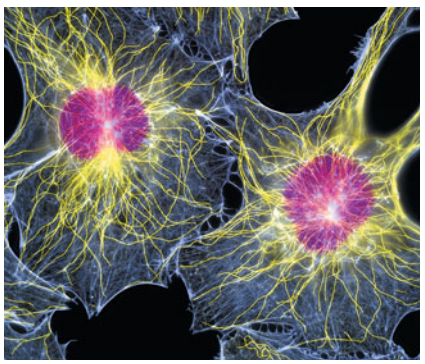
Downloaded From Vet-eBooks.com



► **Chloroplast.** Chloroplasts are the sites of photosynthesis in plants and algae. Light energy is captured by pigments associated with the thylakoid membrane and used to convert carbon dioxide and water to carbohydrates.

Photosynthetic plant cells contain chloroplasts as well as mitochondria. Like mitochondria, chloroplasts have an outer membrane and a complex, highly folded, inner membrane called the thylakoid membrane. Part of the inner membrane forms flattened sacs called grana (singular, granum). The thylakoid membrane, which is suspended in the aqueous stroma, contains chlorophyll and other pigments involved in the capture of light energy. Ribosomes and several circular DNA molecules are also suspended in the stroma. In chloroplasts the energy captured from light is used to drive the formation of carbohydrates from carbon dioxide and water.

Mitochondria and chloroplasts are derived from bacteria that entered into internal symbiotic relationships with primitive eukaryotic cells more than 1 billion years ago. Evidence for the endosymbiotic (*endo-*, "within") origin of mitochondria and chloroplasts includes the presence within these organelles of separate, small genomes and specific ribosomes that resemble those of bacteria. In recent years scientists have compared the sequences of mitochondrial and chloroplast genes (and proteins) with those of many species of bacteria. These studies in molecular evolution have shown that mitochondria are derived from primitive members of a particular group of bacteria called proteobacteria. Chloroplasts are descended from a distantly related class of photosynthetic bacteria called cyanobacteria.



▲ **Micrographs of fluorescently labeled actin filaments and microtubules in mammalian cells. (Left)** Actin filaments in rat muscle cells. **(Right)** Microtubules in human endothelial cells.

D. Specialized Vesicles

Eukaryotic cells contain specialized digestive vesicles called lysosomes. These vesicles are surrounded by a single membrane that encloses a highly acidic interior. The acidity is maintained by proton pumps embedded in the membrane. Lysosomes contain a variety of enzymes that catalyze the breakdown of cellular macromolecules such as proteins and nucleic acids. They can also digest large particles such as retired mitochondria and bacteria ingested by the cell. Lysosomal enzymes are much less active at the near-neutral pH of the cytosol than they are under the acidic conditions inside the lysosome. The compartmentalization of lysosomal enzymes keeps them from accidentally catalyzing the degradation of macromolecules in the cytosol.

Peroxisomes are present in all animal cells and many plant cells. Like lysosomes, they are surrounded by a single membrane. Peroxisomes carry out oxidation reactions, some of which produce the toxic compound hydrogen peroxide, (H_2O_2). Some hydrogen peroxide is used for the oxidation of other compounds. Excess hydrogen peroxide is destroyed by the action of the peroxisomal enzyme catalase, which catalyzes the conversion of hydrogen peroxide to water and oxygen.

Vacuoles are fluid-filled vesicles surrounded by a single membrane. They are common in mature plant cells and some protists. These vesicles are storage sites for water, ions, and nutrients such as glucose. Some vacuoles contain metabolic waste products and some contain enzymes that can catalyze the degradation of macromolecules no longer needed by the plant.

Downloaded From Vet-eBooks.com

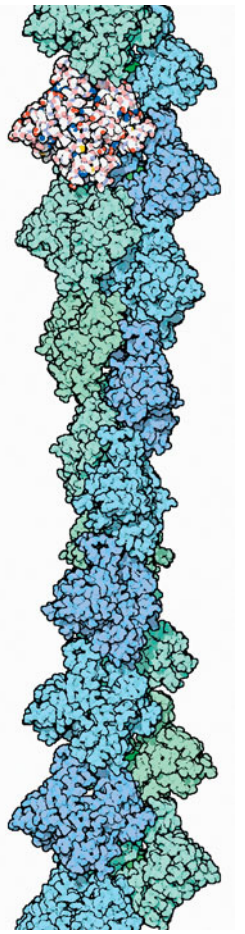
E. The Cytoskeleton

The cytoskeleton is a protein scaffold required for support, internal organization, and even movement of the cell. Some types of animal cells contain a dense cytoskeleton but it is much less prominent in most other eukaryotic cells. The cytoskeleton consists of three types of protein filaments: actin filaments, microtubules, and intermediate filaments. All three types are built of individual protein molecules that combine to form threadlike fibers.

Actin filaments (also called microfilaments) are the most abundant cytoskeletal component. They are composed of a protein called actin that forms ropelike threads with a diameter of about 7 nm. Actin has been found in all eukaryotic cells and is frequently the most abundant protein in the cell. It is also one of the most evolutionarily conserved proteins. This is evidence that actin filaments were present in the ancestral eukaryotic cell from which all modern eukaryotes are descended.

Microtubules are strong, rigid fibers frequently packed in bundles. They have a diameter of about 22 nm—much thicker than actin filaments. Microtubules are composed of a protein called tubulin. Microtubules serve as a kind of internal skeleton in the cytoplasm, but they also form the mitotic spindle during mitosis. In addition, microtubules can form structures capable of directed movement, such as cilia. The flagella that propel sperm cells are an example of very long cilia—they are not related to bacterial flagella. The waving motion of cilia is driven by energy from ATP.

Intermediate filaments are found in the cytoplasm of most eukaryotic cells. These filaments have diameters of approximately 10 nm, which makes them intermediate in size compared to actin filaments and microtubules. Intermediate filaments line the inside of the nuclear envelope and extend outward from the nucleus to the periphery of the cell. They help the cell resist external mechanical stresses.



▲ **Actin.** Actin filament showing the organization in individual subunits of the protein actin. (Courtesy David S. Goodsell)

1.9 A Picture of the Living Cell

We have now introduced the major structures found within cells and described their roles. These structures are immense compared to the molecules and polymers that will be our focus for the rest of this book. Cells contain thousands of different metabolites and many millions of molecules. In the cytosol of every cell there are hundreds of different enzymes, each acting specifically on only one or possibly a few related metabolites. There may be 100,000 copies of some enzymes per cell but only a few copies of other enzymes. Each enzyme is bombarded with potential substrates.

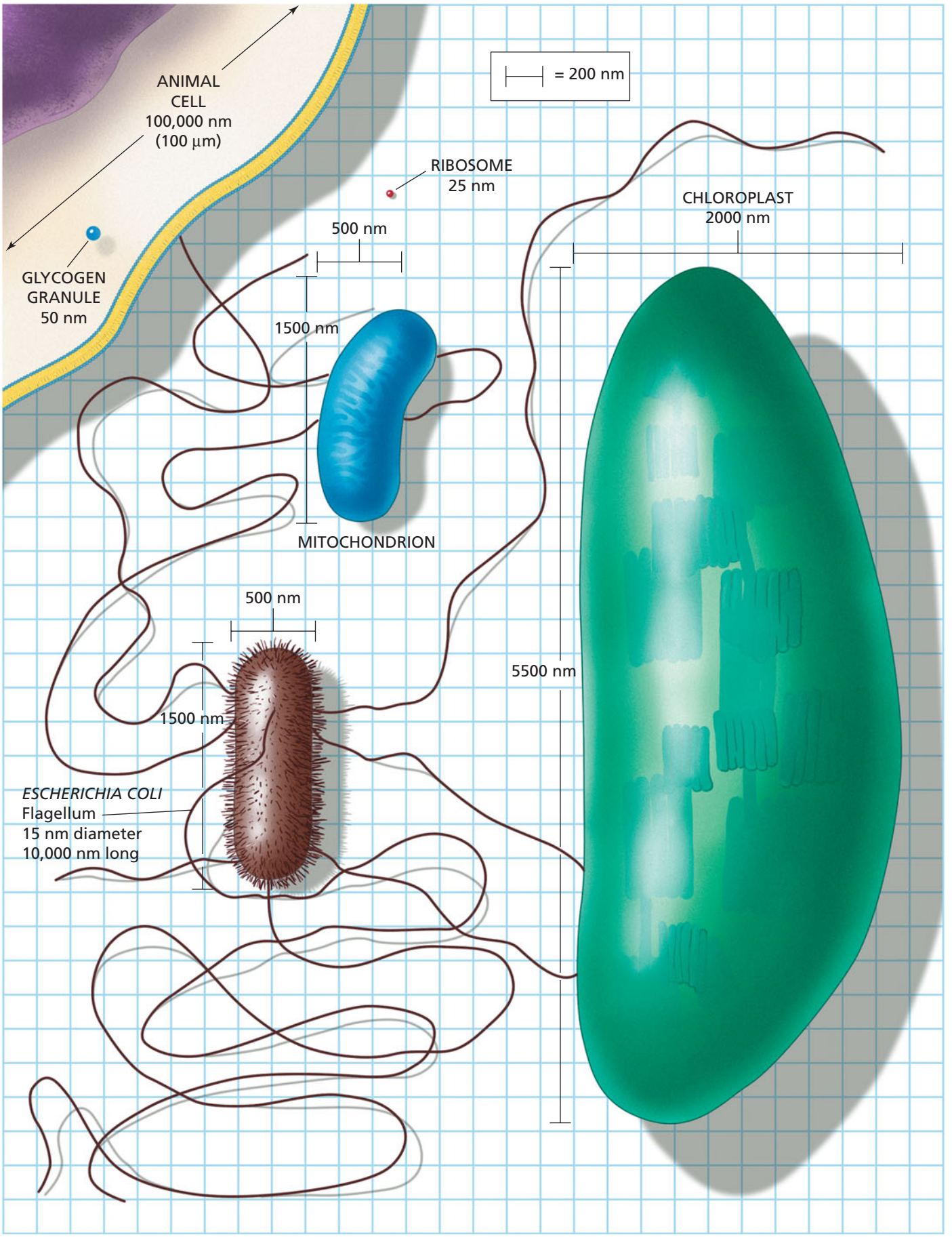
Molecular biologist and artist David S. Goodsell has produced captivating images showing the molecular contents of an *E. coli* cell magnified 1 million times (Figure 1.20 on page 26). Approximately 600 cubes of this size represent the volume of the *E. coli* cell. At this scale individual atoms are smaller than the dot in the letter *i* and small metabolites are barely visible. Proteins are the size of a grain of rice.

A drawing of the molecules in a cell shows how densely packed the cytoplasm can be, but it cannot give a sense of activity at the atomic scale. All the molecules in a cell are moving and colliding with each other. The collisions between molecules are fully elastic—the energy of a collision is conserved in the energy of the rebound. As molecules bounce off each other they travel a wildly crooked path in space, called the random walk of diffusion. For a small molecule such as water, the mean distance traveled between collisions is less than the dimensions of the molecule and the path includes many reversals of direction. Despite its convoluted path, a water molecule can diffuse the length of an *E. coli* cell in 1/10 second.

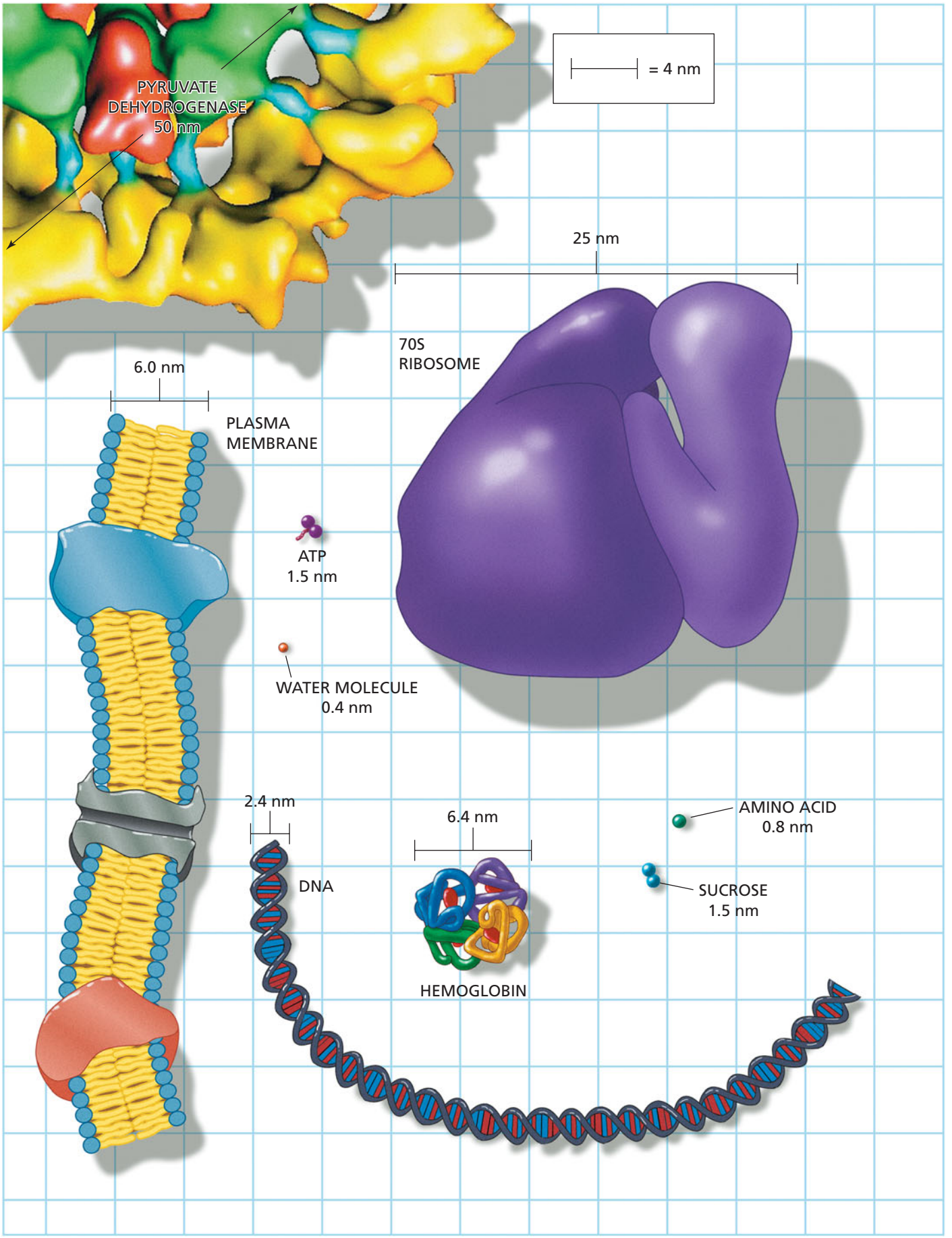
An enzyme and a small molecule will collide 1 million times per second. Under these conditions, a rate of catalysis typical of many enzymes could be achieved even if only 1 in about 1000 collisions results in a reaction. Nevertheless, some enzymes catalyze reactions with an efficiency far greater than 1 reaction per 1000 collisions. In fact, a few enzymes catalyze reactions with almost every molecule of substrate their active sites encounter—an example of the astounding potency of enzyme-directed chemistry. The study of the reaction rates of enzymes, or enzyme kinetics, is one of the most fundamental aspects of biochemistry. It will be covered in Chapter 6.

Lipids in membranes also diffuse vigorously, though only within the two-dimensional plane of the lipid bilayer. Lipid molecules exchange places with neighboring

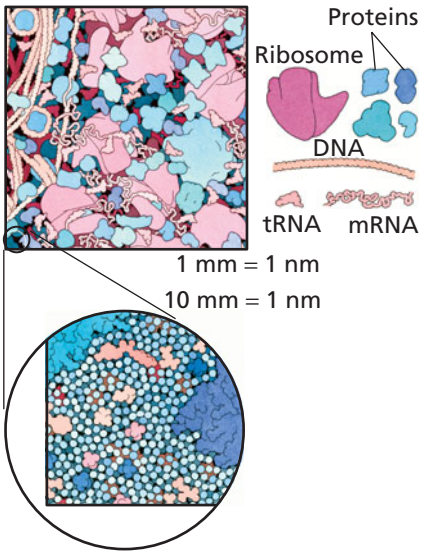
Downloaded From Vet-eBooks.com



Downloaded From Vet-e-Books.com



Downloaded From Vet-e-Books.com



▲ **Figure 1.20**
Portion of the cytosol of an *E. coli* cell. The top illustration, in which the contents are magnified 1 million times, represents a window 100 x 100 nm. Proteins are in shades of blue and green. Nucleic acids are in shades of pink. The large structures are ribosomes. Water and small metabolites are not shown. The contents in the round inset are magnified 10 million times, showing water and other small molecules.

molecules in membranes about 6 million times per second. Some membrane proteins can also diffuse rapidly within the membrane.

Large molecules diffuse more slowly than small ones. In eukaryotic cells the diffusion of large molecules such as enzymes is retarded even further by the complex network of the cytoskeleton. Large molecules diffuse across a given distance as much as 10 times more slowly in the cytosol than in pure water.

The full extent of cytosolic organization is not yet known. A number of proteins and enzymes form large complexes that carry out a series of reactions. We will encounter several such complexes in our study of metabolism. They are often referred to as protein machines. This arrangement has the advantage that metabolites pass directly from one enzyme to the next without diffusing away into the cytosol. Many researchers are sympathetic to the idea that the cytosol is not merely a random mixture of soluble molecules but is highly organized in contrast to the long-held impression that simple solution chemistry governs cytosolic activity. The concept of a highly organized cytosol is a relatively new idea in biochemistry. It may lead to important new insights about how cells work at the molecular level.

1.10 Biochemistry Is Multidisciplinary

One of the goals of biochemists is to integrate a large body of knowledge into a molecular explanation of life. This has been, and continues to be, a challenging task but, in spite of the challenges, biochemists have made a great deal of progress toward defining and understanding the basic reactions common to all cells.

The discipline of biochemistry does not exist in a vacuum. We have already seen how physics, chemistry, cell biology, and evolution contribute to an understanding of biochemistry. Related disciplines, such as physiology and genetics, are also important. In fact, many scientists no longer consider themselves to be just biochemists but are also knowledgeable in several related fields.

Because all aspects of biochemistry are interrelated it is difficult to present one topic without referring to others. For example, function is intimately related to structure and the regulation of individual enzyme activities can be appreciated only in the context of a series of linked reactions. The interrelationship of biochemistry topics is a problem for both students and teachers in an introductory biochemistry course. The material must be presented in a logical and sequential manner but there is no universal sequence of topics that suits every course, or every student. Fortunately, there is general agreement on the broad outline of an approach to understanding the basic principles of biochemistry and this textbook follows that outline. We begin with an introductory chapter on water. We will then describe the structures and functions of proteins and enzymes, carbohydrates, and lipids. The third part of the book makes use of structural information to describe metabolism and its regulation. Finally, we will examine nucleic acids and the storage and transmission of biological information.

Some courses may cover the material in a slightly different order. For example, the structures of nucleic acids can be described before the metabolism section. Wherever possible, we have tried to write chapters so that they can be covered in different orders in a course depending on the particular needs and interests of the students.

Appendix The Special Terminology of Biochemistry

Most biochemical quantities are specified using Système International (SI) units. Some common SI units are listed in Table 1.1 Many biochemists still use more traditional units, although these are rapidly disappearing from the scientific literature. For example, protein chemists sometimes use the angstrom (Å) to report interatomic distances; 1 Å is equal to 0.1 nm, the preferred SI unit. Calories (cal) are sometimes used instead of joules (J); 1 cal is equal to 4.184 J.

The standard SI unit of temperature is the Kelvin, but temperature is most commonly reported in degrees Celsius (°C). One degree Celsius is equal in magnitude to 1 Kelvin, but the Celsius scale begins at the freezing point of water (0°C) and 100°C is

Downloaded From Vet-e-Books.com

TABLE 1.1 SI units commonly used in biochemistry

Physical quantity	SI unit	Symbol
Length	meter	m
Mass	gram	g
Amount	mole	mol
Volume	liter[a]	L
Energy	joule	J
Electric potential	volt	V
Time	second	s
Temperature	Kelvin[b]	K

[a]1 liter = 1000 cubic centimeters.
[b]273 K = 0° C.

Table 1.2 Prefixes commonly used with SI units

Prefix	Symbol	Multiplication factor
giga-	G	10^9
mega-	M	10^6
kilo-	k	10^3
deci-	d	10^{-1}
centi-	c	10^{-2}
milli-	m	10^{-3}
micro-	μ	10^{-6}
nano-	n	10^{-9}
pico-	p	10^{-12}
femto-	f	10^{-15}

the boiling point of water at 1 atm. This scale is often referred to as the centigrade scale (*centi-* = 1/100). Absolute zero is −273 °C, which is equal to 0 K. In warm-blooded mammals biochemical reactions occur at body temperature (37°C in humans).

Very large or very small numerical values for some SI units can be indicated by an appropriate prefix. The commonly used prefixes and their symbols are listed in Table 1.2. In addition to the standard SI units employed in all fields, biochemistry has its own special terminology; for example, biochemists use convenient abbreviations for biochemicals that have long names.

The terms RNA and DNA are good examples. They are shorthand versions of the long names ribonucleic acid and deoxyribonucleic acid. Abbreviations such as these are very convenient, and learning to associate them with their corresponding chemical structures is a necessary step in mastering biochemistry. In this book, we will describe common abbreviations as each new class of compounds is introduced.

Selected Readings

Chemistry

Bruice, P. Y. (2011). *Organic Chemistry*, 6th ed. (Upper Saddle River, NJ: Prentice Hall).

Tinoco, I., Sauer, K., Wang, J. C., and Puglisi, J. D. (2002). *Physical Chemistry: Principles and Applications in Biological Sciences*, 4th ed. (Upper Saddle River, NJ: Prentice Hall).

van Holde, K. E., Johnson, W. C., and Ho, P.S. (2005). *Principles of Physical Biochemistry* 2nd ed. (Upper Saddle River, NJ: Prentice Hall).

Cells

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2004). *Essential Cell Biology* (New York: Garland).

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Kreiger, M., Scott, M. P., Zipursky, L., and Darnell, J. (2003). *Molecular Cell Biology*, 5th ed. (New York: Scientific American Books).

Goodsell, D. S. (1993). *The Machinery of Life* (New York: Springer-Verlag).

Evolution and the Diversity of Life

Doolittle, W. F. (2000). Uprooting the tree of life. *Sci. Am.* 282(2):90–95.

Doolittle, W. F. (2009). Eradicating topological thinking in prokaryotic systematics and evolution. *Cold Spr. Hbr. Symp. Quant. Biol.*

Margulis, L., and Schwartz, K.V. (1998). *Five Kingdoms*, 3rd ed. (New York: W.H. Freeman).

Graur, D., and Li, W.-H. (2000). *Fundamentals of Molecular Evolution* (Sunderland, MA: Sinauer).

Sapp, J. (Ed.) (2005). *Microbial Phylogeny and Evolution: Concepts and Controversies*. (Oxford, UK: Oxford University Press).

Sapp, J. (2009) *The New Foundations of Evolution*. (Oxford, UK: Oxford University Press).

History of Science

Kohler, R. E. (1975). The History of Biochemistry, a Survey. *J. Hist. Biol* 8:275–318.

Downloaded From Vet-e-Books.com





Water

Life on Earth is often described as a carbon-based phenomenon but it would be equally correct to refer to it as a water-based phenomenon. Life probably originated in water more than three billion years ago and all living cells still depend on water for their existence. Water is the most abundant molecule in most cells accounting for 60% to 90% of the mass of the cell. The exceptions are cells from which water is expelled such as those in seeds and spores. Seeds and spores can lie dormant for long periods of time until they are revived by the reintroduction of water.

Life spread from the oceans to the continents about 500 million years ago. This major transition in the history of life required special adaptations to enable terrestrial life to survive in an environment where water was less plentiful. You will encounter many of these adaptations in the rest of this book.

An understanding of water and its properties is important to the study of biochemistry. The macromolecular components of cells—proteins, polysaccharides, nucleic acids, and lipids—assume their characteristic shapes in response to water. For example, some types of molecules interact extensively with water and, as a result, are very soluble while other molecules do not dissolve easily in water and tend to associate with each other in order to avoid water. Much of the metabolic machinery of cells has to operate in an aqueous environment because water is an essential solvent.

We begin our detailed study of the chemistry of life by examining the properties of water. The physical properties of water allow it to act as a solvent for ionic and other polar substances, and the chemical properties of water allow it to form weak bonds with other compounds, including other water molecules. The chemical properties of water are also related to the functions of macromolecules, entire cells, and organisms. These interactions are important sources of structural stability in macromolecules and large cellular structures. We will see how water affects the interactions of substances that have low solubility in water. We will examine the ionization of water and discuss acid–base chemistry—topics that are the foundation for understanding the molecules and processes that we will encounter in subsequent chapters. It's important to keep in mind that water is not just an inert solvent; it is also a substrate for many cellular reactions.

There is nothing softer and weaker than water, And yet there is nothing better for attacking hard and strong things. For this reason there is no substitute for it.

—Lao-Tzu (c. 550 BCE)



▲ **Eureka Dunes evening primrose (*Oenothera californica*)** This species only grows in the sand dunes of Death Valley National Park in California. It has evolved special mechanisms for conserving water.

Top: Earth from space. The earth is a watery planet and water plays a central role in the chemistry of all life.

Downloaded From Vet-e-Books.com

2.1 The Water Molecule Is Polar

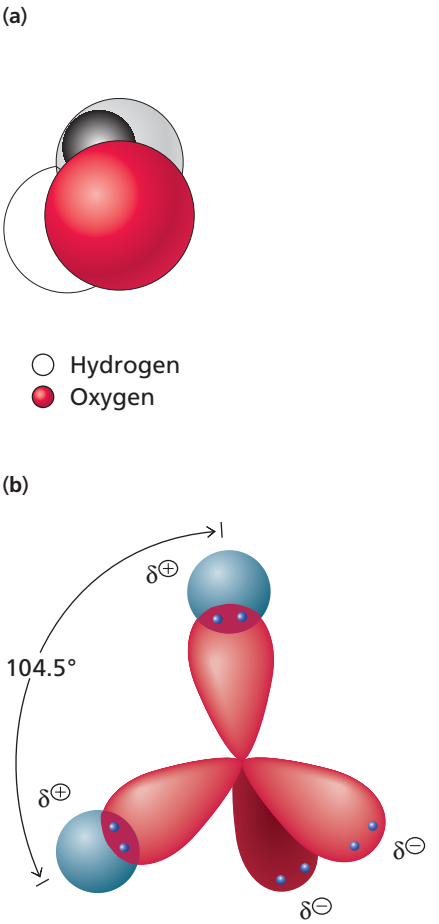
A water molecule (H_2O) is V-shaped (Figure 2.1a) and the angle between the two covalent (O—H) bonds is 104.5°. Some important properties of water arise from its angled shape and the intermolecular bonds that it can form. An oxygen atom has eight electrons and its nucleus has eight protons and eight neutrons. There are two electrons in the inner shell and six electrons in the outer shell. The outer shell can potentially accommodate four pairs of electrons in one *s* orbital and three *p* orbitals. However, the structure of water and its properties can be better explained by assuming that the electrons in the outer shell occupy four sp^3 hybrid orbitals. Think of these four orbitals as occupying the four corners of a tetrahedron that surrounds the central atom of oxygen. Two of the sp^3 hybrid orbitals contain a pair of electrons and the other two each contain a single electron. This means that oxygen can form covalent bonds with other atoms by sharing electrons to fill these single electron orbitals. In water the covalent bonds involve two different hydrogen atoms each of which shares its single electron with the oxygen atom. In Figure 2.1b each electron is indicated by a blue dot showing that each sp^3 hybrid orbital of the oxygen atom is occupied by two electrons including those shared with the hydrogen atoms. The inner shell of the hydrogen atom is also filled because of these two shared electrons in the covalent bond.

The H—O—H bond angle in free water molecules is 104.5° but if the electron orbitals were really pointing to the four corners of a tetrahedron, the angle would be 109.5°. The usual explanation for this difference is that there is strong repulsion between the lone electron pairs and this repulsion pushes the covalent bond orbitals closer together, reducing the angle from 109.5° to 104.5°.

Oxygen atoms are more electronegative than hydrogen atoms because an oxygen nucleus attracts electrons more strongly than the single proton in the hydrogen nucleus. As a result, an uneven distribution of charge occurs within each O—H bond of the water molecule with oxygen bearing a partial negative charge ($\delta^{\ominus}$) and hydrogen bearing a partial positive charge ($\delta^{\oplus}$). This uneven distribution of charge within a bond is known as a dipole and the bond is said to be polar.

The polarity of a molecule depends both on the polarity of its covalent bonds and its geometry. The angled arrangement of the polar O—H bonds of water creates a permanent dipole for the molecule as a whole as shown in Figure 2.2a. A molecule of ammonia also contains a permanent dipole (Figure 2.2b) Thus, even though water and gaseous ammonia are electrically neutral, both molecules are polar. The high solubility of the polar ammonia molecules in water is facilitated by strong interactions with the polar water molecules. The solubility of ammonia in water demonstrates the principle that "like dissolves like."

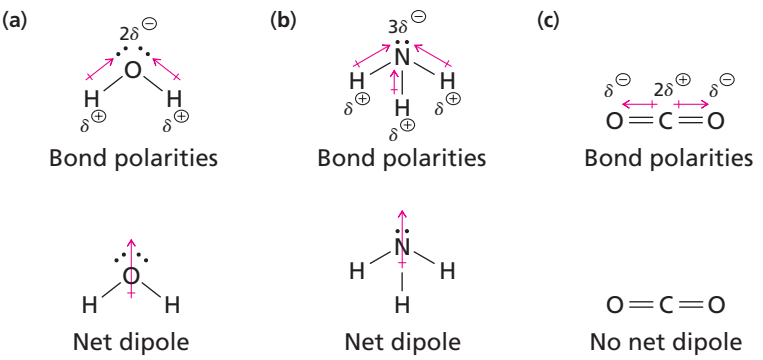
Not all molecules are polar; for example, carbon dioxide also contains polar covalent bonds but the bonds are aligned with each other and oppositely oriented so the polarities cancel each other (Figure 2.2c). As a result, carbon dioxide has no net dipole and is much less soluble in water than ammonia.



▲ **Figure 2.1 A water molecule.** **(a)** Space-filling structure of a water molecule. **(b)** Angle between the covalent bonds of a water molecule. Two of the sp^3 hybrid orbitals of the oxygen atom participate in covalent bonds with *s* orbitals of hydrogen atoms. The other two sp^3 orbitals are occupied by lone pairs of electrons.

KEY CONCEPT

Polar molecules are molecules with an unequal distribution of charge so that one end of the molecules is more negative and another end is more positive.



◄ **Figure 2.2**
Polarity of small molecules. (a) The geometry of the polar covalent bonds of water creates a permanent dipole for the molecule with the oxygen bearing a partial negative charge (symbolized by $2\delta^{\ominus}$) and each hydrogen bearing a partial positive charge (symbolized by $\delta^{\oplus}$). **(b)** The pyramidal shape of a molecule of ammonia also creates a permanent dipole. **(c)** The polarities of the collinear bonds in carbon dioxide cancel each other. Therefore, CO_2 is not polar. (Arrows depicting dipoles point toward the negative charge with a cross at the positive end.)

Downloaded From Vet-eBooks.com

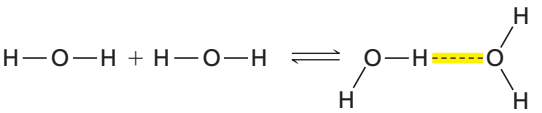
2.2 Hydrogen Bonding in Water

One of the important consequences of the polarity of the water molecule is that water molecules attract one another. The attraction between one of the slightly positive hydrogen atoms of one water molecule and the slightly negative electron pairs in one of the sp^3 hybrid orbitals produces a hydrogen bond (Figure 2.3). In a **hydrogen bond** between two water molecules the hydrogen atom remains covalently bonded to its oxygen atom, the hydrogen donor. At the same time, it is attracted to another oxygen atom, called the hydrogen acceptor. In effect, the hydrogen atom is being shared (unequally) between the two oxygen atoms. The distance from the hydrogen atom to the acceptor oxygen atom is about twice the length of the covalent bond.

KEY CONCEPT

Hydrogen bonds form when a hydrogen atom with a partially positive charge ($\delta^{\oplus}$) is shared between two electronegative atoms ($2\delta^{\ominus}$). Hydrogen bonds are much weaker than covalent bonds.

Water is not the only molecule capable of forming hydrogen bonds; these interactions can occur between any electronegative atom and a hydrogen atom attached to another electronegative atom. (We will examine other examples of hydrogen bonding in Section 2.5B.) Hydrogen bonds are much weaker than typical covalent bonds. The strength of hydrogen bonds in water and in solutions is difficult to measure directly but it is estimated to be about 20 kJ mol^{-1}.



$$\Delta H_f = -20 \text{ kJ mol}^{-1} \qquad (2.1)$$

About 20 kJ mol^{-1} of heat is given off when hydrogen-bonded water molecules form in water under standard conditions. (Recall that standard conditions are 1 atm pressure and a temperature of 25°C.) This value is the standard enthalpy of formation (ΔH_f). It means that the change in enthalpy when hydrogen bonds form is about –20 kJ per mole of water. This is equivalent to saying that +20 kJ mol^{-1} of heat energy is required to disrupt hydrogen bonds between water molecules—the reverse of the reaction shown in Reaction 2.1. This value depends on the type of hydrogen bond. In contrast, the energy required to break a covalent O—H bond in water is about 460 kJ mol^{-1}, and the energy required to break a covalent C—H bond is about 410 kJ mol^{-1}. Thus, the strength of hydrogen bonds is less than 5% of the strength of typical covalent bonds. Hydrogen bonds are weak interactions compared to covalent bonds.

Orientation is important in hydrogen bonding. A hydrogen bond is most stable when the hydrogen atom and the two electronegative atoms associated with it (the two oxygen atoms, in the case of water) are aligned, or nearly in line, as shown in Figure 2.3. Water molecules are unusual because they can form four O—H—O aligned hydrogen bonds with up to four other water molecules (Figure 2.4). They can donate each of their two hydrogen atoms to two other water molecules and accept two hydrogen atoms from two other water molecules. Each hydrogen atom can participate in only one hydrogen bond.

The three-dimensional interactions of liquid water are difficult to study but much has been learned by examining the structure of ice crystals (Figure 2.5). In the common form of ice, every molecule of water participates in four hydrogen bonds, as expected. Each of the hydrogen bonds points to the oxygen atom of an adjacent water molecule and these four adjacent hydrogen-bonded oxygen atoms occupy the vertices of a tetrahedron. This arrangement is consistent with the structure of water shown in Figure 2.1

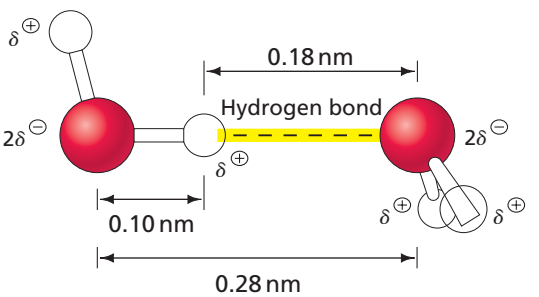


Figure 2.3 ►
Hydrogen bonding between two water molecules. A partially positive ($\delta^{\oplus}$) hydrogen atom of one water molecule attracts the partially negative ($2\delta^{\ominus}$) oxygen atom of a second water molecule, forming a hydrogen bond. The distances between atoms of two water molecules in ice are shown. Hydrogen bonds are indicated by dashed lines highlighted in yellow, as shown here and throughout the book.

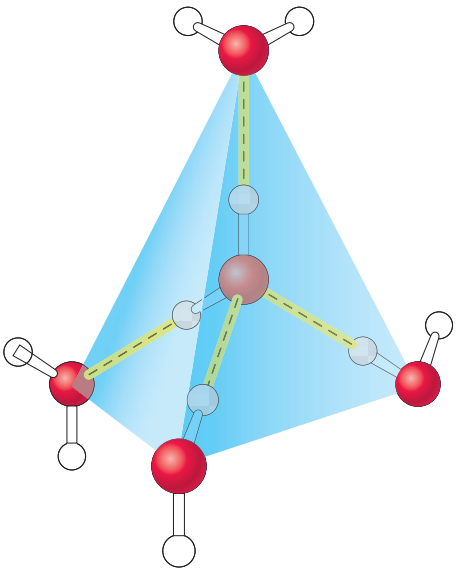
Downloaded From Vet-e-Books.com

except that the bond angles are all equal (109.5°). This is because the polarity of individual water molecules, which distorts the bond angles, is canceled by the presence of hydrogen bonds. The average energy required to break each hydrogen bond in ice has been estimated to be 23 kJ mol^{-1}, making those bonds a bit stronger than those formed in water.

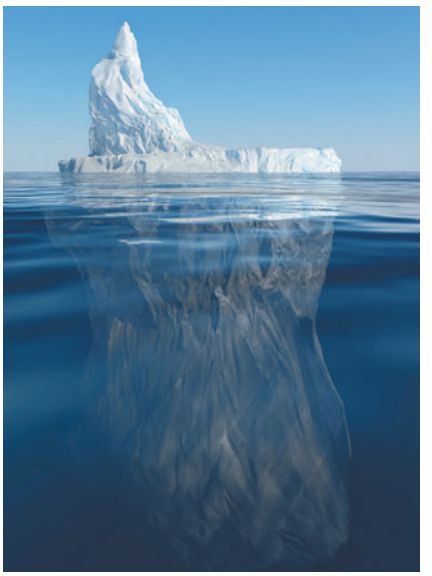
The ability of water molecules in ice to form four hydrogen bonds and the strength of these hydrogen bonds give ice an unusually high melting point because a large amount of energy, in the form of heat, is required to disrupt the hydrogen-bonded lattice of ice. When ice melts most of the hydrogen bonds are retained by liquid water. Each molecule of liquid water can form up to four hydrogen bonds with its neighbors but most participate in only two or three at any given moment. This means that the structure of liquid water is less ordered than that of ice. The fluidity of liquid water is primarily a consequence of the constantly fluctuating pattern of hydrogen bonding as hydrogen bonds break and re-form. At any given time there will be many water molecules participating in two, three, or four hydrogen bonds with other water molecules. There will also be many that participate in only one hydrogen bond or none at all. This is a dynamic structure—the average hydrogen bond lifetime in water is only 10 picoseconds (10^{-11} s).

The density of most substances increases upon freezing as molecular motion slows and tightly packed crystals form. The density of water also increases as it cools—until it reaches a maximum of 1.000 g ml^{-1} at 4°C (277 K). (This value is not a coincidence. Grams are *defined* as the weight of 1 milliliter of water at 4°C.) Water expands as the temperature drops below 4°C. This expansion is caused by the formation of the more open hydrogen-bonded ice crystal in which each water molecule is hydrogen-bonded rigidly to four others. As a result ice is slightly less dense (0.924 g ml^{-1}) than liquid water whose molecules can move enough to pack more closely. Because ice is less dense than liquid water it floats and water freezes from the top down. This has important biological implications since a layer of ice on a pond insulates the creatures below from extreme cold.

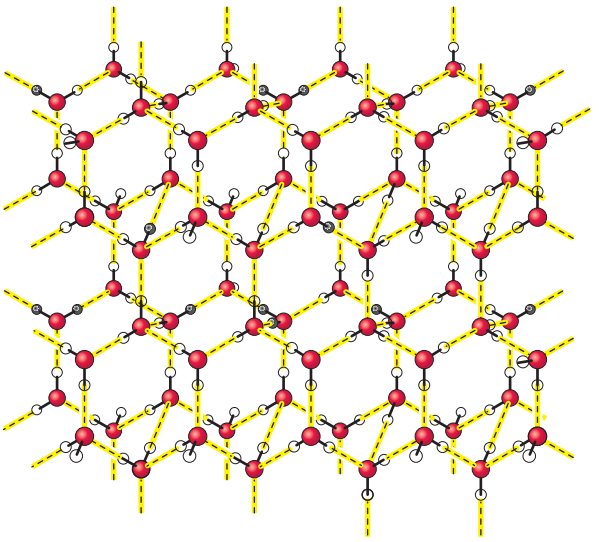
Two additional properties of water are related to its hydrogen-bonding characteristics—its specific heat and its heat of vaporization. The specific heat of a substance is the amount of heat needed to raise the temperature of 1 gram of the substance by 1°C. This property is also called the heat capacity. In the case of water, a relatively large amount of heat is required to raise the temperature because each water molecule participates in multiple hydrogen bonds that must be broken in order for the kinetic energy of the water molecules to increase. The abundance of water in the cells and tissues of all large multicellular organisms means that temperature fluctuations within cells are minimized.



▲ **Figure 2.4**
Hydrogen bonding by a water molecule. A water molecule can form up to four hydrogen bonds: the oxygen atom of a water molecule is the hydrogen acceptor for two hydrogen atoms, and each O—H group serves as a hydrogen donor.



▲ **Icebergs.** Ice floats because it is less dense than water. However, it is only slightly less dense than water so most of the mass of floating ice lies underwater.



◀ **Figure 2.5**
Structure of ice. Water molecules in ice form an open hexagonal lattice in which every water molecule is hydrogen-bonded to four others. The geometrical regularity of these hydrogen bonds contributes to the strength of the ice crystal. The hydrogen-bonding pattern of ice is more regular than that of water. The absolute structure of liquid water has not been determined.

Downloaded From Vet-e-Books.com

BOX 2.1 EXTREME THERMOPHILES

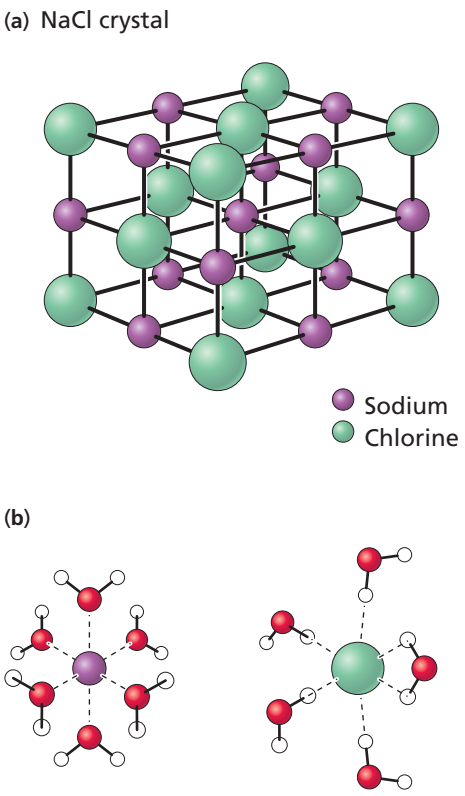
Some species can grow and reproduce at temperatures very close to 0°C, or even lower. There are cold-blooded fish, for example, that survive at ocean temperatures below 0°C (salt lowers the freezing point of water).

At the other extreme are bacteria that live in hot springs where the average temperature is above 80°C. Some bacteria inhabit the environment around deep ocean thermal vents (black smokers) where the average temperature is more than 100°C. (The high pressure at the bottom of the ocean raises the boiling point of water.)

The record for extreme thermophiles is Strain 121, a species of archaebacteria that grows and reproduces at 121°C! These extreme thermophiles are among the earliest branching lineages on the web of life. It's possible that the first living cells arose near deep ocean vents.

Deep ocean hydrothermal vent. ►





▲ **Figure 2.6**
Dissolution of sodium chloride (NaCl) in water. **(a)** The ions of crystalline sodium chloride are held together by electrostatic forces. **(b)** Water weakens the interactions between the positive and negative ions and the crystal dissolves. Each dissolved $Na^{\oplus}$ and $Cl^{\ominus}$ is surrounded by a solvation sphere. Only one layer of solvent molecules is shown. Interactions between ions and water molecules are indicated by dashed lines.

This feature is of critical biological importance since the rates of most biochemical reactions are sensitive to temperature.

The heat of vaporization of water (~2260 J g^{-1}) is also much higher than that of many other liquids. A large amount of heat is required to convert water from a liquid to a gas because hydrogen bonds must be broken to permit water molecules to dissociate from one another and enter the gas phase. Because the evaporation of water absorbs so much heat, perspiration is an effective mechanism for decreasing body temperature.

2.3 Water Is an Excellent Solvent

The physical properties of water combine to make it an excellent solvent. We have already seen that water molecules are polar and this property has important consequences, as we will see below. In addition, water has a low intrinsic viscosity that does not greatly impede the movement of dissolved molecules. Finally, water molecules themselves are small compared to some other solvents such as ethanol and benzene. The small size of water molecules means that many of them can associate with solute particles to make them more soluble.

A. Ionic and Polar Substances Dissolve in Water

Water can interact with and dissolve other polar compounds and compounds that ionize. Ionization is associated with the gain or loss of an electron, or an H^+ ion, giving rise to an atom or a molecule that carries a net charge. Molecules that can dissociate to form ions are called **electrolytes**. Substances that readily dissolve in water are said to be **hydrophilic**, or water loving. (We will discuss hydrophobic, or water fearing, substances in the next section.)

Why are electrolytes soluble in water? Recall that water molecules are polar. This means they can align themselves around electrolytes so that the negative oxygen atoms of the water molecules are oriented toward the cations (positively charged ions) of the electrolytes and the positive hydrogen atoms are oriented toward the anions (negatively charged ions). Consider what happens when a crystal of sodium chloride (NaCl) dissolves in water (Figure 2.6) The polar water molecules are attracted to the charged ions in the crystal. The attractions result in sodium and chloride ions on the surface of the

Downloaded From Vet-eBooks.com

crystal dissociating from one another and the crystal begins to dissolve. Because there are many polar water molecules surrounding each dissolved sodium and chloride ion, the interactions between the opposite electric charges of these ions become much weaker than they are in the intact crystal. As a result of its interactions with water molecules, the ions of the crystal continue to dissociate until the solution becomes saturated. At this point, the ions of the dissolved electrolyte are present at high enough concentrations for them to again attach to the solid electrolyte, or crystallize, and an equilibrium is established between dissociation and crystallization.

BOX 2.2 BLOOD PLASMA AND SEAWATER

There was a time when people believed that the ionic composition of blood plasma resembled that of seawater. This was supposed to be evidence that primitive organisms lived in the ocean and land animals evolved a system of retaining the ocean-like composition of salts.

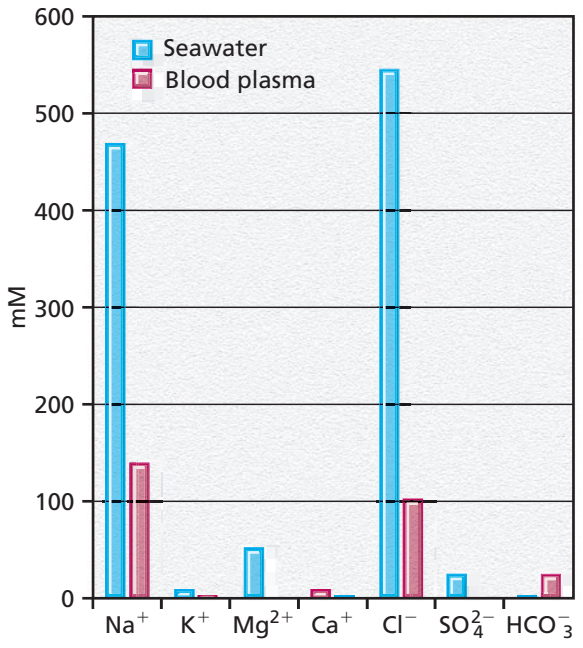
Careful studies of salt concentrations in the early 20th century revealed that the concentration of salts in the ocean were much higher than in blood plasma. Some biochemists tried to explain this discrepancy by postulating that the composition of blood plasma didn't resemble the seawater of today but it did resemble the composition of ancient seawater from several hundred million years ago when multicellular animals arose.

We now know that the saltiness of the ocean hasn't changed very much from the time it first formed over three billion years ago. There is no direct connection between the saltiness of blood plasma and seawater. Not only are the overall concentrations of the major ions (Na^+, K^+, and Cl^-) very different but the relative concentrations of various other ionic species are even more different.

The ionic composition of blood plasma is closely mimicked by Ringer's solution, which also contains lactate as a carbon source. Ringer's solution can be used as a temporary substitute for blood plasma when a patient has suffered blood loss or dehydration.

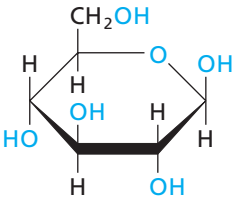
	Blood plasma	Ringer's
Na^+	140 mM	130 mM
K^+	4 mM	4 mM
Cl^-	103 mM	109 mM
Ca^+	2 mM	2 mM
lactate	5 mM	28 mM

▼ The concentrations of various ions in seawater (blue) and human blood plasma (red) are compared. Seawater is much saltier and contains much higher proportions of magnesium and sulfates. Blood plasma is enriched in bicarbonate (see Section 2.10).





Downloaded From Vet-eBooks.com



▲ **Figure 2.7**
Structure of glucose. Glucose contains five hydroxyl groups and a ring oxygen, each of which can form hydrogen bonds with water.

Each dissolved $Na^{\oplus}$ attracts the negative ends of several water molecules whereas each dissolved $Cl^{\ominus}$ attracts the positive ends of several water molecules (Figure 2.6b). The shell of water molecules that surrounds each ion is called a solvation sphere and it usually contains several layers of solvent molecules. A molecule or ion surrounded by solvent molecules is said to be **solvated**. When the solvent is water, such molecules or ions are said to be **hydrated**.

Electrolytes are not the only hydrophilic substances that are soluble in water. Any polar molecule will have a tendency to become solvated by water molecules. In addition, the solubility of many organic molecules is enhanced by formation of hydrogen bonds with water molecules. Ionic organic compounds such as carboxylates and protonated amines owe their solubility in water to their polar functional groups. Other groups that confer water solubility include amino, hydroxyl, and carbonyl groups. Molecules containing such groups disperse among water molecules with their polar groups forming hydrogen bonds with water.

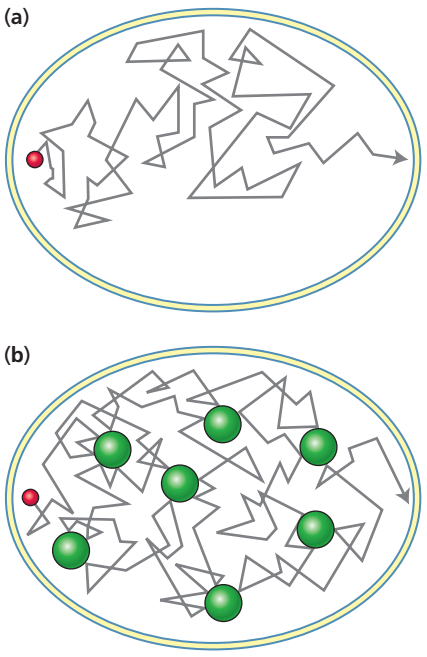
An increase in the number of polar groups in an organic molecule increases its solubility in water. The carbohydrate glucose contains five hydroxyl groups and a ring oxygen (Figure 2.7) and is very soluble in water (up to 83 grams of glucose can dissolve in 100 milliliters of water at 17.5°C). Each oxygen atom of glucose can form hydrogen bonds with water. We will see in other chapters that the attachment of carbohydrates to some otherwise poorly soluble molecules, including lipids and the bases of nucleosides, increases their solubility.

B. Cellular Concentrations and Diffusion

The inside of a cell can be very crowded as suggested by David Goodsell's drawings (Figure 1.17). Consequently, the behavior of solutes in the cytoplasm will be different from their behavior in a simple solution of water. One of the most important differences is reduction of the diffusion rate inside cells.

There are three reasons why solutes diffuse more slowly in cytoplasm.

1. The viscosity of cytoplasm is higher than that of water due to the presence of many solutes such as sugars. This is not an important factor because recent measurements suggest that the viscosity of cytoplasm is only slightly greater than water even in densely packed organelles.
2. Charged molecules bind transiently to each other inside cells and this restricts their mobility. These binding effects have a small but significant effect on diffusion rates.
3. Collisions with other molecules inhibit diffusion due to an effect called **molecular crowding**. This is the main reason why diffusion is slowed in the cytoplasm.



▲ **Figure 2.8**
Diffusion. (a) If the cytoplasm were simply made up of water, a small molecule (red) would diffuse from one end of a cell to the other via a random walk. **(b)** The average time could be about 10 times longer in a crowded cytoplasm, with larger molecules (green).

For small molecules, the diffusion rate inside cells is never more than one-quarter the rate in pure water. For large molecules, such as proteins, the diffusion rate in the cytoplasm may be slowed to about 5% to 10% of the rate in water. This slowdown is due largely to molecular crowding.

For an individual molecule, the rate of diffusion in water at 20°C is described by the diffusion coefficient ($D_{20,w}$). For the protein myoglobin, $D_{20,w} = 11.3 \times 10^{-7}$ cm^2 s^{-1}. From this value we can calculate that the average time to diffuse from one end of a cell to the other (~10 μm) is about 0.44 seconds.

But this diffusion time represents the diffusion time in pure water. In the crowed environment of a typical cell it could take about 10 times longer (4 s). The slower rate is due to the fact that a protein like myoglobin will be constantly bumping into other large molecules. Nevertheless, 4 seconds is still a short time. It means that most molecules, including smaller metabolites and ions, will encounter each other frequently inside a typical cell (Figure 2.8). Recent direct measurements of diffusion inside cells reveal that the effects of molecular crowding are less significant than we used to believe.

C. Osmotic Pressure

If a solvent-permeable membrane separates two solutions that contain different concentrations of dissolved substances, or solutes, then molecules of *solvent* will diffuse from the less concentrated solution to the more concentrated solution in a process

Downloaded From Vet-e-Books.com

called **osmosis**. The pressure required to prevent the flow of solvent is called **osmotic pressure**. The osmotic pressure of a solution depends on the total molar concentration of solute, not on its chemical nature.

Water-permeable membranes separate the cytosol from the external medium. The compositions of intracellular solutions are quite different from those of extracellular solutions with some compounds being more concentrated and some less concentrated inside cells. In general, the concentrations of solutes inside the cell are much higher than their concentrations in the aqueous environment outside the cell. Water molecules tend to move across the cell membrane in order to enter the cell and dilute the solution inside the cell. The influx of water causes the cell's volume to increase but this expansion is limited by the cell membrane. In extreme cases, such as when red blood cells are diluted in pure water, the internal pressure causes the cells to burst. Some species (e.g., plants and bacteria) have rigid cell walls that prevent the membrane expansion. These cells can develop high internal pressures.

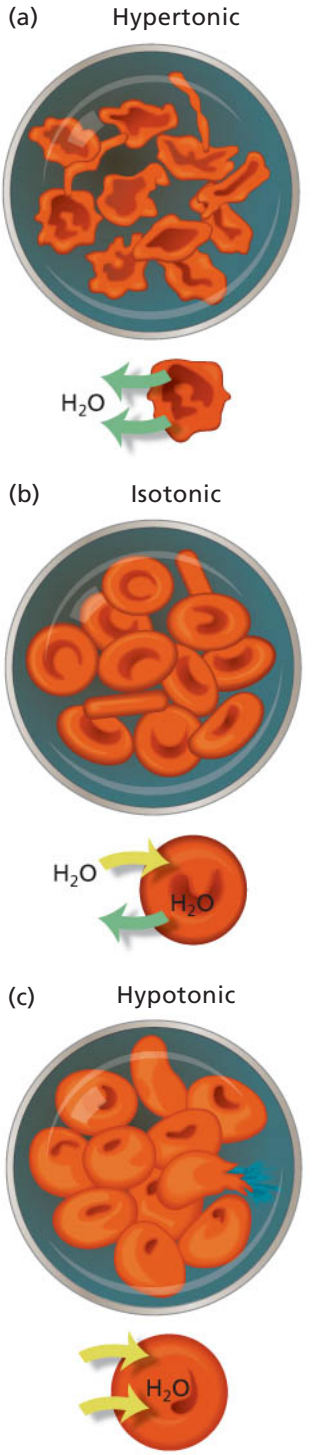
Most cells use several strategies to keep the osmotic pressure from becoming too great and bursting the cell. One strategy involves condensing many individual molecules into a macromolecule. For example, animal cells that store glucose package it as a polymer called glycogen which contains about 50,000 glucose residues. If the glucose molecules were not condensed into a single glycogen molecule the influx of water necessary to dissolve each glucose molecule would cause the cell to swell and burst. Another strategy is to surround cells with an isotonic solution that negates a net efflux or influx of water. Blood plasma, for example, contains salts and other molecules that mimic the osmolarity inside red blood cells (see Box 2.2).

2.4 Nonpolar Substances Are Insoluble in Water

Hydrocarbons and other nonpolar substances have very low solubility in water because water molecules tend to interact with other water molecules rather than with nonpolar molecules. As a result, water molecules exclude nonpolar substances forcing them to associate with each other. For example, tiny oil droplets that are vigorously dispersed in water tend to coalesce to form a single drop thereby minimizing the area of contact between the two substances. This is why the oil in a salad dressing separates if you let it sit for any length of time before putting it on your salad.

Nonpolar molecules are said to be **hydrophobic**, or water fearing, and this phenomenon of exclusion of nonpolar substances by water is called the **hydrophobic effect**. The hydrophobic effect is critical for the folding of proteins and the self-assembly of biological membranes.

The number of polar groups in a molecule affects its solubility in water. Solubility also depends on the ratio of polar to nonpolar groups in a molecule. For example, one-, two-, and three-carbon alcohols are miscible with water but larger hydrocarbons with single hydroxyl groups are much less soluble in water (Table 2.1). In the larger



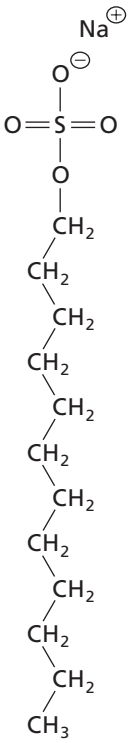
▲ Hypertonic (a), isotonic (b) and hypotonic (c) red blood cells.

Table 2.1 Solubilities of short-chain alcohols in water

Alcohol	Structure	Solubility in water (mol/100 g H_2O at 20°C)[a]
Methanol	CH_3OH	∞
Ethanol	CH_3CH_2OH	∞
Propanol	$CH_3(CH_2)_2OH$	∞
Butanol	$CH_3(CH_2)_3OH$	0.11
Pentanol	$CH_3(CH_2)_4OH$	0.030
Hexanol	$CH_3(CH_2)_5OH$	0.0058
Heptanol	$CH_3(CH_2)_6OH$	0.0008

[a] Infinity (∞) indicates that there is no limit to the solubility of the alcohol in water.

Downloaded From Vet-eBooks.com



▲ Figure 2.9
Sodium dodecyl sulfate (SDS), a synthetic detergent.

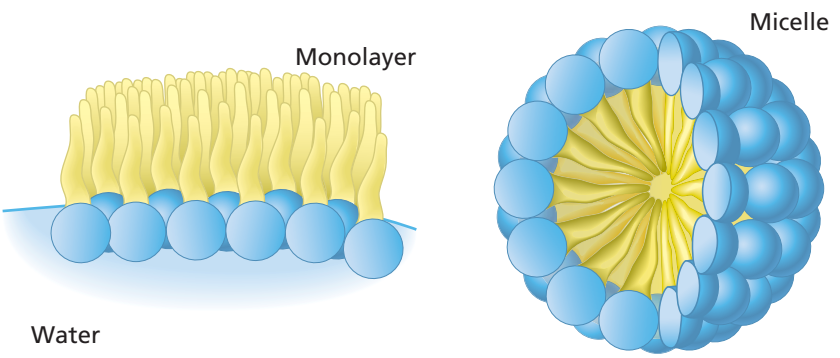
molecules, the properties of the nonpolar hydrocarbon portion of the molecule override those of the polar alcohol group and limit solubility.

Detergents, sometimes called surfactants, are molecules that are both hydrophilic and hydrophobic. They usually have a hydrophobic chain at least 12 carbon atoms long and an ionic or polar end. Such molecules are said to be **amphipathic**. Soaps, which are alkali metal salts of long-chain fatty acids are one type of detergent. The soap sodium palmitate ($CH_3(CH_2)_{14}COO^{\ominus}Na^{\oplus}$), for example, contains a hydrophilic carboxylate group and a hydrophobic tail. One of the synthetic detergents most commonly used in biochemistry is sodium dodecyl sulfate (SDS) which contains a 12-carbon tail and a polar sulfate group (Figure 2.9).

The hydrocarbon portion of a detergent is soluble in nonpolar organic substances and its polar group is soluble in water. When a detergent is spread on the surface of water a monolayer forms in which the hydrophobic, nonpolar tails of the detergent molecules extend into the air groups of detergent molecules aggregate into micelles while the hydrophilic, ionic heads are hydrated, extending into the water (Figure 2.10). When a sufficiently high concentration of detergent is dispersed in water rather than layered on the surface. In one common form of micelle, the nonpolar tails of the detergent molecules associate with one another in the center of the structure minimizing contact with water molecules. Because the tails are flexible, the core of a micelle is liquid hydrocarbon. The ionic heads project into the aqueous solution and are therefore hydrated. Small, compact micelles may contain about 80 to 100 detergent molecules.

The cleansing action of soaps and other detergents derives from their ability to trap water-insoluble grease and oils within the hydrophobic interiors of micelles. SDS and similar synthetic detergents are common active ingredients in laundry detergents. The suspension of nonpolar compounds in water by their incorporation into micelles is termed *solubilization*. Solubilizing nonpolar molecules is a different process than dissolving a polar compound. A number of the structures that we will encounter later in this book, including proteins and biological membranes, resemble micelles in having hydrophobic interiors and hydrophilic surfaces.

Some dissolved ions such as $SCN^{\ominus}$ (thiocyanate) and $ClO_4^{\ominus}$ (perchlorate) are called **chaotropes**. These ions are poorly solvated compared to ions such as $NH4^{\oplus}$, $SO_4^{2\ominus}$, and $H_2PO_4^{\ominus}$. Chaotropes enhance the solubility of nonpolar compounds in water by disordering the water molecules (there is no general agreement on how chaotropes do this). We will encounter other examples of chaotropic agents such as the guanidinium ion and the nonionic compound urea when we discuss denaturation and the three-dimensional structures of proteins and nucleic acids.



▲ Figure 2.10
Cross-sectional views of structures formed by detergents in water. Detergents can form monolayers at the air–water interface. They can also form micelles, aggregates of detergent molecules in which the hydrocarbon tails (yellow) associate in the water-free interior and the polar head groups (blue) are hydrated.

Downloaded From Vet-e-Books.com

2.5 Noncovalent Interactions

So far in this chapter we have introduced two types of noncovalent interactions—hydrogen bonds and hydrophobic interactions. Weak interactions such as these play extremely important roles in determining the structures and functions of macromolecules. Weak forces are also involved in the recognition of one macromolecule by another and in the binding of reactants to enzymes.

There are actually four major noncovalent bonds or forces. In addition to hydrogen bonds and hydrophobicity there are also charge–charge interactions and van der Waals forces. Charge–charge interactions, hydrogen bonds, and van der Waals forces are variations of a more general type of force called **electrostatic interactions**.

A. Charge–Charge Interactions

Charge–charge interactions are electrostatic interactions between two charged particles. These interactions are potentially the strongest noncovalent forces and can extend over greater distances than other noncovalent interactions. The stabilization of NaCl crystals by interionic attraction between the sodium ($Na^{\oplus}$) and chloride ($Cl^{\ominus}$) ions is an example of a charge–charge interaction. The strength of such interactions in solution depends on the nature of the solvent. Since water greatly weakens these interactions, the stability of macromolecules in an aqueous environment is not strongly dependent on charge–charge interactions but they do occur. An example of charge-charge interactions in proteins is when oppositely charged functional groups attract one another. The interaction is sometimes called a **salt bridge** and it's usually buried deep within the hydrophobic interior of a protein where it can't be disrupted by water molecules. The most accurate term for such interactions is **ion pairing**.

Charge–charge interactions are also responsible for the mutual repulsion of similarly charged ionic groups. Charge repulsion can influence the structures of individual biomolecules as well as their interactions with other, like-charged molecules.

In addition to their relatively minor contribution to the stabilization of large molecules, charge–charge interactions play a role in the recognition of one molecule by another. For example, most enzymes have either anionic or cationic sites that bind oppositely charged reactants.

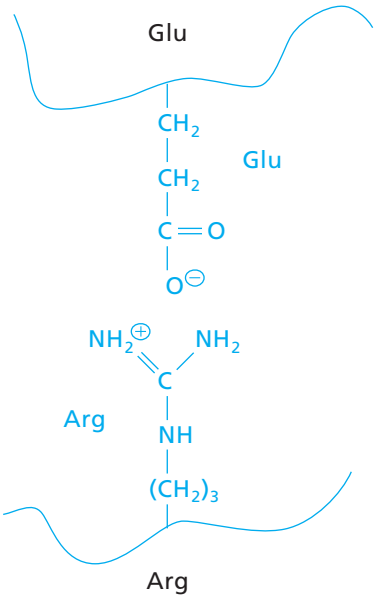
B. Hydrogen Bonds

Hydrogen bonds, which are also a type of electrostatic interaction, occur in many macromolecules and are among the strongest noncovalent forces in biological systems. The strengths of hydrogen bonds such as those between substrates and enzymes and those between the bases of DNA are estimated to be about 25–30 kJ mol^{-1}. These hydrogen bonds are a bit stronger than those formed between water molecules (Section 2.2). Hydrogen bonds in biochemical molecules are strong enough to confer structural stability but weak enough to be broken readily.

In general, when a hydrogen atom is covalently bonded to a strongly electronegative atom, such as nitrogen, oxygen, or sulfur, a hydrogen bond can only form when the hydrogen atom lies approximately 0.2 nm from another strongly electronegative atom with an unshared electron pair. As previously described in the case of hydrogen bonds between water molecules the covalently bonded atom (designated D in Figure 2.11a) is the hydrogen donor and the atom that attracts the proton (designated A in Figure 2.11a) is the hydrogen acceptor. The total distance between the two electronegative atoms participating in a hydrogen bond is typically between 0.27 nm and 0.30 nm. Some common examples of hydrogen bonds are shown in Figure 2.11b.

A hydrogen bond has many of the characteristics of a covalent bond but it is much weaker. You can think of a hydrogen bond as a partial sharing of electrons. (Recall that in a true covalent bond a pair of electrons is shared between two atoms.) The three atoms involved in a hydrogen bond are usually aligned to form a straight line where the center of the hydrogen atoms falls directly on a line drawn between the two electronegative

(a)

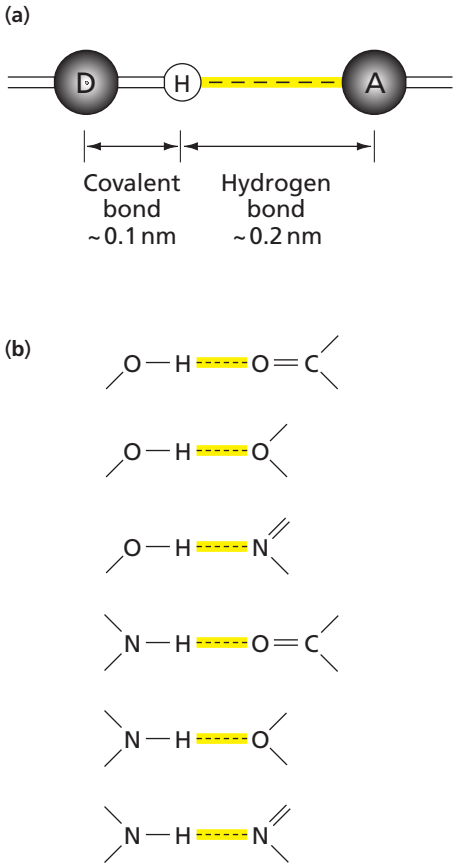


(b)



▲ **Salt bridges. (a)** One kind of salt bridge. **(b)** Another kind of salt bridge.

Downloaded From Vet-e-Books.com



▲ **Figure 2.11**
Hydrogen bonds. (a) Hydrogen bonding between a —D—H group (the hydrogen donor) and an electronegative atom A—(the hydrogen acceptor). A typical hydrogen bond is approximately 0.2 nm long, roughly twice the length of the covalent bond between hydrogen and nitrogen, oxygen, or sulfur. The total distance between the two electronegative atoms participating in a hydrogen bond is therefore approximately 0.3 nm. **(b)** Examples of biologically important hydrogen bonds.

Hydrogen bonding between base pairs in double-stranded DNA makes only a small contribution to the stability of DNA, as described in Section 19.2C.

KEY CONCEPT

Hydrogen bonds between and within biological molecules are easily disrupted by competition with water molecules.

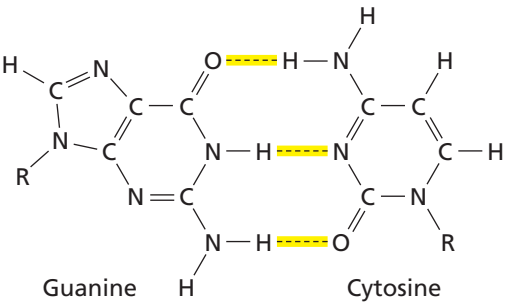


Figure 2.12 ▲
Hydrogen bonding between the complementary bases guanine and cytosine in DNA.

atoms. Small deviations from this alignment are permitted but such hydrogen bonds are weaker than the standard form.

All of the functional groups shown in Figure 2.11 are also capable of forming hydrogen bonds with water molecules. In fact, when they are exposed to water they are far more likely to interact with water molecules because the concentration of water is so high. In order for hydrogen bonds to form between, or within, biochemical macromolecules the donor and acceptor groups have to be shielded from water. In most cases, this shielding occurs because the groups are buried in the hydrophobic interior of the macromolecule where water can't penetrate. In DNA, for example, the hydrogen bonds between complementary base pairs are in the middle of the double helix (Figure 2.12).

C. Van der Waals Forces

The third weak force involves the interactions between permanent or transient dipoles of two molecules. These forces are of short range and small magnitude, about 13 kJ mol^{-1} and 0.8 kJ mol^{-1}, respectively.

These electrostatic interactions are called **van der Waals forces** named after the Dutch physicist Johannes Diderik van der Waals. They only occur when atoms are very close together. Van der Waals forces involve both attraction and repulsion. The attractive forces, also known as London dispersion forces, originate from the infinitesimal dipole generated in atoms by the random movement of the negatively charged electrons around the positively charged nucleus. Thus, van der Waals forces are dipolar, or electrostatic, attractions between the nuclei of atoms or molecules and the electrons of other atoms or molecules. The strength of the interaction between the transiently induced dipoles of nonpolar molecules such as methane is about 0.4 kJ mol^{-1} at an internuclear separation of 0.3 nm. Although they operate over similar distances, van der Waals forces are much weaker than hydrogen bonds.

There is also a repulsive component to van der Waals forces. When two atoms are squeezed together the electrons in their orbitals repel each other. The repulsion increases exponentially as the atoms are pressed together and at very close distances it becomes prohibitive.

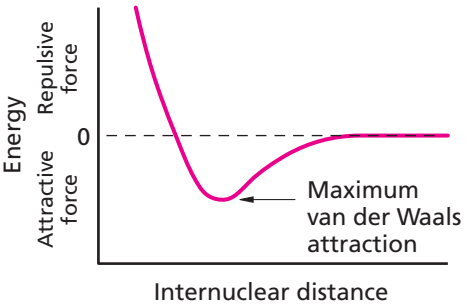
The sum of the attractive and repulsive components of van der Waals forces yields an energy profile like that in Figure 2.13. At large intermolecular distances the two atoms do not interact and there are no attractive or repulsive forces between them. As the atoms approach each other (moving toward the left in the diagram) the attractive force increases. This attractive force is due to the delocalization of the electron cloud around the atoms. You can picture this as a shift in electrons around one of the atoms such that the electrons tend to localize on the side opposite that of the other approaching atom. This shift creates a local dipole where one side of the atom has a slight positive charge and the other side has a slight negative charge. The side with the small positive charge attracts the other negatively charged atom. As the atoms move even closer together the effect of this dipole diminishes and the overall influence of the negatively charged electron cloud becomes more important. At short distances the atoms repel each other.

Downloaded From Vet-e-Books.com

The optimal packing distance is the point at which the attractive forces are maximized. This distance corresponds to the energy trough in Figure 2.13 and it is equal to the sum of the van der Waals radii of the two atoms. When the atoms are separated by the sum of their two van der Waals radii they are said to be in van der Waals contact. Typical van der Waals radii of several atoms are shown in Table 2.2.

In some cases, the shift in electrons is influenced by the approach of another atom. This is an induced dipole. In other cases, the delocalization of electrons is a permanent feature of the molecule as we saw in the case of water (Section 2.1). These permanent dipoles also give rise to van der Waals forces.

Although individual van der Waals forces are weak, the clustering of atoms within a protein, nucleic acid, or biological membrane permits formation of a large number of these weak interactions. Once formed, these cumulative weak forces play important roles in maintaining the structures of the molecules. For example, the heterocyclic bases of nucleic acids are stacked one above another in double-stranded DNA. This arrangement is stabilized by a variety of noncovalent interactions, especially van der Waals forces. These forces are collectively known as stacking interactions (see Chapter 19).



▲ **Figure 2.13**
Effect of internuclear separation on van der Waals forces. Van der Waals forces are strongly repulsive at short internuclear distances and very weak at long internuclear distances. When two atoms are separated by the sum of their van der Waals radii, the van der Waals attraction is maximal.

D. Hydrophobic Interactions

The association of a relatively nonpolar molecule or group with other nonpolar molecules is termed a **hydrophobic interaction**. Although hydrophobic interactions are sometimes called hydrophobic "*bonds*," this description is incorrect. Nonpolar molecules don't aggregate because of mutual attraction but because the polar water molecules surrounding them tend to associate with each other rather than with the nonpolar molecules (Section 2.4). For example, micelles (Figure 2.10) are stabilized by hydrophobic interactions.

The hydrogen-bonding pattern of water is disrupted by the presence of a nonpolar molecule. Thus, water molecules surrounding a less polar molecule in solution are more restricted in their interactions with other water molecules. These restricted water molecules are relatively immobile, or ordered, in the same way that molecules at the surface of water are ordered in the familiar phenomenon of surface tension. However, water molecules in the bulk solvent phase are much more mobile, or disordered. In thermodynamic terms, there is a net gain in the combined entropy of the solvent and the nonpolar solute when the nonpolar groups aggregate and water is freed from its ordered state surrounding the nonpolar groups.

Hydrophobic interactions, like hydrogen bonds, are much weaker than covalent bonds but stronger than van der Waals interactions. For example, the energy required to transfer a —CH_2— group from a hydrophobic to an aqueous environment is about 3 kJ mol^{-1}.

Although individual hydrophobic interactions are weak, the cumulative effect of many hydrophobic interactions can have a significant effect on the stability of a macromolecule. The three-dimensional structure of most proteins, for example, is largely determined by hydrophobic interactions formed during the spontaneous folding of the polypeptide chain. Water molecules are bound to the outside surface of the protein but can't penetrate the interior where most of the nonpolar groups are located.

All four of the interactions covered here are individually weak compared to covalent bonds but the combined effect of many such weak interactions can be quite strong. The most important noncovalent interactions in biomolecules are shown in Figure 2.14.

Table 2.2 Van der Waals radii of several atoms

Atom	Radius (nm)
Hydrogen	0.12
Oxygen	0.14
Nitrogen	0.15
Carbon	0.17
Sulfur	0.18
Phosphorus	0.19

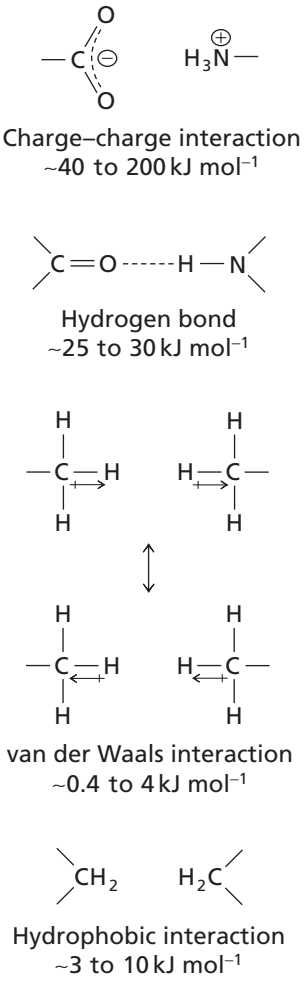
KEY CONCEPT

Weak interactions are individually weak but the combined effect of a large number of weak interactions is a significant organizing force.

2.6 Water Is Nucleophilic

In addition to its physical properties, the chemical properties of water are also important in biochemistry because water molecules can react with biological molecules. The electron-rich oxygen atom determines much of water's reactivity in chemical reactions. Electron-rich chemicals are called **nucleophiles** (nucleus lovers) because they seek positively charged (electron-deficient) species called **electrophiles** (electron lovers). Nucleophiles are either negatively charged or have unshared pairs of electrons. They attack

Downloaded From Vet-e-Books.com



▲ Figure 2.14
Typical noncovalent interactions in biomolecules. Charge–charge interactions, hydrogen bonds, and van der Waals interactions are electrostatic interactions. Hydrophobic interactions depend on the increased entropy of the surrounding water molecules rather than on direct attraction between nonpolar groups. For comparison, the dissociation energy for a covalent bond such as C—H or C—C is approximately 340–450 kJ mol^{-1}.

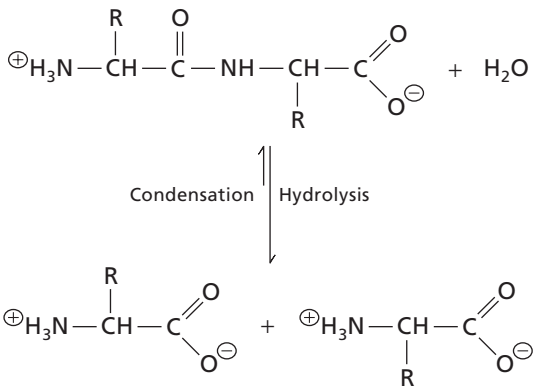


Figure 2.15 ▲
Hydrolysis of a peptide. In the presence of water the peptide bonds in proteins and peptides are hydrolyzed. Condensation, the reverse of hydrolysis, is not thermodynamically favored.

electrophiles during substitution or addition reactions. The most common nucleophilic atoms in biology are oxygen, nitrogen, sulfur, and carbon.

The oxygen atom of water has two unshared pairs of electrons making it nucleophilic. Water is a relatively weak nucleophile but its cellular concentration is so high that one might reasonably expect it to be very reactive. Many macromolecules should be easily degraded by nucleophilic attack by water. This is, in fact, a correct expectation. Proteins, for example, are hydrolyzed, or degraded, by water to release their monomeric units, amino acids (Figure 2.15). The equilibrium for complete hydrolysis of a protein lies far in the direction of degradation; in other words, the ultimate fate of all proteins is destruction by hydrolysis!

If there is so much water in cells then why aren't all biopolymers rapidly degraded? Similarly, if the equilibrium lies toward breakdown, how does biosynthesis occur in an aqueous environment? Cells avoid these problems in several ways. For example, the linkages between the monomeric units of macromolecules, such as the peptide bonds in proteins and the ester linkages in DNA, are relatively stable in solution at cellular pH and temperature in spite of the presence of water. In this case, the stability of linkages refers to their rate of hydrolysis in water and not their thermodynamic stability.

The chemical properties of water combined with its high concentration mean that the Gibbs free energy change for hydrolysis (ΔG) is negative. This means that all hydrolysis reactions are thermodynamically favorable. However, the rate of the reactions inside the cell is so slow that macromolecules are not appreciably degraded by spontaneous hydrolysis during the average lifetime of a cell. It is important to keep in mind the distinction between the preferred direction of a reaction, as indicated by the Gibbs free energy change, and the rate of the reaction, as indicated by the rate constant (Section 1.4D). The key concept is that because of the activation energy there is no direct correlation between the rate of a reaction and the final equilibrium values of the reactants and products.

Cells can synthesize macromolecules in an aqueous environment even though condensation reactions—the reverse of hydrolysis—are thermodynamically unfavorable. They do this by using the chemical potential energy of ATP to overcome an unfavorable thermodynamic barrier. Furthermore, the enzymes that catalyze such reactions exclude water from the active site where the synthesis reactions occur. These reactions usually follow two-step chemical pathways that differ from the reversal of hydrolysis. For example, the simple condensation pathway shown in Figure 2.15 is not the pathway that is used in living cells because the presence of high concentrations of water makes the direct condensation reaction extremely unfavorable. In the first synthetic step, which is thermodynamically uphill, the molecule to be transferred reacts with ATP to form a reactive intermediate. In the second step, the activated group is readily

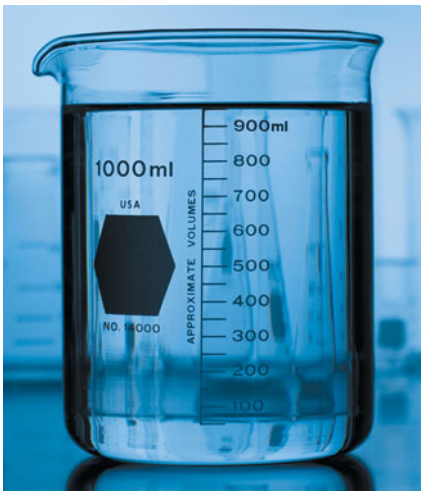
Downloaded From Vet-e-Books.com

BOX 2.3 THE CONCENTRATION OF WATER

The density of water varies with temperature. It is defined as 1.00000 g/ml at 3.98°C. The density is 0.99987 at 0°C and 0.99707 at 25°C.

The molecular mass of the most common form of water is M_r =18.01056. The concentration of pure water at 3.98°C is 55.5 M (1000 ÷ 18.01).

Many biochemical reactions involve water as either a reactant or a product and the high concentration of water will affect the equilibrium of the reaction.



KEY CONCEPT

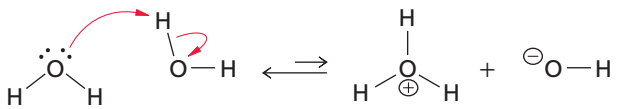
There is a difference between the rate of a reaction and whether it is thermodynamically favorable. Biological molecules are stable because the rate of spontaneous hydrolysis is slow.

transferred to the attacking nucleophile. In Chapter 22 we will see that the reactive intermediate in protein synthesis is an aminoacyl-tRNA that is formed in a reaction involving ATP. The net result of the biosynthesis reaction is to couple the condensation to the hydrolysis of ATP.

The role of ATP in coupled reactions is described in Section 10.7.

2.7 Ionization of Water

One of the important properties of water is its slight tendency to ionize. Pure water contains a low concentration of hydronium ions ($H_3O^{\oplus}$) and an equal concentration of hydroxide ions ($OH^{\ominus}$). The hydronium and hydroxide ions are formed by a nucleophilic attack of oxygen on one of the protons in an adjacent water molecule.



$$H_2O + H_2O \rightleftharpoons H_3O^{\oplus} + OH^{\ominus} \qquad (2.2)$$

The red arrows in Reaction 2.2 show the movement of pairs of electrons. These arrows are used to depict reaction mechanisms and we will encounter many such diagrams throughout this book. One of the free pairs of electrons on the oxygen will contribute to formation of a new O—H covalent bond between the oxygen atom of the hydronium ion and a proton ($H^{\oplus}$) abstracted from a water molecule. An O—H covalent bond is broken in this reaction and the electron pair from that bond remains associated with the oxygen atom of the hydroxide ion.

Note that the atoms in the hydronium ion contain eleven positively charged protons (eight in the oxygen atom and three hydrogen protons) and ten negatively charged electrons (a pair of electrons in the inner orbital of the oxygen atom, one free electron pair associated with the oxygen atom, and three pairs in the covalent bonds). This results in a net positive charge which is why we refer to it as an ion (cation). The positive charge is usually depicted as if it were associated with the oxygen atom but, in fact, it is distributed partially over the hydrogen atoms as well. Similarly, the hydroxide ion (anion) bears a net negative charge because it contains ten electrons whereas the nuclei of the oxygen and hydrogen atoms have a total of only nine positively charged protons.

Downloaded From Vet-e-Books.com

The ionization reaction is a typical reversible reaction. The protonation and deprotonation reactions take place very quickly. Hydroxide ions have a short lifetime in water and so do hydronium ions. Even water molecules themselves have only a transient existence. The average water molecule is thought to exist for about one millisecond (10^{-3}s) before losing a proton to become a hydroxide ion or gaining a proton to become a hydronium ion. Note that the lifetime of a water molecule is still eight orders of magnitude (10^{8}) greater than the lifetime of a hydrogen bond.

Hydronium ($H_3O^{\oplus}$) ions are capable of donating a proton to another ion. Such proton donors are referred to as **acids** according to the Brønsted–Lowry concept of acids and bases. In order to simplify chemical equations we often represent the hydronium ion as simply $H^{\oplus}$ (free proton or hydrogen ion) to reflect the fact that it is a major source of protons in biochemical reactions. The ionization of water can then be depicted as a simple dissociation of a proton from a single water molecule.

$$H_2O \rightleftharpoons H^{\oplus} + OH^{\ominus} \tag{2.3}$$

Reaction 2.3 is a convenient way to show the ionization of water but it does not reflect the true structure of the proton donor which is actually the hydronium ion. Reaction 2.3 also obscures the fact that the ionization of water is actually a bimolecular reaction involving two separate water molecules as shown in Reaction 2.2. Fortunately, the dissociation of water is a reasonable approximation that does not affect our calculations or our understanding of the properties of water. We will make use of this assumption in the rest of the book.

Hydroxide ions can accept a proton and be converted back into water molecules. Proton acceptors are called **bases**. Water can function as either an acid or a base as Reaction 2.2 demonstrates.

The ionization of water can be analyzed quantitatively. Recall that the concentrations of reactants and products in a reaction will eventually reach an equilibrium where there is no net change in concentration. The ratio of these equilibrium concentrations defines the equilibrium constant (K_{eq}). In the case of ionization of water,

$$K_{eq} = \frac{[H^{\oplus}][OH^{\ominus}]}{H_2O} \qquad K_{eq}[H_2O] = [H^{\oplus}][OH^{\ominus}] \tag{2.4}$$

The density of water varies with the temperature (Box 2.2) and so does the ion product. The differences aren't significant in the temperature ranges that we normally encounter in living cells, so we assume that the value 10^{-14} applies at all temperatures. (See Problem 17 at the end of this chapter.)

The equilibrium constant for the ionization of water has been determined under standard conditions of pressure (1 atm) and temperature (25°C). Its value is 1.8×10^{-16} M. We are interested in knowing the concentrations of protons and hydroxide ions in a solution of pure water since these ions participate in many biochemical reactions. These values can be calculated from Equation 2.4 if we know the concentration of water ($[H_2O]$) at equilibrium. Pure water at 25°C has a concentration of approximately 55.5 M (see Box 2.2). A very small percentage of water molecules will dissociate to form $H^{\oplus}$ and $OH^{\ominus}$ when the ionization reaction reaches equilibrium. This will have a very small effect on the final concentration of water molecules at equilibrium. We can simplify our calculations by assuming that the concentration of water in Equation 2.4 is 55.5 M. Substituting this value, and that of the equilibrium constant, gives

$$(1.8 \times 10^{-16}\ \text{M})(55.5\ \text{M}) = 1.0 \times 10^{-14}\ \text{M}^2 = [H^{\oplus}][OH^{\ominus}] \tag{2.5}$$

The product obtained by multiplying the proton and hydroxide ion concentrations ($[H^{\oplus}][OH^{\ominus}]$) is called the **ion product for water**. This is a constant designated K_w (the ion product constant for water). At 25°C the value of K_w is

$$K_w = [H^{\oplus}][OH^{\ominus}] = 1.0 \times 10^{-14}\ \text{M}^2 \tag{2.6}$$

It is a fortunate coincidence that this is a nice round number rather than some awkward fraction because it makes calculations of ion concentrations much easier. Pure water is

Downloaded From Vet-e-Books.com

electrically neutral, so its ionization produces an equal number of protons and hydroxide ions $[H^{\oplus}] = [OH]$. In the case of pure water, Equation 2.6 can therefore be rewritten as

$$K_w = [H^{\oplus}]^2 = 1.0 \times 10^{-14}\ M^2 \quad (2.7)$$

Taking the square root of the terms in Equation 2.7 gives

$$[H^{\oplus}] = 1.0 \times 10^{-7}\ M \quad (2.8)$$

Since $[H^{\oplus}] = [OH^{\ominus}]$, the ionization of pure water produces 10^{-7} M $H^{\oplus}$ and 10^{-7} M $OH^{\ominus}$. Pure water and aqueous solutions that contain equal concentrations of $H^{\oplus}$ and $OH^{\ominus}$ are said to be *neutral*. Of course, not all aqueous solutions have equal concentrations of $H^{\oplus}$ and $OH^{\ominus}$. When an acid is dissolved in water $[H^{\oplus}]$ increases and the solution is described as acidic. Note that when an acid is dissolved in water the concentration of protons increases while the concentration of hydroxide ions decreases. This is because the ion product constant for water (K_w) is unchanged (i.e., constant) and the product of the concentrations of $H^{\oplus}$ and $OH^{\ominus}$ must always be 1.0×10^{-14} M^2 under standard conditions (Equation 2.5). Dissolving a base in water decreases $[H^{\oplus}]$ and increases $[OH^{\ominus}]$ above 1.0×10^{-7} M producing a basic, or alkaline, solution.

2.8 The pH Scale

Many biochemical processes—including the transport of oxygen in the blood, the catalysis of reactions by enzymes, and the generation of metabolic energy during respiration or photosynthesis—are strongly affected by the concentration of protons. Although the concentration of $H^{\oplus}$ (or $H_3O^{\oplus}$) in cells is small relative to the concentration of water, the range of $[H^{\oplus}]$ in aqueous solutions is enormous so it is convenient to use a logarithmic quantity called pH as a measure of the concentration of $H^{\oplus}$. **pH** is defined as the negative logarithm of the concentration of $H^{\oplus}$.

$$pH = -\log[H^{\oplus}] = \log\frac{1}{[H^{\oplus}]} \quad (2.9)$$

In pure water $[H^{\oplus}] = [OH^{\ominus}] = 1.0 \times 10^{-7}$ M (Equations 2.7 and 2.8). As mentioned earlier, pure water is said to be "neutral" with respect to total ionic charge since the concentrations of the positively charged hydrogen ions and the negatively charged hydroxide ions are equal. Neutral solutions have a pH value of 7.0 (the negative value of $\log 10^{-7}$ is 7.0). Acidic solutions have an excess of $H^{\oplus}$ due to the presence of dissolved solute that supplies $H^{\oplus}$ ions. In a solution of 0.01 M HCl, for example, the concentration of $H^{\oplus}$ is 0.01 M (10^{-2} M) because HCl dissociates completely to $H^{\oplus}$ and $Cl^{\ominus}$. The pH of such a solution is $-\log 10^{-2} = 2.0$. Thus, the higher the concentration of $H^{\oplus}$, the lower the pH of the solution. The pH scale is logarithmic, so a change in pH of one unit corresponds to a 10-fold change in the concentration of $H^{\oplus}$.

Aqueous solutions can also contain fewer $H^{\oplus}$ ions than pure water resulting in a pH above 7. In a solution of 0.01 M NaOH, for example, the concentration of $OH^{\ominus}$ is 0.01 M (10^{-2} M) because NaOH, like HCl, is 100% dissociated in water. The $H^{\oplus}$ ions derived from the ionization of water will combine with the hydroxide ions from NaOH to re-form water molecules. This affects the equilibrium for the ionization of water (Reaction 2.3). The resulting solution is very basic because of the low concentration of protons. The actual pH can be determined from the ion product of water, K_w (Equation 2.6), by substituting the concentration of hydroxide ions. Since the product of the $OH^{\ominus}$ and $H^{\oplus}$ concentrations is 10^{-14} M it follows that the $H^{\oplus}$ concentration in a solution of 10^{-2} M $OH^{\ominus}$ is 10^{-12} M. The pH of the solution is 12. Table 2.3 shows this relationship between pH and the concentrations of $H^{\oplus}$ and $OH^{\ominus}$.

Basic solutions have pH values greater than 7.0 and acidic solutions have lower pH values. Figure 2.16 illustrates the pH values of various common solutions.

Table 2.3 Relation of $[H^{\oplus}]$ and $[OH^{\ominus}]$ to pH

pH	$[H^{\oplus}]$ (M)	$[OH^{\ominus}]$ (M)
0	1	10^{-14}
1	10^{-1}	10^{-13}
2	10^{-2}	10^{-12}
3	10^{-3}	10^{-11}
4	10^{-4}	10^{-10}
5	10^{-5}	10^{-9}
6	10^{-6}	10^{-8}
7	10^{-7}	10^{-7}
8	10^{-8}	10^{-6}
9	10^{-9}	10^{-5}
10	10^{-10}	10^{-4}
11	10^{-11}	10^{-3}
12	10^{-12}	10^{-2}
13	10^{-13}	10^{-1}
14	10^{-14}	1

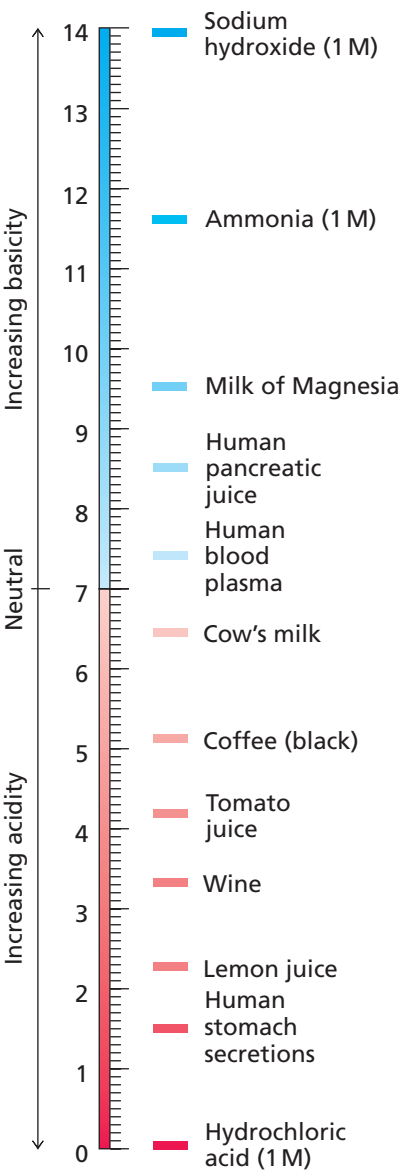


Figure 2.16 ▶
pH values for various fluids at 25°C. Lower values correspond to acidic fluids; higher values correspond to basic fluids.

Downloaded From Vet-e-Books.com

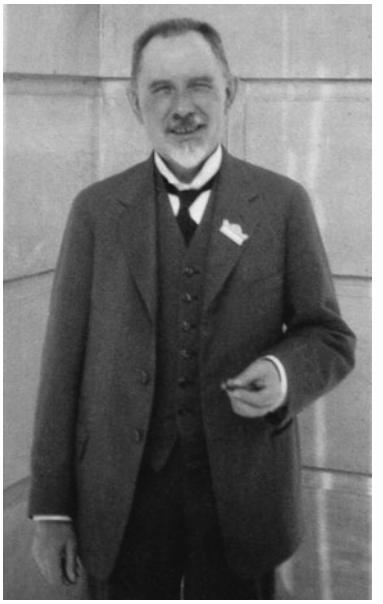


▲ **pH strips.** The approximate pH of solutions can be determined in the lab by placing a drop on a pH strip. Various indicators are bound to a matrix that is affixed to a plastic strip. The indicators change color at different concentrations of $H^{\oplus}$, and the combination of various colors gives a more or less accurate reading of the pH. The strips shown here cover all pH readings from 0 to 14 but other pH strips can be used to cover narrower ranges.

BOX 2.4 THE LITTLE "p" IN pH

The term pH was first used in 1909 by Søren Peter Lauritz Sørensen, director of the Carlsberg Laboratories in Denmark. Sørensen never mentioned what the little "p" stood for (the "H" is obviously hydrogen). Many years later, some of the scientists who write chemistry textbooks began to associate the little "p" with the words *power* or *potential*. This association, as it turns out, is based on a rather tenuous connection in some of Sørensen's early papers. A recent investigation of the historical records by Jens G. Nøby suggests that the little "p" was an arbitrary choice based on Sørensen's use of p and q to stand for unknown variables in much the same way that we might use x and y today.

No matter what the historical origin, it's important to remember that the symbol pH now stands for the negative logarithm of the hydrogen ion concentration.



▲ **Søren Peter Lauritz Sørensen (1868–1939)**

Accurate measurements of pH are routinely made using a pH meter, an instrument that incorporates a selectively permeable glass electrode that is sensitive to $[H^{\oplus}]$. Measurement of pH sometimes facilitates the diagnosis of disease. The normal pH of human blood is 7.4—frequently referred to as physiological pH. The blood of patients suffering from certain diseases, such as diabetes, can have a lower pH, a condition called acidosis. The condition in which the pH of the blood is higher than 7.4, called alkalosis, can result from persistent, prolonged vomiting (loss of hydrochloric acid from the stomach) or from hyperventilation (excessive loss of carbonic acid as carbon dioxide).

KEY CONCEPT

pH is the negative logarithm of the proton ($H^{\oplus}$) concentration.

2.9 Acid Dissociation Constants of Weak Acids

Acids and bases that dissociate completely in water, such as hydrochloric acid and sodium hydroxide, are called strong acids and strong bases. Many other acids and bases, such as the amino acids from which proteins are made and the purines and pyrimidines from DNA and RNA, do not dissociate completely in water. These substances are known as weak acids and weak bases.

KEY CONCEPT

Weak acids and weak bases are compounds that only partially dissociate in water.

In order to understand the relationship between acids and bases let us consider the dissociation of HCl in water. Recall from Section 2.7 that we define an acid as a molecule that can donate a proton and a base as a proton acceptor. Acids and bases always come in pairs since for every proton donor there must be a proton acceptor. Both sides of the dissociation reaction will contain an acid and a base. Thus, the equilibrium reaction for the complete dissociation of HCl is

$$\underset{\text{acid}}{HCl} + \underset{\text{base}}{H_2O} \rightleftharpoons \underset{\text{base}}{Cl^{\ominus}} + \underset{\text{acid}}{H_3O^{\oplus}} \qquad (2.10)$$

HCl is an acid because it can donate a proton. In this case, the proton acceptor is water which is the base in this equilibrium reaction. On the other side of the equilibrium are $Cl^{\ominus}$ and the hydronium ion, $H_3O^{\oplus}$. The chloride ion is the base that corresponds to HCl after it has given up its proton. $Cl^{\ominus}$ is called the **conjugate base** of HCl which indicates that it is a base (i.e., can accept a proton) and is part of an acid–base pair (i.e., $HCl/Cl^{\ominus}$). Similarly, $H_3O^{\oplus}$ is the acid on the right-hand side of the equilibrium because it can donate a proton. $H_3O^{\oplus}$ is the **conjugate acid** of H_2O. Every base

Downloaded From Vet-eBooks.com

has a corresponding conjugate acid and every acid has a corresponding conjugate base. Thus, HCl is the conjugate acid of $Cl^{\ominus}$ and H_2O is the conjugate base of $H_3O^{\oplus}$. Note that H_2O is the conjugate acid of $OH^{\ominus}$ if we are referring to the $H_2O/OH^{\ominus}$ acid–base pair.

In most cases throughout this book we will simplify reactions by ignoring the contribution of water and representing the hydronium ion as a simple proton.

$$HCl \rightleftharpoons H^{\oplus} + Cl^{\ominus} \tag{2.11}$$

This is a standard convention in biochemistry but, on the surface, it seems to violate the rule that both sides of the equilibrium reaction should contain a proton donor and a proton acceptor. Students should keep in mind that in such reactions the contributions of water molecules as proton acceptors and hydronium ions as the true proton donors are implied. In almost all cases we can safely ignore the contribution of water. This is the same principle that we applied to the reaction for the dissociation of water (Section 2.7) which we simplified by ignoring the contribution of one of the water molecules.

KEY CONCEPT

The contribution of water is implied in most acid/base dissociation reactions.

The reason why HCl is such a strong acid is because the equilibrium shown in Reaction 2.11 is shifted so far to the right that HCl is completely dissociated in water. In other words, HCl has a strong tendency to donate a proton when dissolved in water. This also means that the conjugate base, $Cl^{\ominus}$, is a very weak base because it will rarely accept a proton.

Acetic acid is the weak acid present in vinegar. The equilibrium reaction for the ionization of acetic acid is

$$\underset{\substack{\text{Acetic acid}\\\text{(weak acid)}}}{CH_3COOH} \overset{K_a}{\rightleftharpoons} H^{\oplus} + \underset{\substack{\text{Acetate anion}\\\text{(conjugate base)}}}{CH_3COO^{\ominus}} \tag{2.12}$$

We have left out the contribution of water molecules in order to simplify the reaction. We see that the acetate ion is the conjugate base of acetic acid. (We can also refer to acetic acid as the conjugate acid of the acetate ion.)

The equilibrium constant for the dissociation of a proton from an acid in water is called the **acid dissociation constant, K_a**. When the reaction reaches equilibrium, which happens very rapidly, the acid dissociation constant is equal to the concentration of the products divided by the concentration of the reactants. For Reaction 2.12 the acid dissociation constant is

$$K_a = \frac{[H^{\oplus}][CH_3COO^{\ominus}]}{[CH_3COOH]} \tag{2.13}$$

The K_a value for acetic acid at 25°C is 1.76×10^{-5} M. Because K_a values are numerically small and inconvenient in calculations it is useful to place them on a logarithmic scale. The parameter **pK_a** is defined by analogy with pH.

$$pK_a = -\log K_a = \log \frac{1}{K_a} \tag{2.14}$$

A pH value is a measure of the acidity of a solution and a pK_a value is a measure of the acid strength of a particular compound. The pK_a of acetic acid is 4.8.

When dealing with bases we need to consider their protonated forms in order to use Equation 2.13. These conjugate acids are very weak acids. In order to simplify calculations and make easy comparisons we measure the equilibrium constant (K_a) for the dissociation of a proton from the conjugate acid of a weak base. For example, the ammonium ion ($NH_4^{\oplus}$) can dissociate to form the base ammonia (NH_3) and $H^{\oplus}$.

$$NH_4^{\oplus} \rightleftharpoons NH_3 + H^{\oplus} \tag{2.15}$$

The acid dissociation constant (K_a) for this equilibrium is a measure of the strength of the base (ammonia, NH_3) in aqueous solution. The K_a values for several common substances are listed in Table 2.4.

Downloaded From Vet-eBooks.com

Table 2.4 Dissociation constants and pK_a values of weak acids in aqueous solutions at 25°C

Acid	K_a(M)	pK_a
HCOOH (Formic acid)	1.77×10^{-4}	3.8
CH_3COOH (Acetic acid)	1.76×10^{-5}	4.8
$CH_3CHOHCOOH$ (Lactic acid)	1.37×10^{-4}	3.9
H_3PO_4 (Phosphoric acid)	7.52×10^{-3}	2.2
$H_2PO_4^{\ominus}$ (Dihydrogen phosphate ion)	6.23×10^{-8}	7.2
$HPO_4^{2\ominus}$ (Monohydrogen phosphate ion)	2.20×10^{-13}	12.7
H_2CO_3 (Carbonic acid)	4.30×10^{-7}	6.4
$HCO_3^{\ominus}$ (Bicarbonate ion)	5.61×10^{-11}	10.2
$NH_4^{\oplus}$ (Ammonium ion)	5.62×10^{-10}	9.2
$CH_3NH_3^{\oplus}$ (Methylammonium ion)	2.70×10^{-11}	10.7

From Equation 2.13 we see that the K_a for acetic acid is related to the concentration of $H^{\oplus}$ and to the ratio of the concentrations of the acetate ion and undissociated acetic acid. If we represent the conjugate acid as HA and the conjugate base as $A^{\ominus}$ then taking the logarithm of such equations gives the general equation for any acid–base pair.

$$HA \rightleftharpoons H^{\oplus} + A^{\ominus} \qquad \log K_a = \log \frac{[H^{\oplus}][A^{\ominus}]}{[HA]} \tag{2.16}$$

Since $\log(xy) = \log x + \log y$, Equation 2.16 can be rewritten as

$$\log K_a = \log[H^{\oplus}] + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.17}$$

Rearranging Equation 2.17 gives

$$-\log[H^{\oplus}] = -\log K_a + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.18}$$

KEY CONCEPT

The pH of a solution of a weak acid or base at equilibrium can be calculated by combining the pK_a of the ionization reaction and the final concentrations of the proton acceptor and proton donor species.

The negative logarithms in Equation 2.18 have already been defined as pH and pK_a (Equations 2.9 and 2.14, respectively). Thus,

$$pH = pK_a + \log \frac{[A^{\ominus}]}{[HA]} \tag{2.19}$$

or

$$pH = pK_a + \log \frac{[\text{Proton acceptor}]}{[\text{Proton donor}]} \tag{2.20}$$

Equation 2.20 is one version of the **Henderson–Hasselbalch equation**. It defines the pH of a solution in terms of the pK_a of the weak acid form of the acid–base pair and the logarithm of the ratio of concentrations of the dissociated species (conjugate base) to the protonated species (weak acid). Note that the greater the concentration of the proton acceptor (conjugate base) relative to that of the proton donor (weak acid), the lower the concentration of $H^{\oplus}$ and the higher the pH. (Remember that pH is the *negative* log of $H^{\oplus}$ concentration. A high concentration of $H^{\oplus}$ means low pH.) This

Downloaded From Vet-e-Books.com

makes intuitive sense since the concentration of $A^{\ominus}$ is identical to the concentration of $H^{\oplus}$ in simple dissociation reactions. If more HA dissociates the concentration of $A^{\ominus}$ will be higher and so will the concentration of $H^{\oplus}$. When the concentrations of a weak acid and its conjugate base are exactly the same the pH of the solution is equal to the pK_a of the acid (since the ratio of concentrations equals 1.0, and the logarithm of 1.0 equals zero).

The Henderson–Hasselbalch equation is used to determine the final pH of a weak acid solution once the dissociation reaction reaches equilibrium as illustrated in Sample Calculation 2.1 for acetic acid. These calculations are more complicated than those involving strong acids such as HCl. As noted in Section 2.8, the pH of an HCl solution is easily determined from the amount of HCl that is present since the final concentration of $H^{\oplus}$ is equal to the initial concentration of HCl when the solution is made up. In contrast, weak acids are only partially dissociated in water so it makes sense that the pH depends on the acid dissociation constant. The pH decreases (more $H^{\oplus}$) as more weak acid is added to water but the increase in $H^{\oplus}$ is not linear with initial HA concentration. This is because the numerator in Equation 2.16 is the *product* of the $H^{\oplus}$ and $A^{\ominus}$ concentrations.

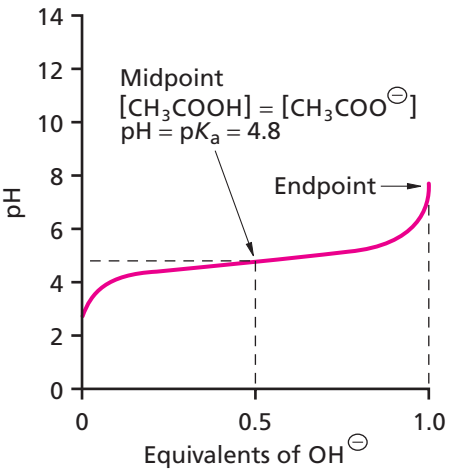
The Henderson–Hasselbalch equation applies to other acid–base combinations as well and not just to those involving weak acids. When dealing with a weak base, for example, the numerator and denominator of Equation 2.20 become [weak base] and [conjugate acid], respectively. The important point to remember is that the equation refers to the concentration of the proton acceptor divided by the concentration of the proton donor.

The pK_a values of weak acids are determined by titration. Figure 2.17 shows the titration curve for acetic acid. In this example, a solution of acetic acid is titrated by adding small aliquots of a strong base of known concentration. The pH of the solution is measured and plotted versus the number of molar equivalents of strong base added during the titration. Note that since acetic acid has only one ionizable group (its carboxyl group) only one equivalent of a strong base is needed to completely titrate acetic acid to its conjugate base, the acetate anion. When the acid has been titrated with one-half an equivalent of base the concentration of undissociated acetic acid exactly equals the concentration of the acetate anion. The resulting pH, 4.8, is thus the experimentally determined pK_a for acetic acid.

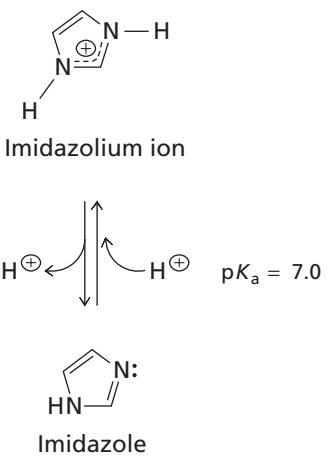
Constructing an ideal titration curve is a useful exercise for reinforcing the relationship between pH and the ionization state of a weak acid. You can use the Henderson–Hasselbalch equation to calculate the pH that results from adding increasing amounts of a strong base such as NaOH to a weak acid such as the imidazolium ion $pK_a = 7.0$. Adding base converts the imidazolium ion to its conjugate base, imidazole (Figure 2.18). The shape of the titration curve is easy to visualize if you calculate the pH when the ratio of conjugate base to acid is 0.01, 0.1, 1, 10, and 100. Calculate pH values at other ratios until you are satisfied that the curve is relatively flat near the midpoint and steeper at the ends.

Similarly shaped titration curves can be obtained for each of the five monoprotic acids (acids having only one ionizable group) listed in Table 2.4. All would exhibit the same general shape as Figure 2.17 but the inflection point representing the midpoint of titration (one-half an equivalent titrated) would fall lower on the pH scale for a stronger acid (such as formic acid or lactic acid) and higher for a weaker acid (such as ammonium ion or methylammonium ion).

Titration curves of weak acids illustrate a second important use of the Henderson–Hasselbalch equation. In this case, the final pH is the result of mixing the weak acid (HA) and a strong base ($OH^{\ominus}$). The base combines with $H^{\oplus}$ ions to form water molecules, H_2O. This reduces the concentration of $H^{\oplus}$ and raises the pH. As the titration of the weak acid proceeds it dissociates in order to restore its equilibrium with $OH^{\ominus}$ and H_2O. The net result is that the final concentration of $A^{\ominus}$ is much higher, and the concentration of HA is much lower, than when we are dealing with the simple case where the pH is determined only by the dissociation of the weak acid in water (i.e., a solution of HA in H_2O).



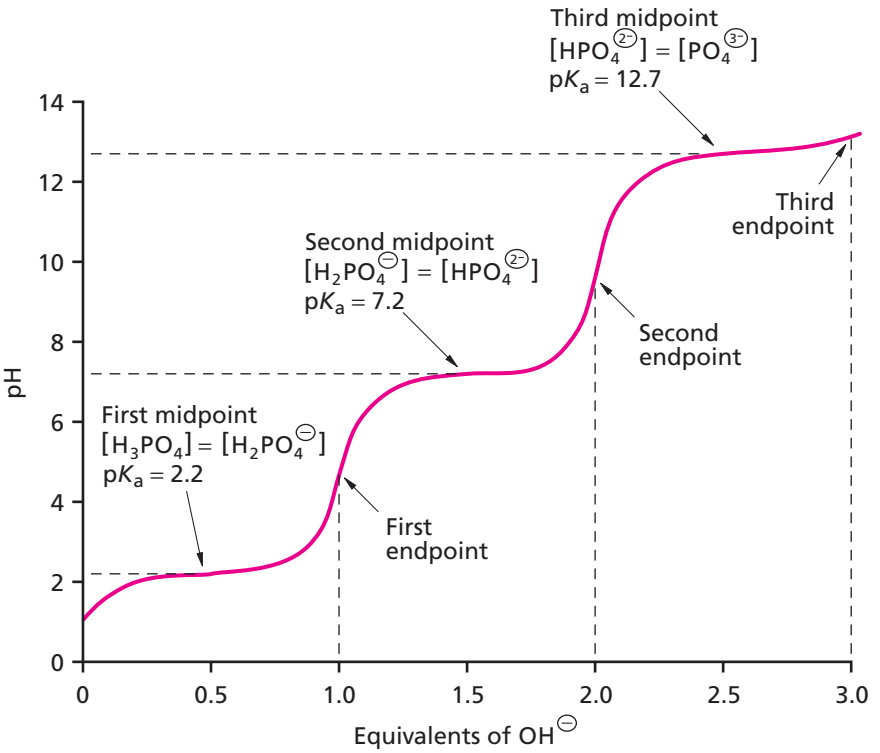
▲ **Figure 2.17**
Titration of acetic acid (CH_3COOH) with aqueous base ($OH^{\ominus}$). There is an inflection point (a point of minimum slope) at the midpoint of the titration, when 0.5 equivalent of base has been added to the solution of acetic acid. This is the point at which $[CH_3COOH] = [CH_3COO^{\ominus}]$ and $pH = pK_a$. The pK_a of acetic acid is thus 4.8. At the endpoint, all the molecules of acetic acid have been titrated to the conjugate base, acetate.



▲ **Figure 2.18**
Titration of the imidazolium ion.

Downloaded From Vet-eBooks.com

Figure 2.19 ▶
Titration curve for H_3PO_4. Three inflection points (at 0.5, 1.5, and 2.5 equivalents of strong base added) correspond to the three pK_a values for phosphoric acid (2.2, 7.2, and 12.7).



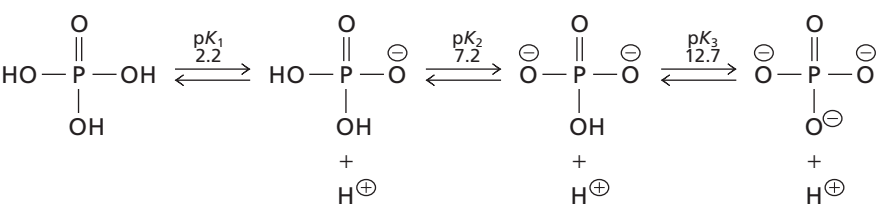


▲ Cola beverages contain phosphoric acid in order to make the drink more acidic. The concentration of phosphoric acid is about 1 mM. This concentration should make the pH about 3 in the absence of any other ingredients that may contribute to acidity.

Phosphoric acid (H_3PO_4) is a polyprotic acid. It contains three different hydrogen atoms that can dissociate to form $H^{\oplus}$ ions and corresponding conjugate bases with one, two, or three negative charges. The dissociation of the first proton occurs readily and is associated with a large acid dissociation constant of 7.53×10^{-3} M and a pK_a of 2.2 in aqueous solution. The dissociations of the second and third protons occur progressively less readily because they have to dissociate from a molecule that is already negatively charged.

Phosphoric acid requires three equivalents of strong base for complete titration and three pK_a values are evident from its titration curve (Figure 2.19). The three pK_a values reflect the three equilibrium constants and thus the existence of four possible ionic species (conjugate acids and bases) of inorganic phosphate. At physiological pH (7.4) the predominant species of inorganic phosphate are $H_2PO_4^{\ominus}$ and $HPO_4^{2\ominus}$. At pH 7.2 these two species exist in equal concentrations. The concentrations of H_3PO_4 and $PO_4^{3\ominus}$ are so low at pH 7.4 that they can be ignored. This is generally the case for a minor species when the pH is more than two units away from its pK_a.

(2.21)



Many biologically important acids and bases, including the amino acids described in Chapter 3, have two or more ionizable groups. The number of pK_a values for such substances is equal to the number of ionizable groups. The pK_a values can be experimentally determined by titration.

Downloaded From Vet-e-Books.com

Sample Calculation 2.1 CALCULATING THE pH OF WEAK ACID SOLUTIONS

Q: What is the pH of a solution of 0.1 M acetic acid?

A: The acid dissociation constant of acetic acid is 1.76×10^{-5} M. Acetic acid dissociates in water to form acetate and $H^{\oplus}$. We need to determine $[H^{\oplus}]$ when the reaction reaches equilibrium.

Let the final $H^{\oplus}$ concentration be represented by the unknown quantity x. At equilibrium the concentration of acetate ion will also be x and the final concentration of acetic acid will be $[0.1\ M - x]$. Thus,

$$1.76 \times 10^{-5} = \frac{[H^{\oplus}][CH_3COO^{\ominus}]}{[CH_3COOH]} = \frac{x^2}{(0.1 - x)}$$

rearranging gives

$$1.76 \times 10^{-6} - 1.76 \times 10^{-5}x = x^2$$
$$x^2 + 1.76 \times 10^{-5}\,x - 1.76 \times 10^{-6} = 0$$

This equation is a typical quadratic equation of the form $ax^2 + bx + c = 0$, where $a = 1$, $b = 1.76 \times 10^{-5}$, and $c = -1.76 \times 10^{-6}$. Solve for x using the standard formula

$$x = \frac{-b \pm \sqrt{(b^2 - 4ac)}}{2a}$$

$$= \frac{-1.76 \times 10^{-5} \pm \sqrt{((1.76 \times 10^{-5})^2 - 4(1.76 \times 10^{-6}))}}{2}$$

$$x = 0.00132 \quad \text{or} \quad -0.00135 \text{ (reject the negative answer)}$$

The hydrogen ion concentration is 0.00132 M and the pH is

$$pH = -\log[H^{\oplus}] = -\log(0.00132) = -(-2.88) = 2.9$$

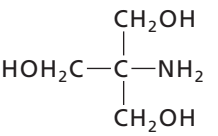
Note that the contribution of hydrogen ions from the dissociation of water $1 10^{-7} 2$ is several orders of magnitude lower than the concentration of hydrogen ions from acetic acid. It is standard practice to ignore the ionization of water in most calculations as long as the initial concentration of weak acid is greater than 0.001 M.

The amount of acetic acid that dissociates to form $H^{\oplus}$ and $CH_3COO^{\ominus}$ is 0.0013 M when the initial concentration is 0.1 M. This means that only 1.3% of the acetic acid molecules dissociate and the final concentration of acetic acid $1[CH_3COOH]2$ is 98.7% of the initial concentration. In general, the percent dissociation of dilute solutions of weak acids is less than 10% and it is a reasonable approximation to assume that the final concentration of the acid form is the same as its initial concentration. This approximation has very little effect on the calculated pH and it has the advantage of avoiding quadratic equations.

Assuming that the concentration of CH_3COOH at equilibrium is 0.1 M and the concentration of $H^{\oplus}$ is x,

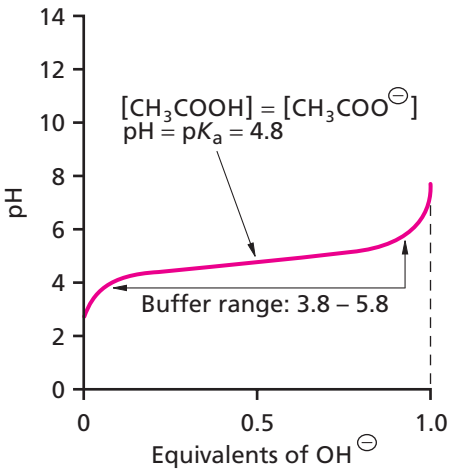
$$K_a = 1.76 \times 10^{-5} = \frac{x^2}{0.1} \qquad x = 1.33 \times 10^{-3}$$

$$pH = -\log(1.33 \times 10^{-3}) = 2.88 = 2.9$$



▲ **Tris buffers.** Tris, or tris (hydroxymethyl) aminomethane, is a common buffer in biochemistry labs. Its pK_a of 8.06 makes it ideal for preparation of buffers in the physiological range.

Downloaded From Vet-e-Books.com



▲ **Figure 2.20**
Buffer range of acetic acid. For CH_3COOH + $CH_3COO^{\ominus}$ the pK_a is 4.8 and the most effective buffer range is from pH 3.8 to pH 5.8.

2.10 Buffered Solutions Resist Changes in pH

If the pH of a solution remains nearly constant when small amounts of strong acid or strong base are added the solution is said to be **buffered**. The ability of a solution to resist changes in pH is known as its buffer capacity. Inspection of the titration curves of acetic acid (Figure 2.17) and phosphoric acid (Figure 2.19) reveals that the most effective buffering, indicated by the region of minimum slope on the curve, occurs when the concentrations of a weak acid and its conjugate base are equal—in other words, when the pH equals the pK_a. The effective range of buffering by a mixture of a weak acid and its conjugate base is usually considered to be from one pH unit below to one pH unit above the pK_a.

Most *in vitro* biochemical experiments involving purified molecules, cell extracts, or intact cells are performed in the presence of a suitable buffer to ensure a stable pH. A number of synthetic compounds with a variety of pK_a values are often used to prepare buffered solutions but naturally occurring compounds can also be used as buffers. For example, mixtures of acetic acid and sodium acetate (pK_a= 4.8) can be used for the pH range from 4 to 6 (Figure 2.20) and mixtures of KH_2PO_4 and K_2HPO_4 (pK_a = 7.2) can be used in the range from 6 to 8. The amino acid glycine (pK_a = 9.8) is often used in the range from 9 to 11.

When preparing buffers the acid solution (e.g., acetic acid) supplies the protons and some of the protons are taken up by combining with the conjugate base (e.g., acetate). The conjugate base is added as a solution of a salt (e.g., sodium acetate). The salt dissociates completely in solution providing free conjugate base and no protons. Sample Calculation 2.2 illustrates one way to prepare a buffer solution.

Sample Calculation 2.2 BUFFER PREPARATION

Q: Acetic acid has a pK_a of 4.8. How many milliliters of 0.1 M acetic acid and 0.1 M sodium acetate are required to prepare 1 liter of 0.1 M buffer solution having a pH of 5.8?

A: Substitute the values for the pK_a and the desired pH into the Henderson–Hasselbalch equation (Equation 2.20).

$$5.8 = 4.8 + \log\frac{[\text{Acetate}]}{[\text{Acetic acid}]}$$

Solve for the ratio of acetate to acetic acid.

$$\log\frac{[\text{Acetate}]}{[\text{Acetic acid}]} = 5.8 - 4.8 = 1.0$$

$$[\text{Acetate}] = 10\,[\text{Acetic acid}]$$

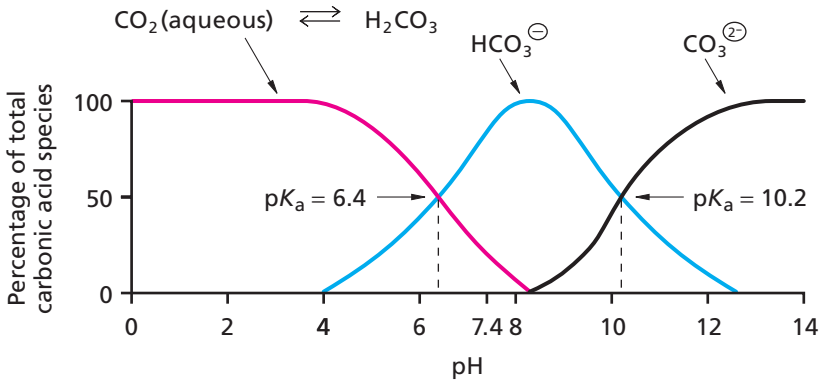
For each volume of acetic acid, 10 volumes of acetate must be added (making a total of 11 volumes of the two ionic species). Multiply the proportion of each component by the desired volume.

$$\text{Acetic acid needed:} \quad \frac{1}{11} \times 1000 \text{ ml} = 91 \text{ ml}$$

$$\text{Acetate needed:} \quad \frac{10}{11} \times 1000 \text{ ml} = 909 \text{ ml}$$

Note that when the ratio of [conjugate base] to [conjugate acid] is 10:1, the pH is exactly one unit above the pK_a. If the ratio were. 1:10, the pH would be one unit below the pK_a.

Downloaded From Vet-eBooks.com



▶ **Figure 2.21**
Percentages of carbonic acid and its conjugate base as a function of pH. In an aqueous solution at pH 7.4 (the pH of blood) the concentrations of carbonic acid (H_2CO_3) and bicarbonate ($HCO_3^{\ominus}$) are substantial, but the concentration of carbonate ($CO_3^{2\ominus}$) is negligible.

An excellent example of buffer capacity is found in the blood plasma of mammals, which has a remarkably constant pH. Consider the results of an experiment that compares the addition of an aliquot of strong acid to a volume of blood plasma with a similar addition of strong acid to either physiological saline (0.15 M NaCl) or water. When 1 milliliter of 10 M HCl (hydrochloric acid) is added to 1 liter of physiological saline or water that is initially at pH 7.0 the pH is lowered to 2.0 (in other words, $[H^{\oplus}]$ from HCl is diluted to 10^{-2} M). However, when 1 milliliter of 10 M HCl is added to 1 liter of human blood plasma at pH 7.4 the pH is lowered to only 7.2—impressive evidence for the effectiveness of physiological buffering.

The pH of blood is primarily regulated by the carbon dioxide–carbonic acid–bicarbonate buffer system. A plot of the percentages of carbonic acid (H_2CO_3) and its conjugate base as a function of pH is shown in Figure 2.21. Note that the major components at pH 7.4 are carbonic acid and the bicarbonate anion ($HCO_3^{\ominus}$).

The buffer capacity of blood depends on equilibria between gaseous carbon dioxide (which is present in the air spaces of the lungs), aqueous carbon dioxide (which is produced by respiring tissues and dissolved in blood), carbonic acid, and bicarbonate. As shown in Figure 2.21, the equilibrium between bicarbonate and its conjugate base, carbonate ($CO_3^{2\ominus}$), does not contribute significantly to the buffer capacity of blood because the pK_a of bicarbonate is 10.2—too far from physiological pH to have an effect on the buffering of blood.

The first of the three relevant equilibria of the carbon dioxide–carbonic acid–bicarbonate buffer system is the dissociation of carbonic acid to bicarbonate.

$$H_2CO_3 \rightleftharpoons H^{\oplus} + HCO_3^{\ominus} \qquad \textbf{(2.22)}$$

This equilibrium is affected by a second equilibrium in which dissolved carbon dioxide is in equilibrium with its hydrated form, carbonic acid.

$$CO_2(\text{aqueous}) + H_2O \rightleftharpoons H_2CO_3 \qquad \textbf{(2.23)}$$

These two reactions can be combined into a single equilibrium reaction where the acid is represented as CO_2 dissolved in water:

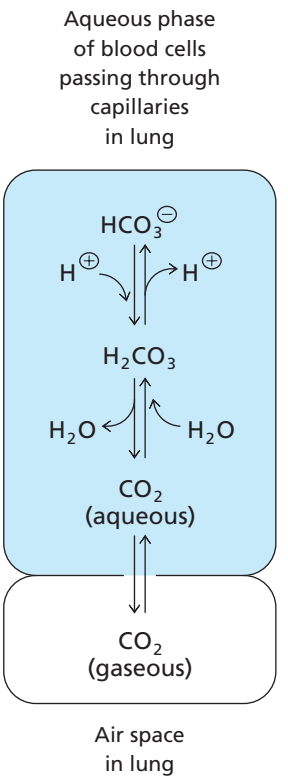
$$CO_2(\text{aqueous}) + H_2O \rightleftharpoons H^{\oplus} + HCO_3^{\ominus} \qquad \textbf{(2.24)}$$

The pK_a of the acid is 6.4.

Finally, CO_2 (gaseous) is in equilibrium with CO_2 (aqueous).

$$CO_2(\text{gaseous}) \rightleftharpoons CO_2(\text{aqueous}) \qquad \textbf{(2.25)}$$

The regulation of the pH of blood afforded by these three equilibria is shown schematically in Figure 2.22. When the pH of blood falls due to a metabolic process that produces excess $H^{\oplus}$ the concentration of H_2CO_3 increases momentarily but H_2CO_3



▲ **Figure 2.22**
Regulation of the pH of blood in mammals. The pH of blood is controlled by the ratio of $[HCO_3^{\ominus}]$ to pCO_2 in the air spaces of the lungs. When the pH of blood decreases due to excess $H^{\oplus}$, pCO_2 increases in the lungs, restoring the equilibrium. When the concentration of $HCO_3^{\ominus}$ rises because the pH of blood increases, CO_2 (gaseous) dissolves in the blood, again restoring the equilibrium.

Downloaded From Vet-e-Books.com

rapidly loses water to form dissolved CO_2 (aqueous) which enters the gaseous phase in the lungs and is expired as CO_2 (gaseous). An increase in the partial pressure of CO_2 (pCO_2) in the air expired from the lungs thus compensates for the increased hydrogen ions. Conversely, when the pH of the blood rises the concentration of $HCO_3^{\ominus}$ increases transiently but the pH is rapidly restored as the breathing rate changes and the CO_2 (gaseous) in the lungs is converted to CO_2 (aqueous) and then to H_2CO_3 in the capillaries of the lungs. Again, the equilibrium of the blood buffer system is rapidly restored by changing the partial pressure of CO_2 in the lungs.

Within cells, both proteins and inorganic phosphate contribute to intracellular buffering. Hemoglobin is the strongest buffer in blood cells other than the carbon dioxide–carbonic acid–bicarbonate buffer. As mentioned earlier, the major species of inorganic phosphate present at physiological pH are $H_2PO_4^{\ominus}$ and $HPO_4^{2\ominus}$ reflecting the second pK_a (pK_2) value for phosphoric acid, 7.2.

Summary

1. The water molecule has a permanent dipole because of the uneven distribution of charge in O—H bonds and their angled arrangement.
2. Water molecules can form hydrogen bonds with each other. Hydrogen bonding contributes to the high specific heat and heat of vaporization of water.
3. Because it is polar, water can dissolve ions. Water molecules form a solvation sphere around each dissolved ion. Organic molecules may be soluble in water if they contain ionic or polar functional groups that can form hydrogen bonds with water molecules.
4. The hydrophobic effect is the exclusion of nonpolar substances by water molecules. Detergents, which contain both hydrophobic and hydrophilic portions, form micelles when suspended in water; these micelles can trap insoluble substances in a hydrophobic interior. Chaotropes enhance the solubility of nonpolar compounds in water.
5. The major noncovalent interactions that determine the structure and function of biomolecules are electrostatic interactions and hydrophobic interactions. Electrostatic interactions include charge–charge interactions, hydrogen bonds, and van der Waals forces.
6. Under cellular conditions, macromolecules do not spontaneously hydrolyze, despite the presence of high concentrations of water. Specific enzymes catalyze their hydrolysis, and other enzymes catalyze their energy-requiring biosynthesis.
7. At 25°C, the product of the proton concentration ($[H^{\oplus}]$) and the hydroxide concentration ($[OH^{\ominus}]$) is 1.0×10^{-14} M^2, a constant designated K_w (the ion-product constant for water). Pure water ionizes to produce 10^{-7} M $H^{\oplus}$ and 10^{-7} M $OH^{\ominus}$.
8. The acidity or basicity of an aqueous solution depends on the concentration of $H^{\oplus}$ and is described by a pH value, where pH is the negative logarithm of the hydrogen ion concentration.
9. The strength of a weak acid is indicated by its pK_a value. The Henderson–Hasselbalch equation defines the pH of a solution of weak acid in terms of the pK_a and the concentrations of the weak acid and its conjugate base.
10. Buffered solutions resist changes in pH. In human blood, a constant pH of 7.4 is maintained by the carbon dioxide–carbonic acid–bicarbonate buffer system.

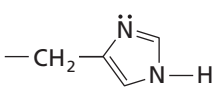
Problems

1. The side chains of some amino acids possess functional groups that readily form hydrogen bonds in aqueous solution. Draw the hydrogen bonds likely to form between water and the following amino acid side chains:

   (a) CH_2OH

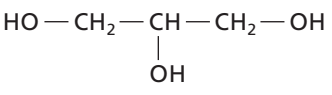
   (b) $CH_2C(O)NH_2$

   (c)

2. State whether each of the following compounds is polar, whether it is amphipathic, and whether it readily dissolves in water.

   (a)

   Glycerol

   (b) $CH_3(CH_2)_{14}—CH_2—OPO_3^{2\ominus}$

   Hexadecanyl phosphate

   (c) $CH_3—(CH_2)_{10}—COO^{\ominus}$

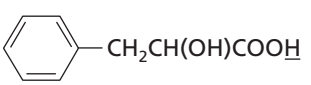
   Laurate

   (d) $H_3\overset{\oplus}{N}—CH_2—COO^{\ominus}$

   Glycine

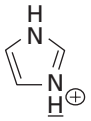
3. Osmotic lysis is a gentle method of breaking open animal cells to free intracellular proteins. In this technique, cells are suspended in a solution that has a total molar concentration of solutes much less than that found naturally inside cells. Explain why this technique might cause cells to burst.
4. Each of the following molecules is dissolved in buffered solutions of: (a) pH = 2 and (b) pH = 11. For each molecule, indicate the solution in which the charged species will predominate. (Assume that the added molecules do not appreciably change the pH of the solution.)

   (a) Phenyl lactic acid $pK_a = 4$

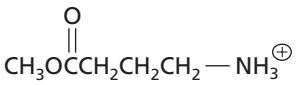
   

Downloaded From Vet-e-Books.com

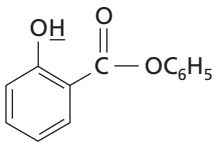
(b) Imidazole pK_a = 7



(c) *O*-methyl-γ-aminobutyrate pK_a = 9.5



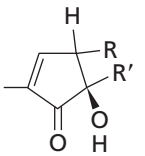
(d) Phenyl salicylate pK_a = 9.6



5. Use Figure 2.16 to determine the concentration of $H^{\oplus}$ and $OH^{\ominus}$ in:

(a) tomato juice
(b) human blood plasma
(c) 1 M ammonia

6. The interaction between two (or more) molecules in solution can be mediated by specific hydrogen bond interactions. Phorbol esters can act as a tumor promoter by binding to certain amino acids that are part of the enzyme protein kinase C (PKC). Draw the hydrogen bonds expected in the complex formed between the tumor promoter phorbol and the glycine portion of PKC: —$NHCH_2C(O)$—

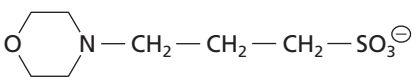


Phorbol

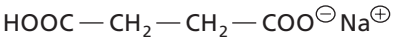
7. What is the concentration of a lactic acid buffer (pK_a = 3.9) that contains 0.25 M $CH_3CH(OH)COOH$ and 0.15 M $CH_3CH(OH)COO^{\ominus}$? What is the pH of this buffer?

8. You are instructed to prepare 100 ml of a 0.02 M sodium phosphate buffer, pH 7.2, by mixing 50 ml of solution A (0.02M Na_2HPO_4) and 50 ml of solution B (0.02 M NaH_2PO_4). Refer to Table 2.4 to explain why this procedure provides an effective buffer at the desired pH and concentration.

9. What are the effective buffering ranges of MOPS (3-(*N*-morpholino)propanesulfonic acid) and SHS (sodium hydrogen succinate)?



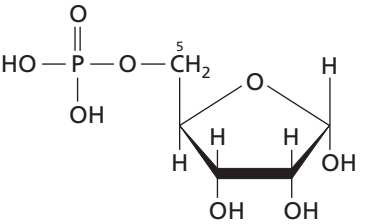
MOPS



SHS

The nitrogen atom of MOPS can be protonated (pK_a = 7.2). The carboxyl group of SHS can be ionized (pK_a = 5.5). Calculate the ratio of basic to acidic species for each buffer at pH 6.5.

10. Many phosphorylated sugars (phosphate esters of sugars) are metabolic intermediates. The two ionizable —OH groups of the phosphate group of the monophosphate ester of ribose (ribose 5-phosphate) have pK_a values of 1.2 and 6.6. The fully protonated form of α-D-ribose 5-phosphate has the structure shown below.

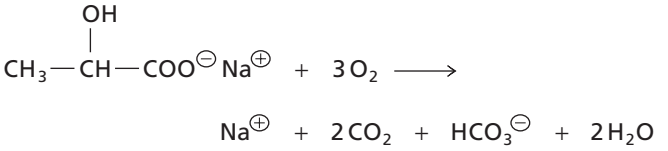


α-D-Ribose 5-phosphate

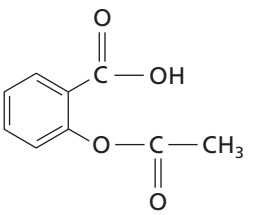
(a) Draw, in order, the ionic species formed upon titration of this phosphorylated sugar from pH 0.0 to pH 10.0.
(b) Sketch the titration curve for ribose 5-phosphate.

11. Normally, gaseous CO_2 is efficiently expired in the lungs. Under certain conditions, such as obstructive lung disease or emphysema, expiration is impaired. The resulting excess of CO_2 in the body may lead to respiratory acidosis, a condition in which excess acid accumulates in bodily fluids. How does excess CO_2 lead to respiratory acidosis?

12. Organic compounds in the diets of animals are a source of basic ions and may help combat nonrespiratory types of acidosis. Many fruits and vegetables contain salts of organic acids that can be metabolized, as shown below for sodium lactate. Explain how the salts of dietary acids may help alleviate metabolic acidosis.



13. Absorption of food in the stomach and intestine depends on the ability of molecules to penetrate the cell membranes and pass into the bloodstream. Because hydrophobic molecules are more likely to be absorbed than hydrophilic or charged molecules, the absorption of orally administered drugs may depend on their pK_a values and the pH in the digestive organs. Aspirin (acetylsalicylic acid) has an ionizable carboxyl group (pK_a = 3.5). Calculate the percentage of the protonated form of aspirin available for absorption in the stomach (pH = 2.0) and in the intestine (pH = 5.0).



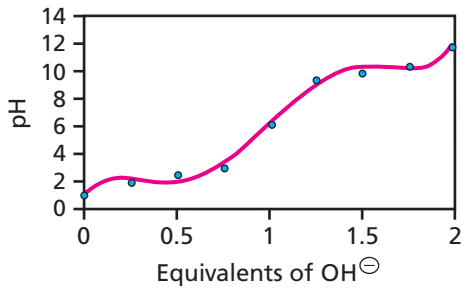
Aspirin

14. What percent of glycinamide, $^{\oplus}H_3NCH_2CONH_2$ (pK_a = 8.20) is unprotonated at (a) pH 7.5, (b) pH 8.2, and (c) pH 9.0?

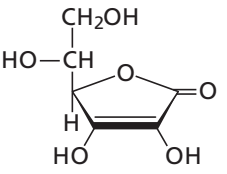
Downloaded From Vet-e-Books.com

15. Refer to the following table and titration curve to determine which compound from the table is illustrated by the titration curve.

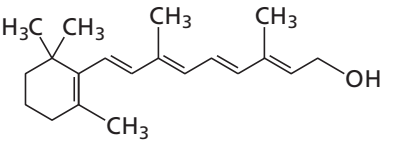
Compound	pK_1	pK_2	pK_3
Phosphoric acid	2.15	7.20	12.15
Acetic acid	4.76		
Succinic acid	4.21	5.64	
Boric acid	9.24	12.74	
Glycine	2.40	9.80	



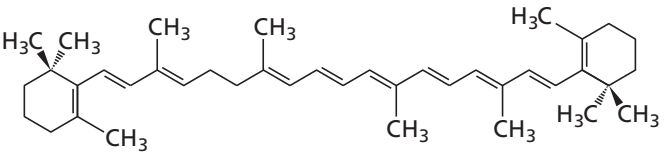
16. Predict which of the following substances are soluble in water.



(a) Vitamin C



(b) Vitamin A



(c) β-carotene

17. The ion product for water at 0°C is 1.14×10^{-15}, and at 100°C it is about 4.0×10^{-13}. What is the actual neutral pH for extremophiles living at 0°C and 100°C?

18. What is the approximate pH of a solution of 6 M HCl? Why doesn't the scale in Figure 2.16 accommodate the pH of such a solution?

Selected Readings

Water

Chaplin, M. F. (2001). Water, its importance to life. *Biochem. and Mol. Biol. Education* 29:54–59.

Dix, J. A. and Verkman, A. S. (2008). Crowding effects on diffusion in solutions and cells. *Annu. Rev. Biophys.* 37:247–263.

Stillinger, F. H. (1980). Water revisited. *Science* 209:451–457.

Verkman, A. S. (2001). Solute and macromolecular diffusion in cellular aqueous compartments. *Trends Biochem Sci.* 27:27–33.

Noncovalent Interactions

Fersht, A. R. (1987). The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 12:301–304.

Frieden, E. (1975). Non-covalent interactions. *J. Chem. Educ.* 52:754–761.

Tanford, C. (1980). *The Hydrophobic Effect: Formation of Micelles and Biological Membranes,* 2nd ed. (New York: John Wiley & Sons).

Biochemical Calculations

Montgomery, R., and Swenson, C. A. (1976). *Quantitative Problems in Biochemical Sciences,* 2nd ed. (San Francisco: W. H. Freeman).

Segel, I. H. (1976). *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd ed. (New York: John Wiley & Sons).

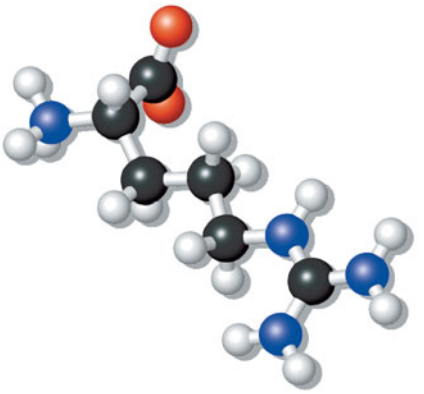
pH and Buffers

Stoll, V. S., and Blanchard, J. S. (1990). Buffers: principles and practice. *Methods Enzymol.* 182:24–38.

Nørby, J. G. (2000). The origin and meaning of the little p in pH. *Trends Biochem. Sci.* 25:36–37.

Downloaded From Vet-e-Books.com





Amino Acids and the Primary Structures of Proteins

The relationship between structure and function is a fundamental part of biochemistry. In spite of its importance, we sometimes forget to mention structure-function relationships, thinking that the concept is obvious from the examples. In this book we will try and remind you from time to time how the study of structure leads to a better understanding of function. This is especially important when studying proteins.

In this chapter and the next one we will cover the basic rules of protein structure. In Chapters 5 and 6, we will learn how enzymes work and how their structure contributes to the mechanisms of enzyme action.

Before beginning, let's review the various kinds of proteins. The following list, although not exhaustive, covers most of the important biological functions of proteins:

1. Many proteins function as enzymes, the biochemical catalysts. Enzymes catalyze nearly all reactions that occur in living organisms.
2. Some proteins bind other molecules for storage and transport. For example, hemoglobin binds and transports O_2 and CO_2 in red blood cells and other proteins bind fatty acids and lipids.
3. Several types of proteins serve as pores and channels in membranes, allowing for the passage of small, charged molecules.
4. Some proteins, such as tubulin, actin, and collagen, provide support and shape to cells and hence to tissues and organisms.
5. Assemblies of proteins can do mechanical work, such as the movement of flagella, the separation of chromosomes at mitosis, and the contraction of muscles.
6. Many proteins play a role in information flow in the cell. Some are involved in translation whereas others play a role in regulating gene expression by binding to nucleic acids.
7. Some proteins are hormones, which regulate biochemical activities in target cells or tissues; other proteins serve as receptors for hormones.

"Amino acids are literally raining down from the sky, and if that's not a big deal then I don't know what is."

Max Bernstein,
SETI Institute

KEY CONCEPT

The functions of biochemical molecules can only be understood by knowing their structures.

Top: L-Arginine, one of the 20 common amino acids.

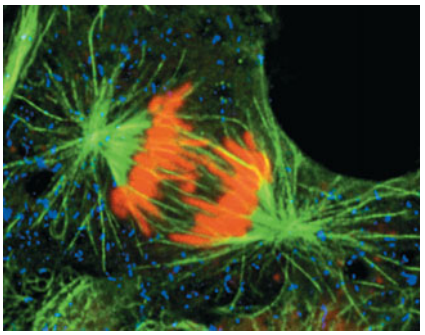
Downloaded From Vet-eBooks.com

KEY CONCEPT

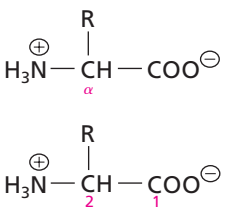
There are many different kinds of proteins with many different roles in metabolism and cell structure.

8. Proteins on the cell surface can act as receptors for various ligands and as modifiers of cell-cell interactions.
9. Some proteins have highly specialized functions. For example, antibodies defend vertebrates against bacterial and viral infections, and toxins, produced by bacteria, can kill larger organisms.

We begin our study of proteins by exploring the structures and chemical properties of their constituent amino acids. In this chapter we will also discuss the purification, analysis, and sequencing of polypeptides.



▲ **Spindle fibers.** Spindle fibers (green) help separate chromosomes at mitosis. The fibers are microtubules formed from the structural protein tubulin.



▲ **Numbering conventions for amino acids.** In traditional names, the carbon atoms adjacent to the carboxyl group are identified by the Greek letters α, β, γ, etc. In the official IUPAC/IUBMB chemical names or systematic names, the carbon atom in the carboxyl group is number 1 and the adjacent carbons are numbered sequentially. Thus, the α-carbon atom in traditional names is the carbon 2 atom in systematic names.

The IUPAC-IUBMB website for *Nomenclature and Symbolism for Amino Acids and Peptides* is: www.chem.qmul.ac.uk/iupac/AminoAcid/.

3.1 General Structure of Amino Acids

All organisms use the same 20 amino acids as building blocks for the assembly of protein molecules. These 20 amino acids are called the *common*, or *standard*, amino acids. Despite the limited number of amino acids, an enormous variety of different polypeptides can be produced by connecting the 20 common amino acids in various combinations.

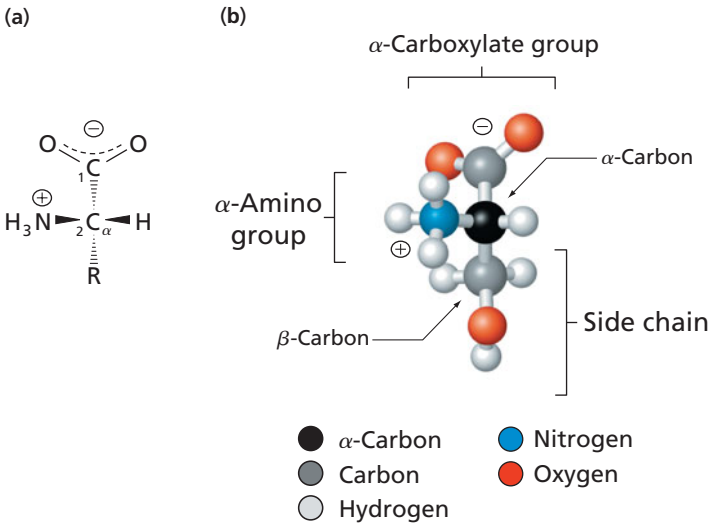
Amino acids are called amino acids because they are amino derivatives of carboxylic acids. In the 20 common amino acids the amino group and the carboxyl group are bonded to the same carbon atom: the α-carbon atom. Thus, all of the standard amino acids found in proteins are α-amino acids. Two other substituents are bound to the α-carbon—a hydrogen atom and a side chain (R) that is distinctive for each amino acid. In the chemical names of amino acids, carbon atoms are identified by numbers, beginning with the carbon atom of the carboxyl group. [The correct chemical name, or systematic name, follows rules established by the International Union of Pure and Applied Chemistry (IUPAC) and the International Union of Biochemistry and Molecular Biology (IUBMB).] If the R group is —CH_3 then the *systematic name* for that amino acid would be 2-aminopropanoic acid. (Propanoic acid is CH_3—CH_2—COOH.) The *trivial name* for CH_3—$CH(NH_2)$—COOH is *alanine*. The old nomenclature uses Greek letters to identify the α-carbon atom and the carbon atoms of the side chain. This nomenclature identifies the carbon atom relative to the carboxyl group so the carbon atom of the carboxyl group is not specified, unlike in the systematic nomenclature, where this carbon atom is number 1 in the numbering system. Biochemists have traditionally used the old, alternate nomenclature.

Inside a cell, under normal physiological conditions, the amino group is protonated (—$NH_3^{\oplus}$) because the pK_a of this group is close to 9. The carboxyl group is ionized (—$COO^{\ominus}$) because the pK_a of that group is below 3, as we saw in Section 2.9. Thus, in the physiological pH range of 6.8 to 7.4, amino acids are **zwitterions**, or dipolar ions, even though their net charge may be zero. We will see in Section 3.4 that some side chains can also ionize. Biochemists always represent the structures of amino acids in the form that is biologically relevant which is why you will see the zwitterions in the following figures.

Figure 3.1a shows the general three-dimensional structure of an amino acid. Figure 3.1b shows a ball-and-stick model of a representative amino acid, serine, whose side chain is —CH_2OH. The first carbon atom that's directly bound to the carboxylate carbon is the α-carbon so the other carbon atoms of a side chain are sequentially labeled β, γ, δ, and ε, referring to carbons 3, 4, 5, and 6, respectively, in the newer convention. The systematic name for serine is 2-amino-3-hydroxypropanoic acid.

In 19 of the 20 common amino acids the α-carbon atom is **chiral**, or asymmetric, since it has four different groups bonded to it. The exception is glycine, whose R group is simply a hydrogen atom. The molecule is not chiral because the α-carbon atom is bonded to two identical hydrogen atoms. The 19 chiral amino acids can therefore exist as stereoisomers. **Stereoisomers** are compounds that have the same molecular formula but differ in the arrangement, or **configuration**, of their atoms in space. The two stereoisomers are distinct molecules that can't be easily converted from one form to the other since a change in configuration requires the breaking of one or more bonds. Amino acid stereoisomers are nonsuperimposable mirror images called **enantiomers**. Two of the 19 chiral amino acids, isoleucine and threonine, have two chiral carbon atoms each. Isoleucine and threonine can each form four different stereoisomers.

Downloaded From Vet-e-Books.com



◀ **Figure 3.1**
Two representations of an L-amino acid at neutral pH. (a) General structure. An amino acid has a carboxylate group (whose carbon atom is designated C-1), an amino group, a hydrogen atom, and a side chain (or R group), all attached to C-2 (the α-carbon). Solid wedges indicate bonds above the plane of the paper; dashed wedges indicate bonds below the plane of the paper. The blunt ends of wedges are nearer the viewer than the pointed ends. **(b)** Ball-and-stick model of serine (whose R group is (—CH_2OH).

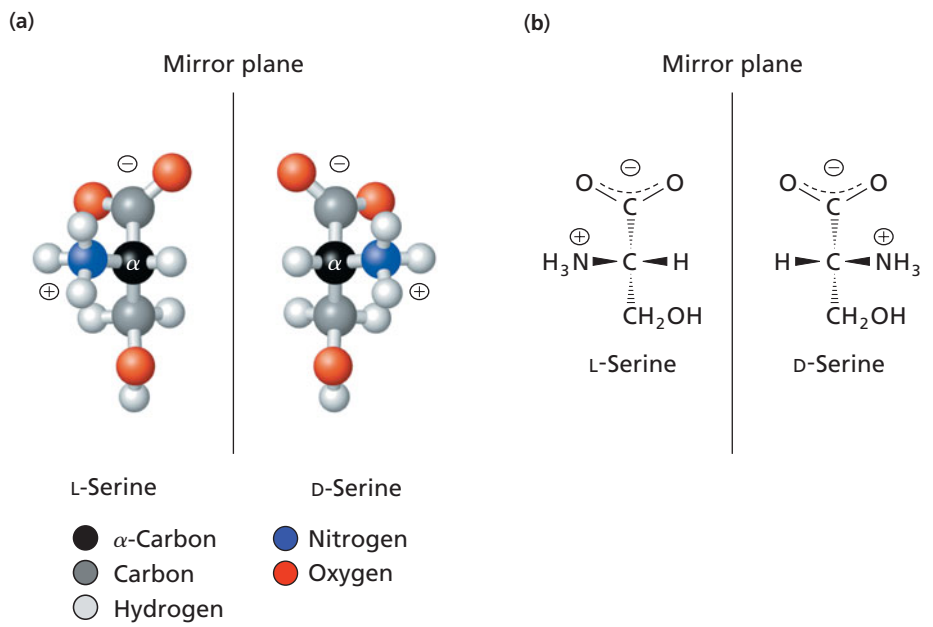
By convention, the mirror-image pairs of amino acids are designated D (for dextro, from the Latin *dexter*, "right") and L (for levo, from the Latin *laevus*, "left"). The configuration of the amino acid in Figure 3.1a is L and that of its mirror image is D. To assign the stereochemical designation, one draws the amino acid vertically with its α-carboxylate group at the top and its side chain at the bottom, both pointing away from the viewer. In this orientation, the α-amino group of the L isomer is on the left of the α-carbon, and that of the D isomer is on the right, as shown in Figure 3.2. (The four atoms attached to the α-carbon occupy the four corners of a tetrahedron much like the bonding of hydrogen atoms to oxygen in water, as shown in Figure 2.4.)

The 19 chiral amino acids used in the assembly of proteins are all of the L configuration, although a few D-amino acids occur in nature. By convention, amino acids are assumed to be in the L configuration unless specifically designated D. Often it is convenient to draw the structures of L-amino acids in a form that is stereochemically uncommitted, especially when a correct stereochemical representation is not critical to a given discussion.

The fact that all living organisms use the same standard amino acids in protein synthesis is evidence that all species on Earth are descended from a common ancestor. Like modern organisms, the last common ancestor (LCA) must have used L-amino



▲ **Meteorites and amino acids.** The Murchison meteorite fell in 1969 near Murchison, Australia. There are many similar carbonaceous meteorites and many of them contain spontaneously formed amino acids, including some of the common amino acids found in proteins. These amino acids are found in the meteorites as almost equal mixtures of the L and D configurations.



See Section 8.1 for a more complete description of the convention for displaying stereoisomers (Fischer projection).

◀ **Figure 3.2**
Mirror-image pairs of amino acids. (a) Ball-and-stick models of L-serine and D-serine. Note that the two molecules are not identical; they cannot be superimposed. **(b)** L-Serine and D-serine. The common amino acids all have the L configuration.

Downloaded From Vet-e-Books.com

acids and not D-amino acids. Mixtures of L- and D-amino acids are formed under conditions that mimic those present when life first arose on Earth 4 billion years ago and both enantiomers are found in meteorites and in the vicinity of stars. It is not known how or why primitive life forms selected L-amino acids from the presumed mixture of the enantiomers present when life first arose. It's likely that the first proteins were composed of a small number of simple amino acids and selection of L-amino acids over D-amino acids was a chance event. Modern living organisms do not select L-amino acids from a mixture because only the L-amino acids are synthesized in sufficient quantities. Thus, the predominance of L-amino acids in modern species is due to the evolution of metabolic pathways that produce L-amino acids and not D-amino acids (Chapter 17).

3.2 Structures of the 20 Common Amino Acids

The structures of the 20 amino acids commonly found in proteins are shown in the following figures as Fischer projections. In Fischer projections, horizontal bonds at a chiral center extend toward the viewer and vertical bonds extend away (as in Figures 3.1 and 3.2). Examination of the structures reveals considerable variation in the side chains of the 20 amino acids. Some side chains are nonpolar and thus hydrophobic whereas others are polar or ionized at neutral pH and are therefore hydrophilic. The properties of the side chains greatly influence the overall three-dimensional shape, or conformation, of a protein. For example, most of the hydrophobic side chains of a water-soluble protein fold into the interior giving the protein a compact, globular shape.

Some nonstandard amino acids are described in Section 3.3.

Both the three-letter and one-letter abbreviations for each amino acid are shown in the figures. The three-letter abbreviations are self-evident but the one-letter abbreviations are less obvious. Several amino acids begin with the same letter so other letters of the alphabet have to be used in order to provide a unique label; for example, threonine = T, tyrosine = Y, and tryptophan = W. These labels have to be memorized.

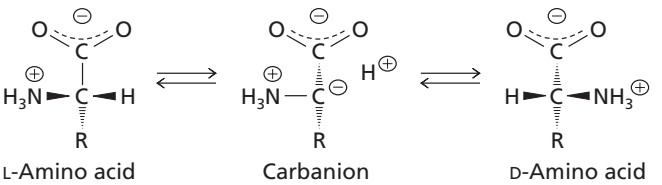
BOX 3.1 FOSSIL DATING BY AMINO ACID RACEMIZATION

Amino acids can spontaneously convert from the D configuration to the L configuration and *vice versa*. This is a chemical reaction that usually proceeds through a carbanion intermediate.

The racemization reaction is normally very slow but it can be sped up at high temperatures. For example, the half-life for conversion of L-aspartate to D-aspartate is about 30 days at 100°C. The half-life of this reaction at 37°C is about 350 years and at 18°C it's about 50,000 years.

The amino acid composition of mammalian tooth enamel can be used to determine the age of a fossil if the average temperature of the environment is known or can be estimated. When the amino acids are first synthesized they are exclusively of the L configuration. Over time, the amount of the D enantiomer increases and the D/L ratio can be measured very precisely.

Fossil dating by measuring amino acid racemization has been superceded by more reliable methods but it's an interesting example of a slow chemical reaction. Some organisms contain specific racemases that catalyze the interconversion of an L-amino acid and a D-amino acid; for example, bacteria have alanine racemase for converting L-alanine to D-alanine (see Section 8.7B). These enzymes catalyze thousands of reactions per second.





▲ The Badegoule Jaw from a stone age juvenile. *Homo sapiens* (Natural History Museum, Lyon, France)

Downloaded From Vet-eBooks.com

It is important to learn the structures of the standard amino acids because we refer to them frequently in the chapters on protein structure, enzymes, and protein synthesis. In the following sections we have grouped the standard amino acids by their general properties and the chemical structures of their side chains. The side chains fall into the following chemical classes: aliphatic, aromatic, sulfur-containing, alcohols, positively charged, negatively charged, and amides. Of the 20 amino acids five are further classified as highly hydrophobic (blue) and seven are classified as highly hydrophilic (red). Understanding the classification of the R groups will simplify memorizing the structures and names.

A. Aliphatic R Groups

Glycine (Gly, G) is the smallest amino acid. Since its R group is simply a hydrogen atom, the α-carbon of glycine is not chiral. The two hydrogen atoms of the α-carbon of glycine impart little hydrophobic character to the molecule. We will see that glycine plays a unique role in the structure of many proteins because its side chain is small enough to fit into niches that cannot accommodate any other amino acid.

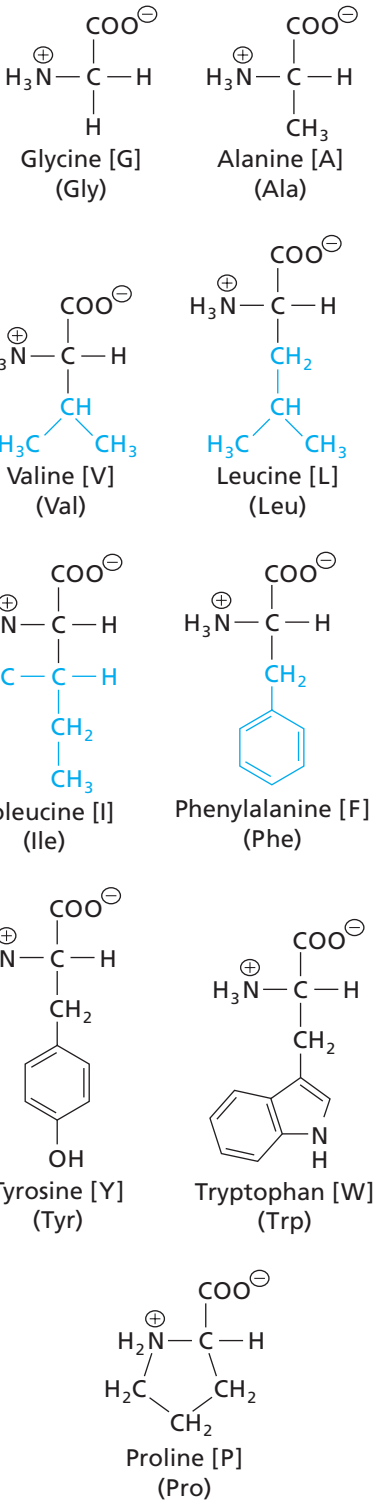
Four amino acids—alanine (Ala, A), valine (Val, V), leucine (Leu, L), and the structural isomer of leucine, isoleucine (Ile, I)—have saturated aliphatic side chains. The side chain of alanine is a methyl group whereas valine has a three-carbon branched side chain and leucine and isoleucine each contain a four-carbon branched side chain. Both the α- and β-carbon atoms of isoleucine are asymmetric. Because isoleucine has two chiral centers, it has four possible stereoisomers. The stereoisomer used in proteins is called L-isoleucine and the amino acid that differs at the β-carbon is called L-alloisoleucine (Figure 3.3). The other two stereoisomers are D-isoleucine and D-alloisoleucine.

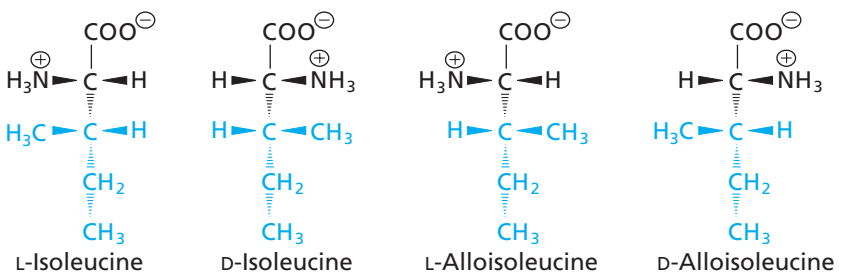
Alanine, valine, leucine, and isoleucine play an important role in establishing and maintaining the three-dimensional structures of proteins because of their tendency to cluster away from water. Valine, leucine, and isoleucine are known collectively as the branched chain amino acids because their side chains of carbon atoms contain branches. All three amino acids are highly hydrophobic and they share biosynthesis and degradation pathways (Chapter 17).

Proline (Pro, P) differs from the other 19 amino acids because its three-carbon side chain is bonded to the nitrogen of its α-amino group as well as to the α-carbon creating a cyclic molecule. As a result, proline contains a secondary rather than a primary amino group. The heterocyclic pyrrolidine ring of proline restricts the geometry of polypeptides sometimes introducing abrupt changes in the direction of the peptide chain. The cyclic structure of proline makes it much less hydrophobic than valine, leucine, and isoleucine.

B. Aromatic R Groups

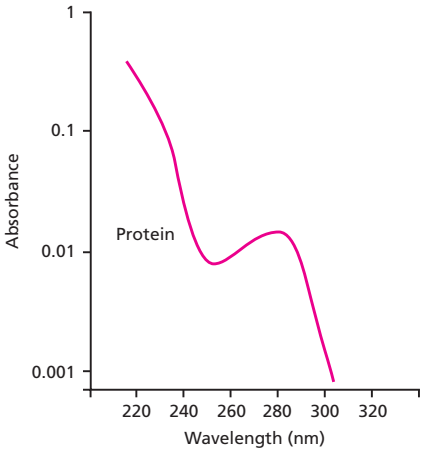
Phenylalanine (Phe, F), tyrosine (Tyr, Y), and tryptophan (Trp, W) have side chains with aromatic groups. Phenylalanine has a hydrophobic benzyl side chain. Tyrosine is structurally similar to phenylalanine except that the *para* hydrogen of phenylalanine is replaced in tyrosine by a hydroxyl group (—OH) making tyrosine a phenol. The hydroxyl group of tyrosine is ionizable but retains its hydrogen under normal physiological conditions. The side chain of tryptophan contains a bicyclic indole group. Tyrosine and



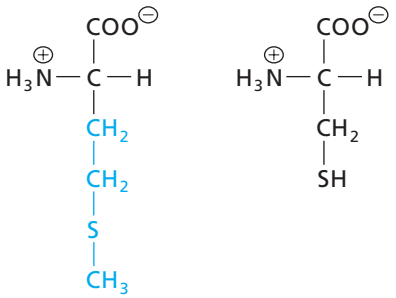


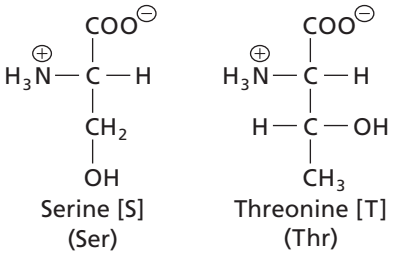
◀ **Figure 3.3**
Stereoisomers of isoleucine. Isoleucine and threonine are the only two common amino acids with more than one chiral center. The other DL pair of isoleucine isomers is called alloleucine. Note that in L-isoleucine the $—NH_3^{\oplus}$ and $—CH_3$ groups are both on the left in this projection, while in D-isoleucine they are both on the right, so that D-isoleucine and L-isoleucine are mirror images.

Downloaded From Vet-e-Books.com



▲ **UV absorbance of proteins.** The peak of absorbance of most proteins peaks at 280 nm. Most of the absorbance is due to the presence of tryptophan and tyrosine residues in the protein.







▲ **A sulfur bridge.** Natural stone bridge, Puente del Inca, in Mendoza, Argentina. Over the years the bridge has been covered with sulfur deposits.

tryptophan are not as hydrophobic as phenylalanine because their side chains include polar groups (Table 3.1, page 62).

All three aromatic amino acids absorb ultraviolet (UV) light because, unlike the saturated aliphatic amino acids, the aromatic amino acids contain delocalized π-electrons. At neutral pH both tryptophan and tyrosine absorb light at a wavelength of 280 nm whereas phenylalanine is almost transparent at 280 nm and absorbs light weakly at 260 nm. Since most proteins contain tryptophan and tyrosine they will absorb light at 280 nm. Absorbance at 280 nm is routinely used to estimate the concentration of proteins in solutions.

C. R Groups Containing Sulfur

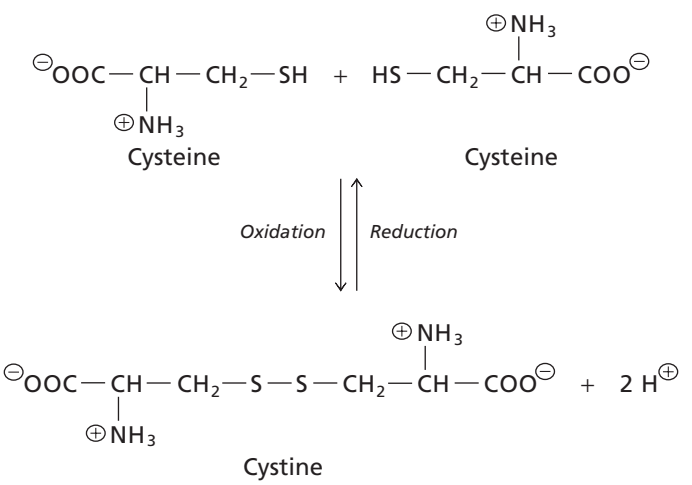
Methionine (Met, M) and cysteine (Cys, C) are the two amino acids whose side chains contain a sulfur atom. Methionine contains a nonpolar methyl thioether group in its side chain and this makes it one of the more hydrophobic amino acids. Methionine plays a special role in protein synthesis because it is almost always the first amino acid in a growing polypeptide chain. The structure of cysteine resembles that of alanine with a hydrogen atom replaced by a sulfhydryl group (—SH).

Although the side chain of cysteine is somewhat hydrophobic, it is also highly reactive. Because the sulfur atom is polarizable the sulfhydryl group of cysteine can form weak hydrogen bonds with oxygen and nitrogen. Moreover, the sulfhydryl group of cysteine residues in proteins can be a weak acid which allows it to lose its proton to become a negatively charged thiolate ion. (The pK_a of the sulfhydryl group of the free amino acid is 8.3 but this can range from 5-10 in proteins.)

A compound called cystine can be isolated when some proteins are hydrolyzed. Cystine is formed from two oxidized cysteine molecules linked by a disulfide bond (Figure 3.4). Oxidation of the sulfhydryl groups of cysteine molecules proceeds most readily at slightly alkaline pH values because the sulfhydryl groups are ionized at high pH. The two cysteine side chains must be adjacent in three-dimensional space in order to form a disulfide bond but they don't have to be close together in the amino acid sequence of the polypeptide chain. They may even be found in different polypeptide chains. Disulfide bonds, or disulfide bridges, may stabilize the three-dimensional structures of some proteins by covalently cross-linking cysteine residues in peptide chains. Most proteins do not contain disulfide bridges because conditions inside the cell do not favor oxidation; however, many secreted, or extracellular, proteins contain disulfide bridges.

D. Side Chains with Alcohol Groups

Serine (Ser, S) and threonine (Thr, T) have uncharged polar side chains containing β-hydroxyl groups. These alcohol groups give a hydrophilic character to the aliphatic



▲ **Figure 3.4**
Formation of cystine. When oxidation links the sulfhydryl groups of two cysteine molecules, the resulting compound is a disulfide called cystine.

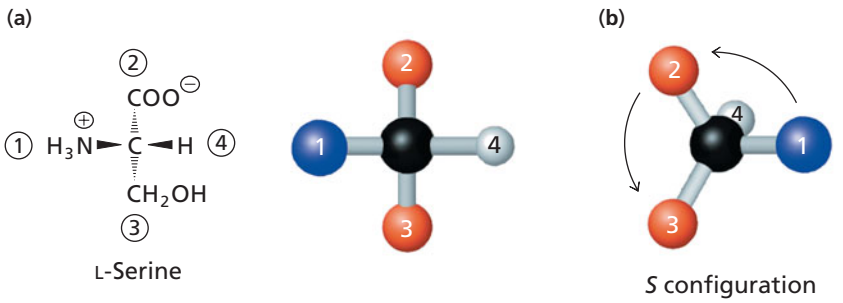
Downloaded From Vet-e-Books.com

BOX 3.2 AN ALTERNATIVE NOMENCLATURE

The *RS* system of configurational nomenclature is also sometimes used to describe the chiral centers of amino acids. The *RS* system is based on the assignment of a priority sequence to the four groups bound to a chiral carbon atom. Once assigned, the group priorities are used to establish the configuration of the molecule. Priorities are numbered 1 through 4 and are assigned to groups according to the following rules:

1. For atoms directly attached to the chiral carbon, the one with the lowest atomic mass is assigned the lowest priority (number 4).
2. If there are two identical atoms bound to the chiral carbon, the priority is decided by the atomic mass of the next atoms bound. For example, a $-CH_3$ group has a lower priority than a $-CH_2Br$ group because hydrogen has a lower atomic mass than bromine.
3. If an atom is bound by a double or triple bond, the atom is counted once for each formal bond. Thus, —CHO, with a double-bonded oxygen, has a higher priority than $-CH_2OH$. The order of priority for the most common groups, from lowest to highest, is $-H$, $-CH_3$, $-C_6H_5$, $-CH_2OH$, —CHO, —COOH, —COOR, $-NH_2$, —NHR, —OH, —OR, and —SH.

With these rules in mind, imagine the molecule as the steering wheel of a car, with the group of lowest priority (numbered 4) pointing away from you (like the steering column) and the other three groups arrayed around the rim of the steering wheel. Trace the rim of the wheel, moving from the group of highest priority to the group of lowest priority (1, 2, 3). If the movement is clockwise, the configuration is *R* (from the Latin *rectus*, "right-handed"). If the movement is counterclockwise, the configuration is *S* (from the Latin, *sinister*, "left-handed"). The figure demonstrates the assignment of *S* configuration to L-serine by the *RS* system. L-Cysteine has the opposite configuration, *R*. The DL system is used more often in biochemistry because not all amino acids found in proteins have the same *RS* designation.



◀ Assignment of configuration by the *RS* system. **(a)** Each group attached to a chiral carbon is assigned a priority based on atomic mass, 4 being the lowest priority. **(b)** By orienting the molecule with the priority 4 group pointing away (behind the chiral carbon) and tracing the path from the highest priority group to the lowest, the absolute configuration can be established. If the sequence 1, 2, 3 is clockwise, the configuration is *R*. If the sequence 1, 2, 3 is counterclockwise, the configuration is *S*. L-Serine has the *S* configuration.

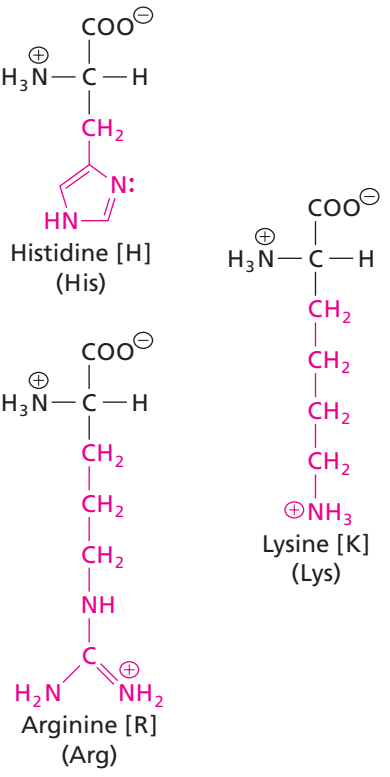
side chains. Unlike the more acidic phenolic side chain of tyrosine the hydroxyl groups of serine and threonine have the weak ionization properties of primary and secondary alcohols. The hydroxymethyl group of serine ($-CH_2OH$) does not appreciably ionize in aqueous solutions; nevertheless, this alcohol can react within the active sites of a number of enzymes as though it were ionized. Threonine, like isoleucine, has two chiral centers—the α- and β-carbon atoms. L-Threonine is the only one of the four stereoisomers that commonly occurs in proteins. (The other stereoisomers are called D-threonine, L-allothreonine, and D-allothreonine.)

E. Positively Charged R Groups

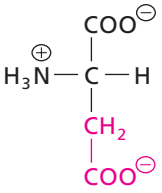
Histidine (His, H), lysine (Lys, K), and arginine (Arg, R) have hydrophilic side chains that are nitrogenous bases. The side chains can be positively charged at physiological pH.

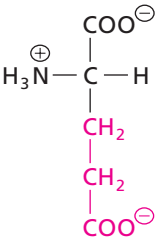
The side chain of histidine contains an imidazole ring substituent. The protonated form of this ring is called an imidazolium ion (Section 3.4). At pH 7 most histidines are neutral (base form) as shown in the accompanying figure but the form with a positively charged side chain is present and it becomes more common at slightly lower pH.

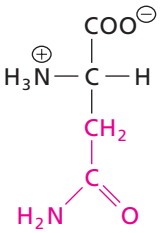
Lysine is a diamino acid with both α- and ε-amino groups. The ε-amino group exists as an alkylammonium ion ($-CH_2-NH_3^{\oplus}$) at neutral pH and confers a positive charge on proteins. Arginine is the most basic of the 20 amino acids because its



Downloaded From Vet-eBooks.com


Aspartate [D]
(Asp)


Glutamate [E]
(Glu)


Asparagine [N]
(Asn)

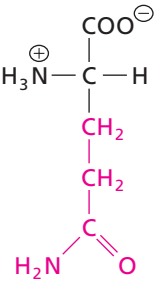

Glutamine [Q]
(Gln)

Table 3.1 Hydropathy scale

Amino acid	Free energy change of transfer[a] (kj mol^{-1})
Highly hydrophobic	
Isoleucine	3.1
Phenylalanine	2.5
Valine	2.3
Leucine	2.2
Methionine	1.1
Less hydrophobic	
Tryptophan	1.5[b]
Alanine	1.0
Glycine	0.67
Cysteine	0.17
Tyrosine	0.08
Proline	−0.29
Threonine	−0.75
Serine	−1.1
Highly hydrophilic	
Histidine	−1.7
Glutamate	−2.6
Asparagine	−2.7
Glutamine	−2.9
Aspartate	−3.0
Lysine	−4.6
Arginine	−7.5

[a]The free-energy change is for transfer of an amino acid residue from the interior of a lipid bilayer to water.

[b]On other scales, tryptophan has a lower hydropathy value.

[Adapted from Eisenberg, D., Weiss, R. M., Terwilliger, T. C., Wilcox, W. (1982). Hydrophobic moments in protein structure. *Faraday Symp. Chem. Soc.* 17:109–120.]

side-chain guanidinium ion is protonated under all conditions normally found within a cell. Arginine side chains also contribute positive charges in proteins.

F. Negatively Charged R Groups and Their Amide Derivatives

Aspartate (Asp, D) and glutamate (Glu, E) are dicarboxylic amino acids and have negatively charged hydrophilic side chains at pH 7. In addition to α-carboxyl groups, aspartate possesses a β-carboxyl group and glutamate possesses a γ-carboxyl group. Aspartate and glutamate confer negative charges on proteins because their side chains are ionized at pH 7. Aspartate and glutamate are sometimes called aspartic acid and glutamic acid but under most physiological conditions they are found as the conjugate bases and, like other carboxylates, have the suffix *-ate*. Glutamate is probably familiar as its monosodium salt, monosodium glutamate (MSG), which is used in food as a flavor enhancer.

Asparagine (Asn, N) and glutamine (Gln, Q) are the amides of aspartic acid and glutamic acid, respectively. Although the side chains of asparagine and glutamine are uncharged these amino acids are highly polar and are often found on the surfaces of proteins where they can interact with water molecules. The polar amide groups of asparagine and glutamine can also form hydrogen bonds with atoms in the side chains of other polar amino acids.

G. The Hydrophobicity of Amino Acid Side Chains

The various side chains of amino acids range from highly hydrophobic, through weakly polar, to highly hydrophilic. The relative hydrophobicity or hydrophilicity of each amino acid is called its **hydropathy**.

There are several ways of measuring hydropathy, but most of them rely on calculating the tendency of an amino acid to prefer a hydrophobic environment over a hydrophilic environment. A commonly used hydropathy scale is shown in Table 3.1. Amino acids with highly positive hydropathy values are considered hydrophobic whereas those with the largest negative values are hydrophilic. It is difficult to determine the hydropathy values of some amino acid residues that lie near the center of the scale. For example, there is disagreement over the hydropathy of the indole group of tryptophan and in some tables tryptophan has a much lower hydropathy value. Conversely, cysteine can have a higher hydropathy value in some tables.

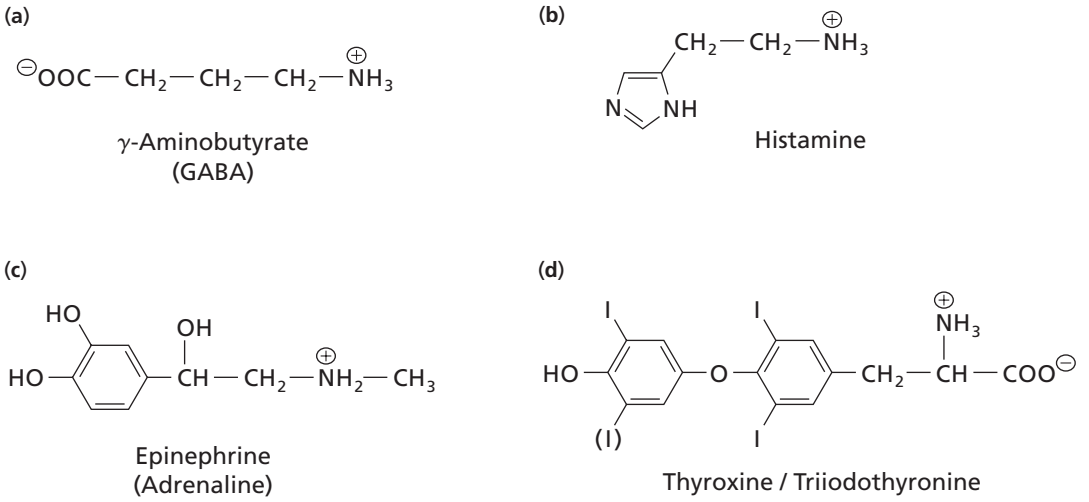
Hydropathy is an important determinant of protein folding because hydrophobic side chains tend to be clustered in the interior of a protein and hydrophilic residues are usually found on the surface (Section 4.10). However, it is not yet possible to predict accurately whether a given residue will be found in the nonaqueous interior of a protein or on the solvent-exposed surface. On the other hand, hydropathy measurements of free amino acids can be successfully used to predict which segments of membrane-spanning proteins are likely to be embedded in a hydrophobic lipid bilayer (Chapter 9).

3.3 Other Amino Acids and Amino Acid Derivatives

More than 200 different amino acids are found in living organisms. In addition to the 20 common amino acids covered in the previous section there are three others that are incorporated into proteins during protein synthesis. The 21st amino acid is *N*-formylmethionine which serves as the initial amino acid during protein synthesis in bacteria (Section 22.5). The 22nd amino acid is selenocysteine which contains selenium in place of the sulfur of cysteine. It is incorporated into a few proteins in almost every species. Selenocysteine is formed from serine during protein synthesis. The 23rd amino acid is pyrrolysine, found in some species of archaebacteria. Pyrrolysine is a modified form of lysine that is synthesized before being added to a growing polypeptide chain by the translation machinery.

N-formylmethionine, selenocysteine, and pyrrolysine are incorporated at specific codons and that's why they are considered additions to the standard repertoire of protein precursors. Because of post-translational modifications many complete proteins have more than the standard 23 amino acids used in protein synthesis (see below).

Downloaded From Vet-e-Books.com

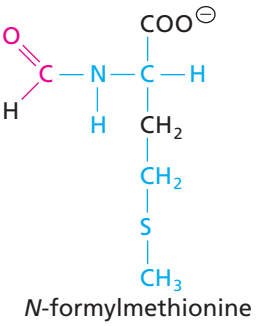


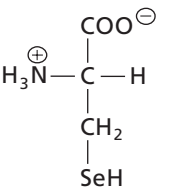
▲ **Figure 3.5**
Compounds derived from common amino acids. (a) γ-Aminobutyrate. a derivative of glutamate. **(b)** Histamine, a derivative of histidine. **(c)** Epinephrine, a derivative of tyrosine. **(d)** Thyroxine and triiodothyronine, derivatives of tyrosine. Thyroxine contains one more atom of iodine (in parentheses) than does triiodothyronine.

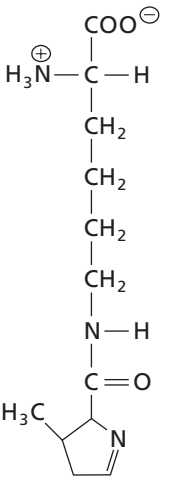
In addition to the common 23 amino acids that are incorporated into proteins, all species contain a variety of L-amino acids that are either precursors of the common amino acids or intermediates in other biochemical pathways. Examples are homocysteine, homoserine, ornithine, and citrulline (see Chapter 17). *S*-Adenosylmethionine (SAM) is a common methyl donor in many biochemical pathways (Section 7.2). Many species of bacteria and fungi synthesize D-amino acids that are used in cell walls and in complex peptide antibiotics such as actinomycin.

Several common amino acids are chemically modified to produce biologically important amines. These are synthesized by enzyme-catalyzed reactions that include decarboxylation and deamination. In the mammalian brain, for example, glutamate is converted to the neurotransmitter γ-aminobutyrate (GABA) (Figure 3.5a). Mammals can also synthesize histamine (Figure 3.5b) from histidine. Histamine controls the constriction of certain blood vessels and also the secretion of hydrochloric acid by the stomach. In the adrenal medulla, tyrosine is metabolized to epinephrine, also known as adrenaline (Figure 3.5c). Epinephrine and its precursor, norepinephrine (a compound whose amino group lacks a methyl substituent), are hormones that help regulate metabolism in mammals. Tyrosine is also the precursor of the thyroid hormones thyroxine and triiodothyronine (Figure 3.5d). Biosynthesis of the thyroid hormones requires iodide. Small amounts of sodium iodide are commonly added to table salt to prevent goiter, a condition of hypothyroidism caused by a lack of iodide in the diet.

Some amino acids are chemically modified after they have been incorporated into polypeptides. In fact, there are hundreds of known post-translational modifications. For example, some proline residues in the protein collagen are oxidized to form hydroxyproline residues (Section 4.11). Another common modification is the addition of complex carbohydrate chains—a process known as glycosylation (Chapters 8 and 22). Many proteins are phosphorylated, usually by the addition of phosphoryl groups to the side chains of serine, threonine, or tyrosine (histidine, lysine, cysteine, aspartate, and glutamate can also be phosphorylated). The oxidation of pairs of cysteine residues to form cystine also occurs after a polypeptide has been synthesized.


N-formylmethionine


Selenocysteine


Pyrrolysine

3.4 Ionization of Amino Acids

The physical properties of amino acids are influenced by the ionic states of the α-carboxyl and α-amino groups and of any ionizable groups in the side chains. Each ionizable group is associated with a specific pK_a value that corresponds to the pH at which the

Downloaded From Vet-e-Books.com

BOX 3.3 COMMON NAMES OF AMINO ACIDS

Alanine: probably from aldehyde + "an" (for convenience) + amine (1849)

Arginine: crystallizes as a silver salt, from Latin *argentum* (silver) (1886)

Asparagine: first isolated from asparagus (1813)

Aspartate: similar to asparagine (1836)

Glutamate: first identified in the plant protein gluten (1866)

Glutamine: similar to glutamate (1866)

Glycine: from the Greek *glykys* (sweet), tastes sweet (1848)

Cysteine: from the Greek *kystis* (bladder), discovered in bladder stones (1882)

Histidine: first isolated from sturgeon sperm, named for the Greek *histidin* (tissue) (1896)

Isoleucine: isomer of leucine

Leucine: from the Greek *leukos* (white), forms white crystals (1820)

Lysine: product of protein hydrolysis, from the Greek *lysis* (loosening) (1891)

Methionine: side chain is a sulfur (Greek *theion*) atom with a methyl group (1928)

Phenylalanine: alanine with a phenyl group (1883)

Proline: a corrupted form of "pyrrolidine" because it forms a pyrrolidine ring (1904)

Serine: from the Latin *sericum* (silk), serine is common in silk (1865)

Threonine: similar to the four-carbon sugar threose (1936)

Tryptophan: isolated from a tryptic digest of protein 1 Greek *phanein* (to appear) (1890)

Tyrosine: found in cheese, from the Greek *tyros* (cheese) (1890)

Valine: derivative of valeric acid from the plant genus *Valeriana* (1906)

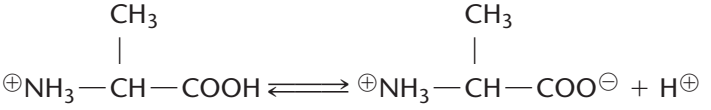
Sources: *Oxford English Dictionary* 2nd ed., and Leung, S.H. (2000) Amino acids, aromatic compounds, and carboxylic acids: how did they get their common names? *J. Chem. Educ.* 77: 48–49.

concentrations of the protonated and unprotonated forms are equal (Section 2.9). When the pH of the solution is below the pK_a the protonated form predominates and the amino acid is then a true acid that is capable of donating a proton. When the pH of the solution is above the pK_a of the ionizable group the unprotonated form of that group predominates and the amino acid exists as the conjugate base, which is a proton acceptor. Every amino acid has at least two pK_a values corresponding to the ionization of the α-carboxyl and α-amino groups. In addition, seven of the common amino acids have ionizable side chains with additional, measurable pK_a values. These values differ among the amino acids. Thus, at a given pH, amino acids frequently have different net charges. Many of the modified amino acids have additional ionizable groups contributing to the diversity of charged amino acid side chains in proteins. Phosphoserine and phosphotyrosine, for example, will be negatively charged.

Knowing the ionic states of amino acid side chains is important for two reasons. First, the charged state influences protein folding and the three-dimensional structure of proteins (Section 4.10). Second, an understanding of the ionic properties of amino acids in the active site of an enzyme helps one understand enzyme mechanisms (Chapter 6).

The pK_a values of amino acids are determined from titration curves such as those we saw in the previous chapter. The titration of alanine is shown in Figure 3.6. Alanine has two ionizable groups—the α-carboxyl and the protonated α-amino group. As more base is added to the solution of acid, the titration curve exhibits two pK_a values, at pH 2.4 and pH 9.9. Each pK_a value is associated with a buffering zone where the pH of the solution changes relatively little when more base is added.

The pK_a of an ionizable group corresponds to a midpoint of its titration curve. It is the pH at which the concentration of the acid form (proton donor) exactly equals the concentration of its conjugate base (proton acceptor). In the example shown in Figure 3.6 the concentrations of the positively charged form of alanine and of the zwitterion are equal at pH 2.4.

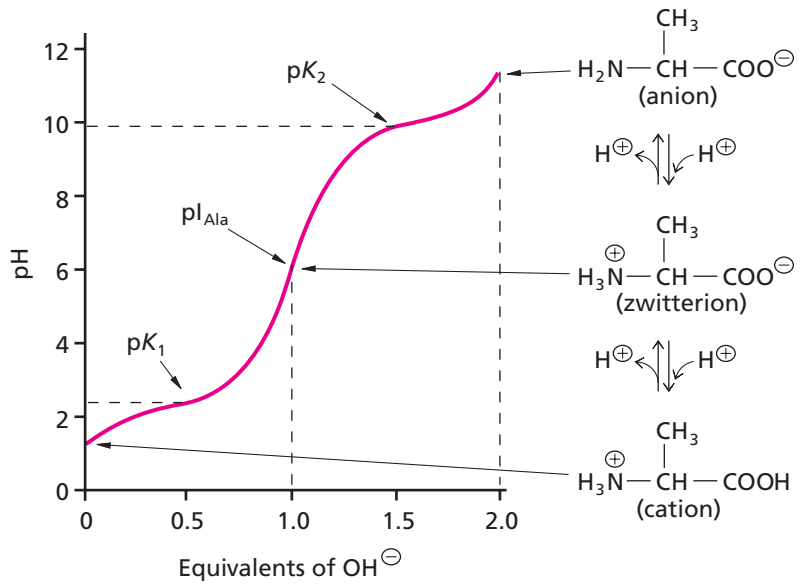


(3.1)

KEY CONCEPT

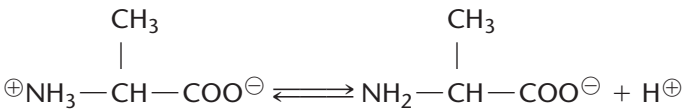
For every acid-base pair the p*Ka* is the pH at which the concentrations of the two forms are equal.

Downloaded From Vet-eBooks.com



◀ **Figure 3.6**
Titration curve for alanine. The first pK_a value is 2.4; the second is 9.9. pI_{Ala} represents the isoelectric point of alanine.

At pH 9.9 the concentration of the zwitterion equals the concentration of the negatively charged form.



(3.2)

Note that in the acid–base pair shown in the first equilibrium (Reaction 3.1) the zwitterion is the conjugate base of the acid form of alanine. In the second acid–base pair (Reaction 3.2) the zwitterion is the proton donor, or conjugate acid, of the more basic form that predominates at higher pH.

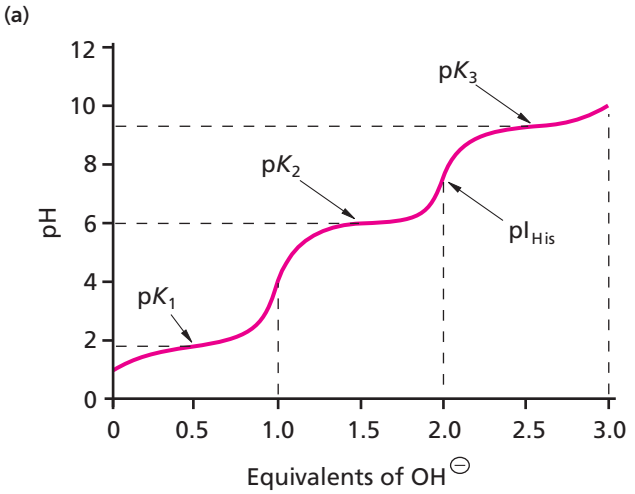
One can deduce that the net charge on alanine molecules at pH 2.4 averages +0.5 because there are equal amounts of neutral zwitterion (+/–) and cation (+). The net charge at pH 9.9 averages –0.5. Midway between pH 2.4 and pH 9.9, at pH 6.15, the average net charge on alanine molecules in solution is zero. For this reason, pH 6.15 is referred to as the isoelectric point (pI), or isoelectric pH, of alanine. If alanine were placed in an electric field at a pH below its pI it would carry a net positive charge (in other words, its cationic form would predominate), and it would therefore migrate toward the cathode (the negative electrode). At a pH higher than its pI alanine would carry a net negative charge and would migrate toward the anode (the positive electrode). At its isoelectric point (pH = 6.15) alanine would not migrate in either direction.

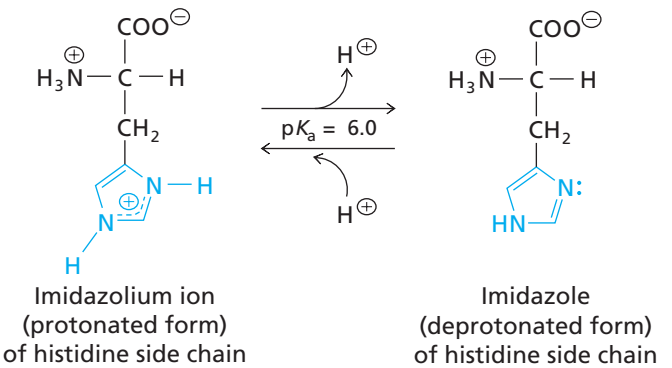
Histidine contains an ionizable side chain. The titration curve for histidine contains an additional inflection point that corresponds to the pK_a of its side chain (Figure 3.7a).

KEY CONCEPT

The ionic state of a particular amino acid side chain is determined by its pK_a value and the pH of the local environment.

▼ **Figure 3.7**
Ionization of histidine. (a) Titration curve for histidine. The three pK_a values are 1.8, 6.0, and 9.3. pI_{His} represents the isoelectric point of histidine. **(b)** Deprotonation of the imidazolium ring of the side chain of histidine.





Downloaded From Vet-e-Books.com

Table 3.2 pK_a values of acidic and basic constituents of free amino acids at 25°C

Amino acid	p*K*a value		
	Carboxyl group	Amino group	Side chain
Glycine	2.4	9.8	
Alanine	2.4	9.9	
Valine	2.3	9.7	
Leucine	2.3	9.7	
Isoleucine	2.3	9.8	
Methionine	2.1	9.3	
Proline	2.0	10.6	
Phenylalanine	2.2	9.3	
Tryptophan	2.5	9.4	
Serine	2.2	9.2	
Threonine	2.1	9.1	
Cysteine	1.9	10.7	8.4
Tyrosine	2.2	9.2	10.5
Asparagine	2.1	8.7	
Glutamine	2.2	9.1	
Aspartic acid	2.0	9.9	3.9
Glutamic acid	2.1	9.5	4.1
Lysine	2.2	9.1	10.5
Arginine	1.8	9.0	12.5
Histidine	1.8	9.3	6.0

As is the case with alanine, the first pK_a (1.8) represents the ionization of the α-COOH carboxyl group and the most basic pK_a value (9.3) represents the ionization of the α-amino group. The middle pK_a (6.0) corresponds to the deprotonation of the imidazolium ion of the side chain of histidine (Figure 3.7b). At pH 7.0 the ratio of imidazole (conjugate base) to imidazolium ion (conjugate acid) is 10:1. Thus, the protonated and neutral forms of the side chain of histidine are both present in significant concentrations near physiological pH. A given histidine side chain in a protein may be either protonated or unprotonated depending on its immediate environment within the protein. In other words, the actual pK_a value of the side-chain group may not be the same as its value for the free amino acid in solution. This property makes the side chain of histidine ideal for the transfer of protons within the catalytic sites of enzymes. (A famous example is described in Section 6.7c.)

The isoelectric point of an amino acid that contains only two ionizable groups (the α-amino and the α-carboxyl groups) is the arithmetic mean of its two pK_a values (i.e., pI = (pK_1 + pK_2)/2). However, for an amino acid that contains three ionizable groups, such as histidine, one must assess the net charge of each ionic species. The isoelectric point for histidine lies between the pK_a values on either side of the species with no net charge, that is, midway between 6.0 and 9.3, or 7.65.

As shown in Table 3.2 the pK_a values of the α-carboxyl groups of free amino acids range from 1.8 to 2.5. These values are lower than those of typical carboxylic acids such as acetic acid (pK_a = 4.8) because the neighboring —$NH_3^{\oplus}$ group withdraws electrons from the carboxylic acid group and this favors the loss of a proton from the α-carboxyl group. The side chains, or R groups, also influence the pK_a value of the α-carboxyl group which is why different amino acids have different pK_a values. (We have just seen that the values for histidine and alanine are not the same.)

The α-COOH group of an amino acid is a weak acid. We can use the Henderson–Hasselbalch equation (Section 2.9) to calculate the fraction of the group that is ionized at any given pH.

$$\text{pH} = \text{p}K_a + \log\frac{[\text{proton acceptor}]}{[\text{proton donor}]} \qquad \textbf{(3.3)}$$

For a typical amino acid whose α-COOH group has a pK_a of 2.0, the ratio of proton acceptor (carboxylate anion) to proton donor (carboxylic acid) at pH 7.0 can be calculated using the Henderson–Hasselbalch equation.

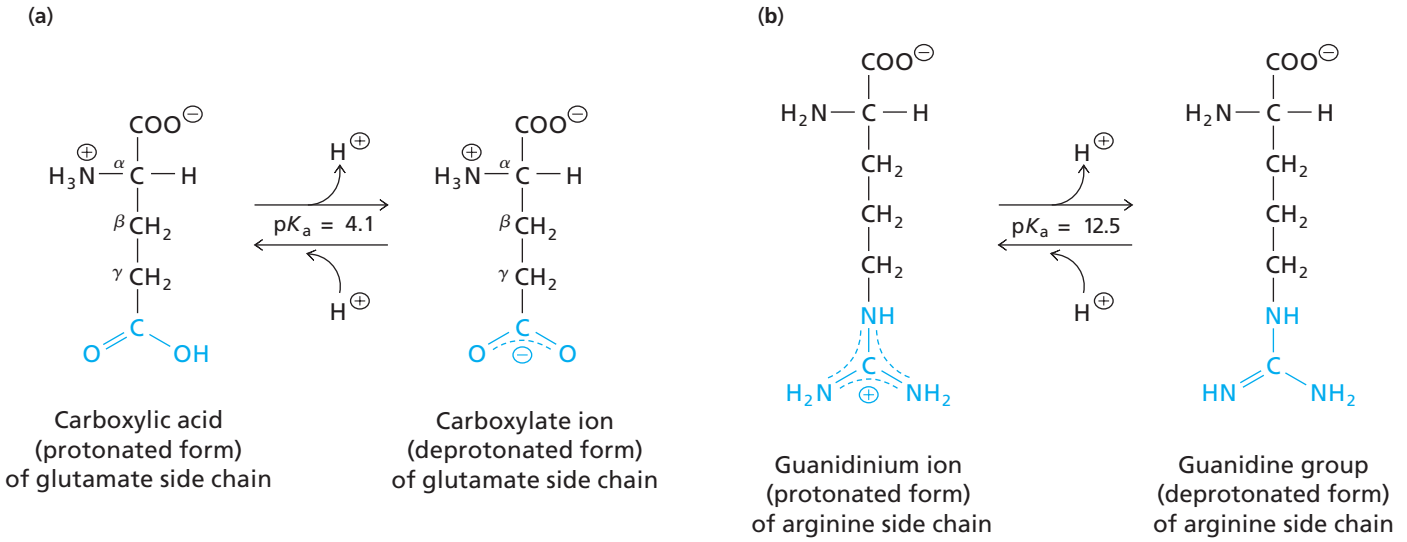
$$7.0 = 2.0 + \log\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]} \qquad \textbf{(3.4)}$$

In this case, the ratio of carboxylate anion to carboxylic acid is 100,000:1. This means that under the conditions normally found inside a cell the carboxylate anion is the predominant species.

The α-amino group of a free amino acid can exist as a free amine, —NH_2 (proton acceptor) or as a protonated amine, —$NH_3^{\oplus}$ (proton donor). The pK_a values range from 8.7 to 10.7 as shown in Table 3.2. For an amino acid whose α-amino group has a pK_a value of 10.0 the ratio of proton acceptor to proton donor is 1:1000 at pH 7.0. In other words, under physiological conditions the α-amino group is mostly protonated and positively charged. These calculations verify our earlier statement that free amino acids exist predominantly as zwitterions at neutral pH. They also show that it is inappropriate to draw the structure of an amino acid with both —COOH and —NH groups since there is no pH at which a significant number of molecules contain a protonated carboxyl group and an unprotonated amino group (see Problem 19). Note that the secondary amino group of proline (pK_a = 10.6) is also protonated at neutral pH so proline—despite the bonding of the side chain to the α-amino group—is also zwitterionic at pH 7.

The seven standard amino acids with readily ionizable groups in their side chains are aspartate, glutamate, histidine, cysteine, tyrosine, lysine, and arginine. Ionization of these groups obeys the same principles as ionization of the α-carboxyl and α-amino groups and the Henderson–Hasselbalch equation can be applied to each ionization. The ionization of the γ-carboxyl group of glutamate (pK_a = 4.1) is shown in Figure 3.8a.

Downloaded From Vet-eBooks.com



▲ **Figure 3.8**
Ionization of amino acid side chains. (a) Ionization of the protonated γ-carboxyl group of glutamate. The negative charge of the carboxylate anion is delocalized. **(b)** Deprotonation of the guanidinium group of the side chain of arginine. The positive charge is delocalized.

Note that the γ-carboxyl group is further removed from the influence of the α-ammonium ion and behaves as a weak acid with a pK_a of 4.1. This makes it similar in strength to acetic acid ($pK_a = 4.8$) whereas the α-carboxyl group is a stronger acid ($pK_a = 2.1$). Figure 3.8b shows the deprotonation of the guanidinium group of the side chain of arginine in a strongly basic solution. Charge delocalization stabilizes the guanidinium ion contributing to its high pK_a value of 12.5.

As mentioned earlier, the pK_a values of ionizable side chains in proteins can differ from those of the free amino acids. Two factors cause this perturbation of ionization constants. First, α-amino and α-carboxyl groups lose their charges once they are linked by peptide bonds in proteins—consequently, they exert weaker inductive effects on their neighboring side chains. Second, the position of an ionizable side chain within the three dimensional structure of a protein can affect its pK_a. For example, the enzyme ribonuclease A has four histidine residues but the side chain of each residue has a slightly different pK_a as a result of differences in their immediate surroundings, or microenvironments.

3.5 Peptide Bonds Link Amino Acids in Proteins

The linear sequence of amino acids in a polypeptide chain is called the **primary structure** of a protein. Higher levels of structure are referred to as secondary, tertiary, and quaternary. The structure of proteins is covered more thoroughly in the next chapter but it's important to understand peptide bonds and primary structure before discussing some of the remaining topics in this chapter.

The structure of peptide bonds is described in Section 4.3.

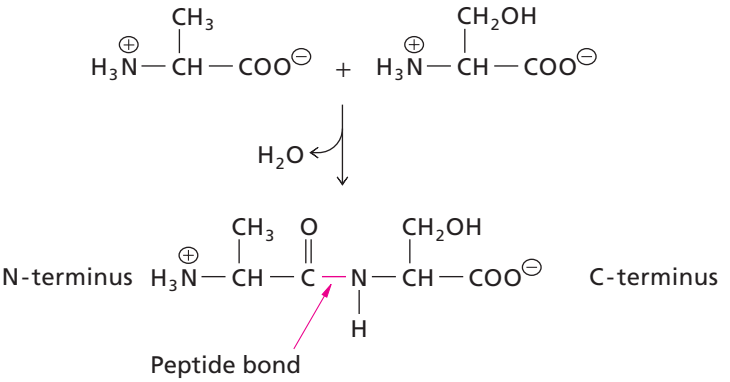
The linkage formed between amino acids is an amide bond called a **peptide bond** (Figure 3.9). This linkage can be thought of as the product of a simple condensation reaction between the α-carboxyl group of one amino acid and the α-amino group of another. A water molecule is lost from the condensing amino acids in the reaction. (Recall from Section 2.6 that such simple condensation reactions are extremely unfavorable in aqueous solutions due to the huge excess of water molecules. The actual pathway of protein synthesis involves reactive intermediates that overcome this limitation.) Unlike the carboxyl and amino groups of free amino acids in solution the groups involved in peptide bonds carry no ionic charges.

Protein synthesis (translation) is described in Chapter 22.

Linked amino acids in a polypeptide chain are called amino acid *residues*. The names of residues are formed by replacing the ending *-ine* or *-ate* with *-yl*. For example, a glycine residue in a polypeptide is called glycyl and a glutamate residue is called glutamyl.

Downloaded From Vet-e-Books.com

Figure 3.9 ►
Peptide bond between two amino acids. The structure of the peptide linkage can be viewed as the product of a condensation reaction in which the α-carboxyl group of one amino acid condenses with the α-amino group of another amino acid. The result is a dipeptide in which the amino acids are linked by a peptide bond. Here, alanine is condensed with serine to form alanylserine.

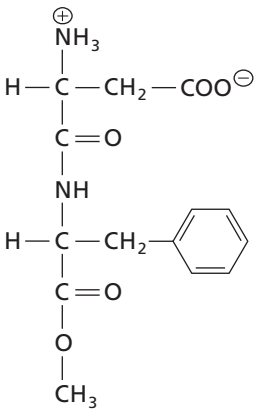


In the cases of asparagine, glutamine, and cysteine, *-yl* replaces the final *-e* to form asparaginyl, glutaminyl, and cysteinyl, respectively. The *-yl* ending indicates that the residue is an acyl unit (a structure that lacks the hydroxyl of the carboxyl group). The dipeptide in Figure 3.9 is called alanylserine because alanine is converted to an acyl unit but the amino acid serine retains its carboxyl group.

The free amino group and free carboxyl group at the opposite ends of a peptide chain are called the N-terminus (amino terminus) and the C-terminus (carboxyl terminus), respectively. At neutral pH each terminus carries an ionic charge. By convention, amino acid residues in a peptide chain are numbered from the N-terminus to the C-terminus and are usually written from left to right. This convention corresponds to the direction of protein synthesis (Section 22.6). Synthesis begins with the N-terminal amino acid—almost always methionine (Section 22.5)—and proceeds sequentially toward the C-terminus by adding one residue at a time.

Both the standard three-letter abbreviations for the amino acids (e.g., Gly–Arg–Phe–Ala–Lys) and the one-letter abbreviations (e.g., GRFAK) are used to describe the sequence of amino acid residues in peptides and polypeptides. It's important to know both abbreviation systems. The terms *dipeptide*, *tripeptide*, *oligopeptide*, and *polypeptide* refer to chains of two, three, several (up to about 20), and many (usually more than 20) amino acid residues, respectively. A dipeptide contains one peptide bond, a tripeptide contains two peptide bonds, and so on. As a general rule, each peptide chain, whatever its length, possesses one free α-amino group and one free α-carboxyl group. (Exceptions include covalently modified terminal residues and circular peptide chains.) Note that the formation of a peptide bond eliminates the ionizable α-carboxyl and α-amino groups found in free amino acids. As a result, most of the ionic charges associated with a protein molecule are contributed by the side chains of the amino acids. This means that the solubility and ionic properties of a protein are largely determined by its amino acid composition. Furthermore, the side chains of the residues interact with each other and these interactions contribute to the three dimensional shape and stability of a protein molecule (Chapter 4).

Some peptides are important biological compounds and the chemistry of peptides is an active area of research. Several hormones are peptides; for example, endorphins are the naturally occurring molecules that modulate pain in vertebrates. Some very simple peptides are useful as food additives; for example, the sweetening agent aspartame is the methyl ester of aspartylphenylalanine (Figure 3.10). Aspartame is about 200 times sweeter than table sugar and is widely used in diet drinks. There are also many peptide toxins such as those found in snake venom and poisonous mushrooms.



▲ **Figure 3.10**
Aspartame (aspartylphenylalanine methyl ester).

3.6 Protein Purification Techniques

In order to study a particular protein in the laboratory it must be separated from all other cell components including other, similar proteins. Few analytical techniques will work with crude mixtures of cellular proteins because they contain hundreds (or thousands) of different proteins. The purification steps are different for each protein. They are worked

Downloaded From Vet-eBooks.com

out by trying a number of different techniques until a procedure is developed that reproducibly yields highly purified protein that is still biologically active. Purification steps usually exploit minor differences in the solubilities, net charges, sizes, and binding specificities of proteins. In this section, we consider some of the common methods of protein purification. Most purification techniques are performed at 0°C to 4°C to minimize temperature-dependent processes such as protein degradation and denaturation (unfolding).

The first step in protein purification is to prepare a solution of proteins. The source of a protein is often whole cells in which the target protein accounts for less than 0.1% of the total dry weight. Isolation of an intracellular protein requires that cells be suspended in a buffer solution and homogenized, or disrupted into cell fragments. Under these conditions most proteins dissolve. (Major exceptions include membrane proteins which require special purification procedures.) Let's assume that the desired protein is one of many proteins in this solution.

One of the first steps in protein purification is often a relatively crude separation that makes use of the different solubilities of proteins in salt solutions. Ammonium sulfate is frequently used in such fractionations. Enough ammonium sulfate is mixed with the solution of proteins to precipitate the less soluble impurities, which are removed by centrifugation. The target protein and other more soluble proteins remain in the fluid called the supernatant fraction. Next, more ammonium sulfate is added to the supernatant fraction until the desired protein is precipitated. The mixture is centrifuged, the fluid removed, and the precipitate dissolved in a minimal volume of buffer solution. Typically, fractionation using ammonium sulfate gives a two- to threefold purification (i.e., one-half to two-thirds of the unwanted proteins have been removed from the resulting enriched protein fraction). At this point the solvent containing residual ammonium sulfate is exchanged by dialysis for a buffer solution suitable for chromatography.

In dialysis, a protein solution is sealed in a cylinder of cellophane tubing and suspended in a large volume of buffer. The cellophane membrane is semipermeable—high molecular weight proteins are too large to pass through the pores of the membrane so proteins remain inside the tubing while low molecular weight solutes (including, in this case, ammonium and sulfate ions) diffuse out and are replaced by solutes in the buffer.

Column chromatography is often used to separate a mixture of proteins. A cylindrical column is filled with an insoluble material such as substituted cellulose fibers or synthetic beads. The protein mixture is applied to the column and washed through the matrix of insoluble material by the addition of solvent. As solvent flows through the column the eluate (the liquid emerging from the bottom of the column) is collected in many fractions, a few of which are represented in Figure 3.11a. The rate at which proteins travel through the matrix depends on interactions between matrix and protein. For a given column different proteins are eluted at different rates. The concentration of protein in each fraction can be determined by measuring the absorbance of the eluate at a wavelength of 280 nm (Figure 3.11b). (Recall from Section 3.2B that at neutral pH, tyrosine and tryptophan absorb UV light at 280 nm.) To locate the target protein the fractions containing protein must then be assayed, or tested, for biological activity or some other characteristic property. Column chromatography may be performed under high pressure using small, tightly packed columns with solvent flow controlled by a computer. This technique is called **HPLC**, for high-performance liquid chromatography.

Chromatographic techniques are classified according to the type of matrix. In **ion-exchange chromatography** the matrix carries positive charges (anion-exchange resins) or negative charges (cation-exchange resins). Anion-exchange matrices bind negatively charged proteins retaining them in the matrix for subsequent elution. Conversely, cation-exchange materials bind positively charged proteins. The bound proteins can be serially eluted by gradually increasing the salt concentration in the solvent. As the salt concentration is increased it eventually reaches a concentration where the salt ions outcompete proteins in binding to the matrix. At this concentration the protein is released and is collected in the eluate. Individual bound proteins are eluted at different salt concentrations and this fractionation makes ion-exchange chromatography a powerful tool in protein purification.

Gel-filtration chromatography separates proteins on the basis of molecular size. The gel is a matrix of porous beads. Proteins that are smaller than the average pore size



▲ There is only one correct way to write the sequence of a polypeptide- from N-teminus to C-terminus.



▲ **Green mamba (*Dendroapsis angusticeps*).** One of the toxins in the venom of this poisonous snake is a large peptide with the sequence MICYSHKTPQPSATITCEEKT-CYKKSVRKL PAVVAGRGCGCPSKEMLVAIH CCRSDKCNE [Viljoen and Botes (1974). *J.Biol.Chem.* 249:366]

Downloaded From Vet-eBooks.com

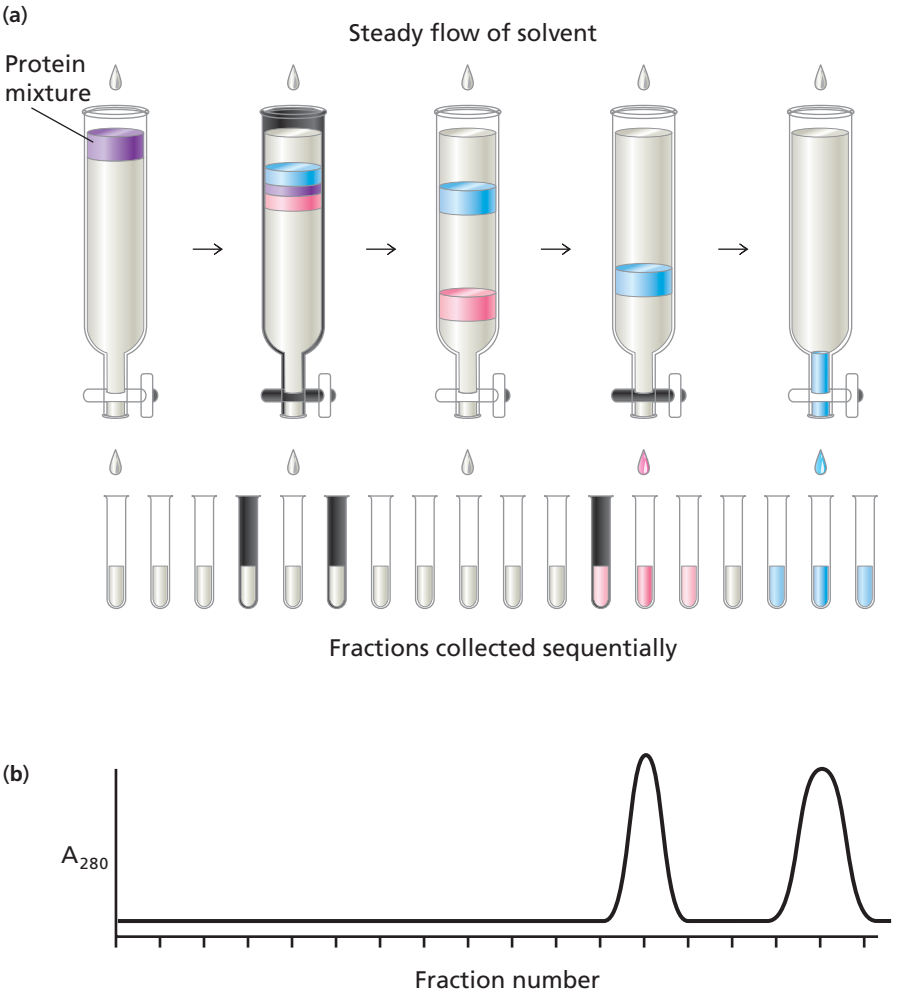
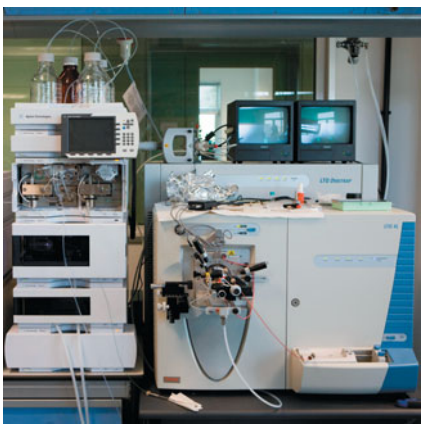


Figure 3.11 ►
Column chromatography. (a) A mixture of proteins is added to a column containing a solid matrix. Solvent then flows into the column from a reservoir. Washed by solvent, different proteins (represented by red and blue bands) travel through the column at different rates, depending on their interactions with the matrix. Eluate is collected in a series of fractions, a few of which are shown. **(b)** The protein concentration of each fraction is determined by measuring the absorbance at 280 nm. The peaks correspond to the elution of the protein bands shown in (a). The fractions are then tested for the presence of the target protein.



▲ A typical high-performance liquid chromatography (HPLC) system in a research lab (left). The large instrument on the right is a mass spectrometer (Istituto di Ricerche Farmacologiche, Milan, Italy)

penetrate much of the internal volume of the beads and are therefore retarded by the matrix as the buffer solution flows through the column. The smaller the protein, the later it elutes from the column. Fewer of the pores are accessible to larger protein molecules. Consequently, the largest proteins flow past the beads and elute first.

Affinity chromatography is the most selective type of column chromatography. It relies on specific binding interactions between the target protein and some other molecule that is covalently bound to the matrix of the column. The molecule bound to the matrix may be a substance or a ligand that binds to a protein *in vivo*, an antibody that recognizes the target protein, or another protein that is known to interact with the target protein inside the cell. As a mixture of proteins passes through the column only the target protein specifically binds to the matrix. The column is then washed with buffer several times to rid it of nonspecifically bound proteins. Finally, the target protein can be eluted by washing the column with a solvent containing a high concentration of salt that disrupts the interaction between the protein and column matrix. In some cases, bound protein can be selectively released from the affinity column by adding excess ligand to the elution buffer. The target protein preferentially binds to the ligand in solution instead of the lower concentration of ligand that is attached to the insoluble matrix of the column. This method is most effective when the ligand is a small molecule. Affinity chromatography alone can sometimes purify a protein 1000- to 10,000-fold.

3.7 Analytical Techniques

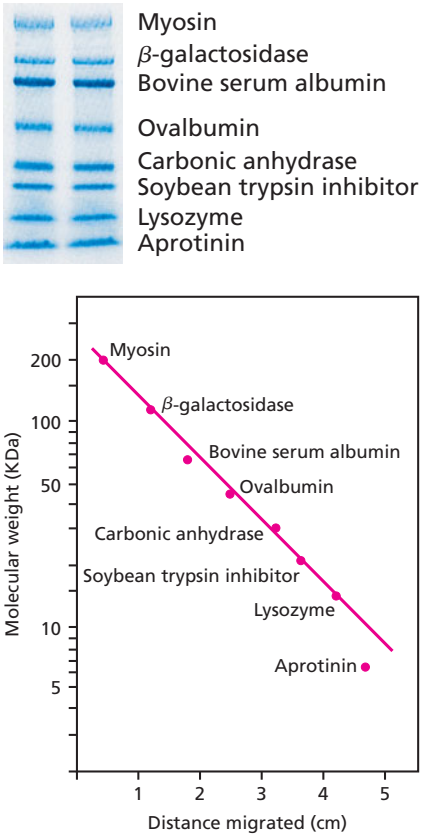
Electrophoresis separates proteins based on their migration in an electric field. In **polyacrylamide gel electrophoresis (PAGE)** protein samples are placed on a highly cross-linked gel matrix of polyacrylamide and an electric field is applied. The matrix is

Downloaded From Vet-eBooks.com

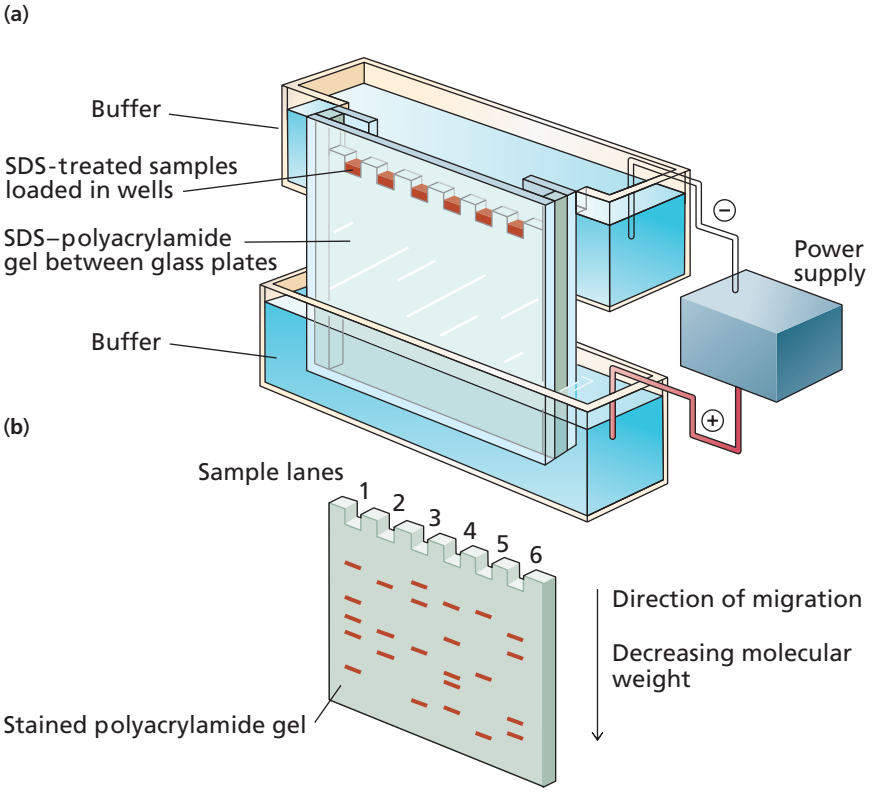
buffered to a mildly alkaline pH so that most proteins are anionic and migrate toward the anode. Typically, several samples are run at once together with a reference sample. The gel matrix retards the migration of large molecules as they move in the electric field. Hence, proteins are fractionated on the basis of both charge and mass.

A modification of the standard electrophoresis technique uses the negatively charged detergent sodium dodecyl sulfate (SDS) to overwhelm the native charge on proteins so that they are separated on the basis of mass only. SDS–polyacrylamide gel electrophoresis (**SDS–PAGE**) is used to assess the purity and to estimate the molecular weight of a protein. In SDS–PAGE the detergent is added to the polyacrylamide gel as well as to the protein samples. A reducing agent is also added to the samples to reduce any disulfide bonds. The dodecyl sulfate anion, which has a long hydrophobic tail ($CH_3(CH_2)_{11}OSO_3^{\ominus}$, Figure 2.8) binds to hydrophobic side chains of amino acid residues in the polypeptide chain. SDS binds at a ratio of approximately one molecule for every two residues of a typical protein. Since larger proteins bind proportionately more SDS the charge-to-mass ratios of all treated proteins are approximately the same. All the SDS–protein complexes are highly negatively charged and move toward the anode as diagrammed in Figure 3.12a. However, their rate of migration through the gel is inversely proportional to the logarithm of their mass—larger proteins encounter more resistance and therefore migrate more slowly than smaller proteins. This sieving effect differs from gel-filtration chromatography because in gel filtration larger molecules are excluded from the pores of the gel and hence travel faster. In SDS–PAGE all molecules penetrate the pores of the gel so the largest proteins travel most slowly. The protein bands that result from this differential migration (Figure 3.13) can be visualized by staining. Molecular weights of unknown proteins can be estimated by comparing their migration to the migration of reference proteins on the same gel.

Although SDS–PAGE is primarily an analytical tool, it can be adapted for purifying proteins. Denatured proteins can be recovered from SDS–PAGE by cutting out the bands of a gel. The protein is then electroeluted by applying an electric current to allow the protein to migrate into a buffer solution. After concentration and the removal of salts such protein preparations can be used for structural analysis, preparation of antibodies, or other purposes.



▲ **Figure 3.13**
Proteins separated on an SDS–polyacrylamide gel. (a) Stained proteins after separation. The high molecular weight proteins are at the top of the gel. **(b)** Graph showing the relationship between the molecular weight of a protein and the distance it migrates in the gel.



◀ **Figure 3.12**
SDS–PAGE. (a) An electrophoresis apparatus includes an SDS–polyacrylamide gel between two glass plates and buffer in the upper and lower reservoirs. Samples are loaded into the wells of the gel, and voltage is applied. Because proteins complexed with SDS are negatively charged, they migrate toward the anode. **(b)** The banding pattern of the proteins after electrophoresis can be visualized by staining. The smallest proteins migrate fastest, so the proteins of lowest molecular weight are at the bottom of the gel.

Downloaded From Vet-e-Books.com

Mass spectrometry, as the name implies, is a technique that determines the mass of a molecule. The most basic type of mass spectrometer measures the time that it takes for a charged gas phase molecule to travel from the point of injection to a sensitive detector. This time depends on the charge of a molecule and its mass and the result is reported as the mass/charge ratio. The technique has been used in chemistry for almost 100 years but its application to proteins was limited because, until recently, it was not possible to disperse charged protein molecules into a gaseous stream of particles.

This problem was solved in the late 1980s with the development of two new types of mass spectrometry. In **electrospray mass spectrometry** the protein solution is pumped through a metal needle at high voltage to create tiny droplets. The liquid rapidly evaporates in a vacuum and the charged proteins are focused on a detector by a magnetic field. The second new technique is called **matrix-assisted laser desorption ionization (MALDI)**. In this method the protein is mixed with a chemical matrix and the mixture is precipitated on a metal substrate. The matrix is a small organic molecule that absorbs light at a particular wavelength. A laser pulse at the absorption wavelength imparts energy to the protein molecules via the matrix. The proteins are instantly released from the substrate (desorbed) and directed to the detector (Figure 3.14). When time-of-flight (TOF) is measured, the technique is called MALDI–TOF.

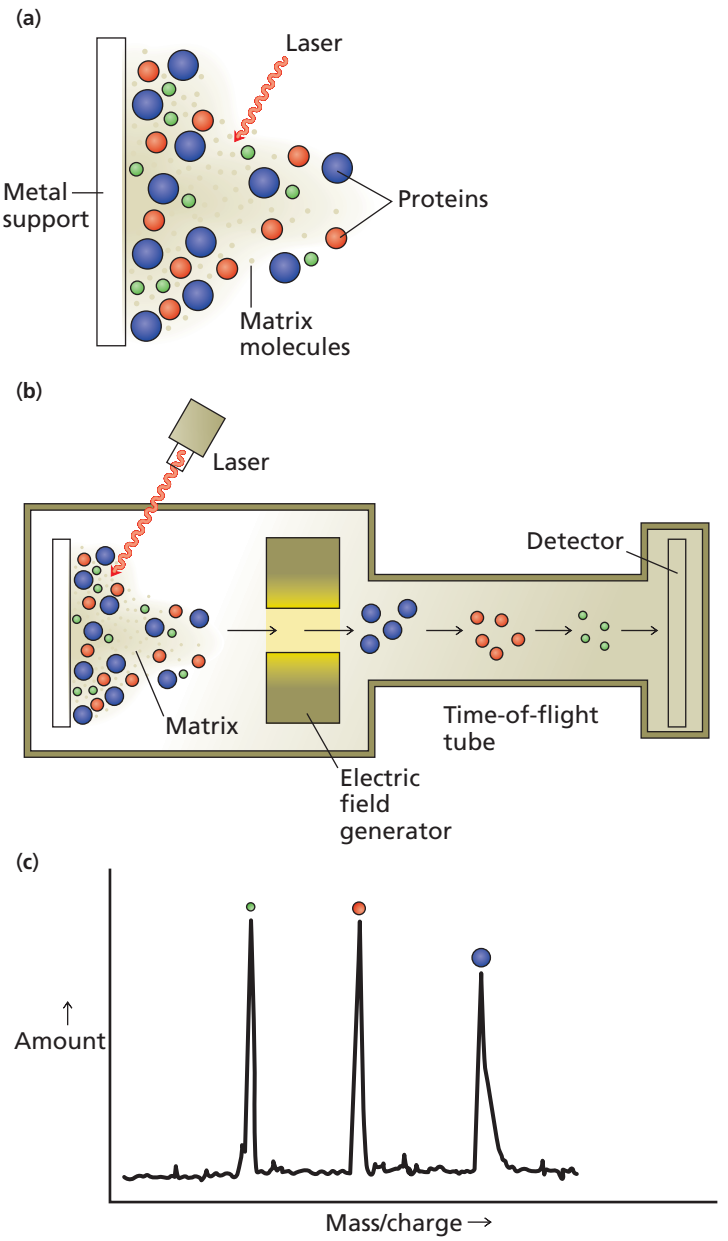


Figure 3.14 ▶
MALDI–TOF mass spectrometry. (a) A burst of light releases proteins from the matrix. **(b)** Charged proteins are directed toward the detector by an electric field. **(c)** The time of arrival at the detector depends on the mass and the charge of the protein.

Downloaded From Vet-e-Books.com

The raw data from a mass spectrometry experiment can be quite simple as shown in Figure 3.14. There, a single species with one positive charge is detected so the mass/charge ratio gives the mass directly. In other cases the spectra can be more complicated, especially in electrospray mass spectrometry. Often there are several different charged species and the correct mass has to be calculated by analyzing a collection of molecules with charges of +1, +2, +3, etc. The spectrum can be daunting when the source is a mixture of different proteins. Fortunately, there are sophisticated computer programs that can analyze the data and calculate the correct masses. The current popularity of mass spectrometry owes as much to the development of this software as it does to the new hardware and new methods of sample preparation.

Mass spectrometry is very sensitive and highly accurate. Often the mass of a protein can be obtained from picomole (10^{-12} mol) quantities that are isolated from an SDS–PAGE gel. The correct mass can be determined with an accuracy of less than the mass of a single proton.



John B. Fenn (1917–)



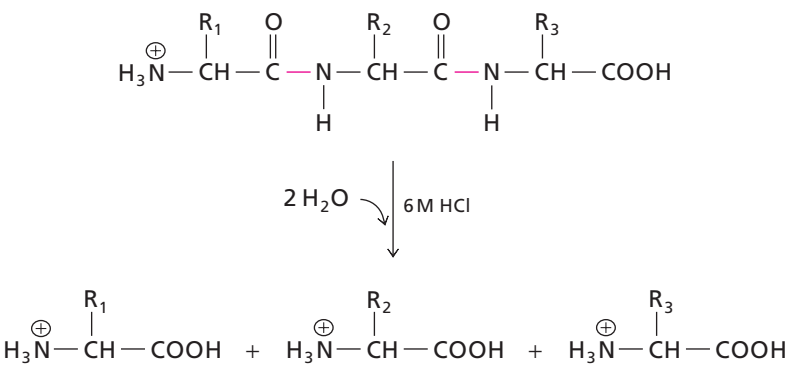
Koichi Tanaka (1959–)

▲ John B. Fenn and Koichi Tanaka were awarded the Nobel Prize in Chemistry in 2002 "for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules."

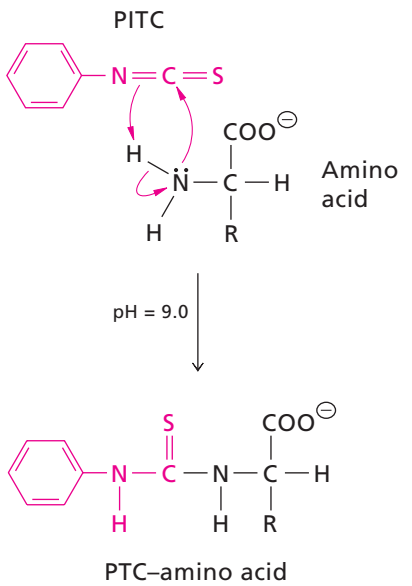
3.8 Amino Acid Composition of Proteins

Once a protein has been isolated its amino acid composition can be determined. First, the peptide bonds of the protein are cleaved by acid hydrolysis, typically using 6 M HCl (Figure 3.15). Next, the hydrolyzed mixture, or hydrolysate, is subjected to a chromatographic procedure in which each of the amino acids is separated and quantitated, a process called **amino acid analysis**. One method of amino acid analysis involves treatment of the protein hydrolysate with phenylisothiocyanate (PITC) at pH 9.0 to generate phenylthiocarbamoyl (PTC)–amino acid derivatives (Figure 3.16). The PTC–amino acid mixture is then subjected to HPLC in a column of fine silica beads to which short hydrocarbon chains have been attached. The amino acids are separated by the hydrophobic properties of their side chains. As each PTC–amino acid derivative is eluted it is detected and its concentration is determined by measuring the absorbance of the eluate at 254 nm (the peak absorbance of the PTC moiety). Since different PTC–amino acid derivatives are eluted at different rates the time at which an amino acid derivative elutes from the column identifies the amino acid relative to known standards. The amount of each amino acid in the hydrolysate is proportional to the area under its peak. With this method, amino acid analysis can be performed on samples as small as 1 picomole of a protein that contains approximately 200 residues.

Despite its usefulness, acid hydrolysis cannot yield a complete amino acid analysis. Since the side chains of asparagine and glutamine contain amide bonds the acid used to cleave the peptide bonds of the protein also converts asparagine to aspartic acid and glutamine to glutamic acid. Other limitations of the acid hydrolysis method include small losses of serine, threonine, and tyrosine. In addition, the side chain of tryptophan is almost totally destroyed by acid hydrolysis. There are several ways of overcoming these limitations. For example, proteins can be hydrolyzed to amino acids by enzymes



▲ **Figure 3.15**
Acid-catalyzed hydrolysis of a peptide. Incubation with 6 M HCl at 110°C for 16 to 72 hours releases the constituent amino acids of a peptide.



▲ **Figure 3.16**
Amino acid treated with phenylisothiocyanate (PITC). The α-amino group of an amino acid reacts with phenylisothiocyanate to give a phenylthiocarbamoyl–amino acid (PTC–amino acid).

Downloaded From Vet-e-Books.com

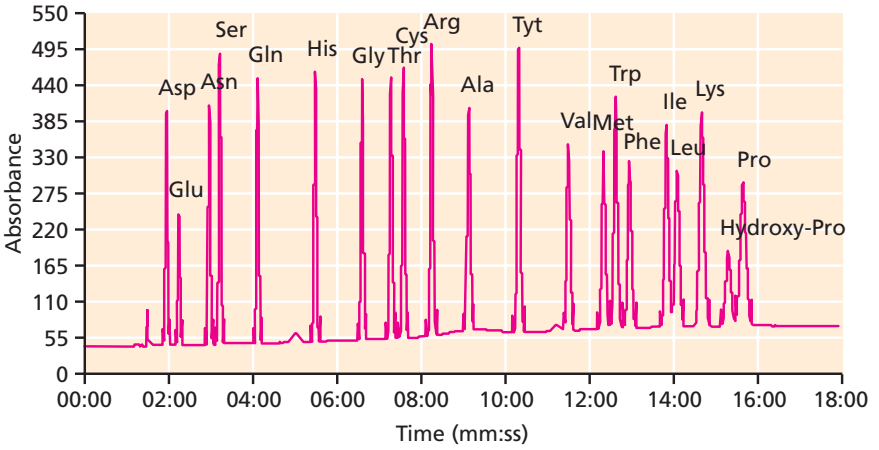


Figure 3.17 ►
HPLC separation of amino acids. Amino acids obtained from the enzymatic hydrolysis of a protein are treated with o-phthalaldehyde and separated by HPLC.

The frequency of amino acids in proteins is correlated with the number of codons for each amino acid (Section 22.1)

instead of using acid hydrolysis. The free amino acids are then attached to a chemical that absorbs light in the ultraviolet and the derivatized amino acids are analyzed by HPLC (Figure 3.17).

Using various analytical techniques the complete amino acid compositions of many proteins have been determined. Dramatic differences in composition have been found, illustrating the tremendous potential for diversity based on different combinations of the 20 amino acids.

The amino acid composition (and sequence) of proteins can also be determined from the sequence of its gene. In fact, these days it is often much easier to clone and sequence DNA than it is to purify and sequence a protein. Table 3.3 shows the average frequency of amino acid residues in more than 1000 different proteins whose sequences are deposited in protein databases. The most common amino acids are leucine, alanine, and glycine, followed by serine, valine, and glutamate. Tryptophan, cysteine, and histidine are the least abundant amino acids in typical proteins.

If you know the amino acid composition of a protein you can calculate the molecular weight using the molecular weights of the amino acids in Table 3.4. Be sure to subtract the molecular weight of one water molecule for each peptide bond (Section 3.5). You can get a rough estimate of the molecular weight of a protein by using the average molecular weight of a residue (= 110). Thus, a protein of 650 amino acid residues has an approximate relative molecular mass of 71,500 (M_r = 71,500).

Table 3.3 Amino acid compositions of proteins

Amino acid	Frequency in proteins (%)
Highly hydrophobic	
Ile (I)	5.2
Val (V)	6.6
Leu (L)	9.0
Phe (F)	3.9
Met (M)	2.4
Less hydrophobic	
Ala (A)	8.3
Gly (G)	7.2
Cys (C)	1.7
Trp (W)	1.3
Tyr (Y)	3.2
Pro (P)	5.1
Thr (T)	5.8
Ser (S)	6.9
Highly hydrophilic	
Asn (N)	4.4
Gln (Q)	4.0
Acidic	
Asp (D)	5.3
Glu (E)	6.2
Basic	
His (H)	2.2
Lys (K)	5.7
Arg (R)	5.7

3.9 Determining the Sequence of Amino Acid Residues

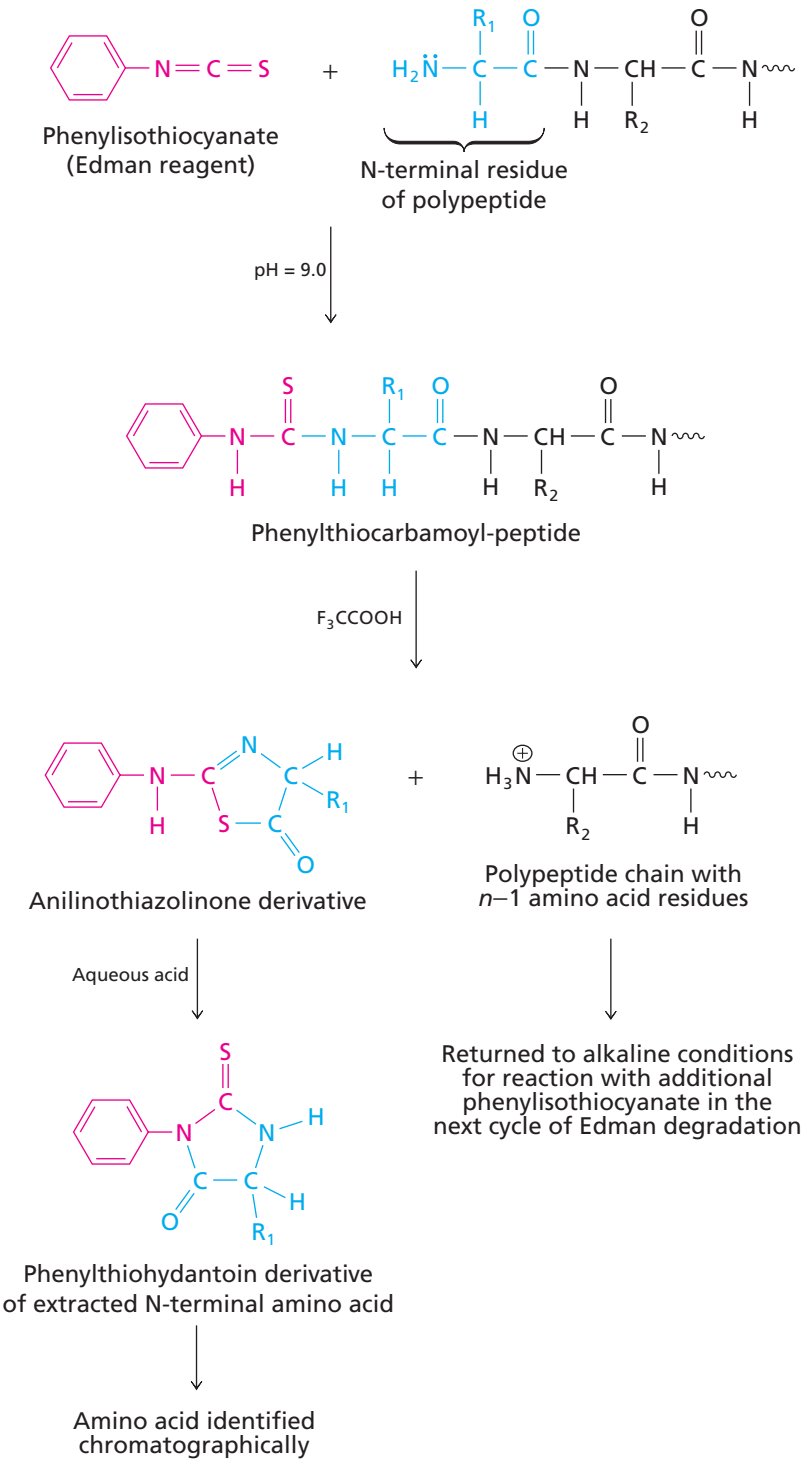
Amino acid analysis provides information on the composition of a protein but not its primary structure (sequence of residues). In 1950, Pehr Edman developed a technique that permits removal and identification of one residue at a time from the N-terminus of a protein. The **Edman degradation procedure** involves treating a protein at pH 9.0 with PITC, also known as the Edman reagent. (Recall that PITC can also be used in the measurement of free amino acids as shown in Figure 3.16.) PITC reacts with the free N-terminus of the chain to form a phenylthiocarbamoyl derivative, or PTC-peptide (Figure 3.18, on the next page). When the PTC-peptide is treated with an anhydrous acid, such as trifluoroacetic acid the peptide bond of the N-terminal residue is selectively cleaved releasing an anilinothiazolinone derivative of the residue. This derivative can be extracted with an organic solvent, such as butyl chloride, leaving the remaining peptide in the aqueous phase. The unstable anilinothiazolinone derivative is then treated with aqueous acid which converts it to a stable phenylthiohydantoin derivative of the amino acid that had been the N-terminal residue (PTH–amino acid). The polypeptide chain in the aqueous phase, now one residue shorter (residue 2 of the original protein is now the N-terminus), can be adjusted back to pH 9.0 and treated again with PITC. The entire procedure can be repeated serially using an automated instrument known as a sequenator. Each cycle yields a PTH–amino acid that can be identified chromatographically, usually by HPLC.

Downloaded From Vet-e-Books.com

The yield of the Edman degradation procedure under carefully controlled conditions approaches 100% and a few picomoles of sample protein can yield sequences of 30 residues or more before further measurement is obscured by the increasing concentration of unrecovered sample from previous cycles of the procedure. For example, if the Edman degradation procedure had an efficiency of 98% the cumulative yield at the 30th cycle would be 0.98^{30}, or 0.55. In other words, only about half of the PTH–amino acids generated in the 30th cycle would be derived from the 30th residue from the N-terminus.

Table 3.4 Molecular weights of amino acids

Amino acid	M_r
Ala(A)	89
Arg(R)	174
Asn(N)	132
Asp(D)	133
Cys(C)	121
Gln(O)	146
Glu(E)	147
Gly(G)	75
His(H)	155
He(I)	131
Leu(L)	131
Lys(K)	146
Met(M)	149
Phe(F)	165
Pro(P)	115
Ser(S)	105
Thr(T)	119
Trp(W)	204
Tyr(Y)	181
Val(V)	117



◄ **Figure 3.18**
Edman degradation procedure. The N-terminal residue of a polypeptide chain reacts with phenylisothiocyanate to give a phenylthiocarbamoyl–peptide. Treating this derivative with trifluoroacetic acid (F_3CCOOH) releases an anilinothiazolinone derivative of the N-terminal amino acid residue. The anilinothiazolinone is extracted and treated with aqueous acid, which rearranges the derivative to a stable phenylthiohydantoin derivative that can then be identified chromatographically. The remainder of the polypeptide chain, whose new N-terminal residue was formerly in the second position, is subjected to the next cycle of Edman degradation.

Downloaded From Vet-e-Books.com

3.10 Protein Sequencing Strategies

Most proteins contain too many residues to be completely sequenced by Edman degradation proceeding only from the N-terminus. Therefore, proteases (enzymes that catalyze the hydrolysis of peptide bonds in proteins) or certain chemical reagents are used to selectively cleave some of the peptide bonds of a protein. The smaller peptides formed are then isolated and subjected to sequencing by the Edman degradation procedure.

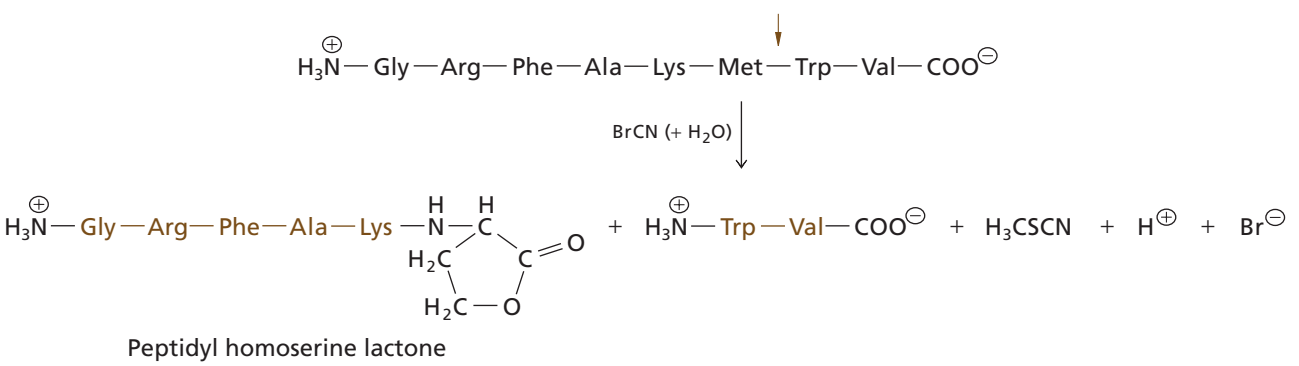
The chemical reagent cyanogen bromide (CNBr) reacts specifically with methionine residues to produce peptides with C-terminal homoserine lactone residues and new N-terminal residues (Figure 3.19). Since most proteins contain relatively few methionine residues treatment with CNBr usually produces only a few peptide fragments. For example, reaction of CNBr with a polypeptide chain containing three internal methionine residues should generate four peptide fragments. Each fragment can then be sequenced from its N-terminus.

Many different proteases can be used to generate fragments for protein sequencing. For example, trypsin specifically catalyzes the hydrolysis of peptide bonds on the carbonyl side of lysine and arginine residues both of which bear positively charged side chains (Figure 3.20a). *Staphylococcus aureus* V8 protease catalyzes the cleavage of peptide bonds on the carbonyl side of negatively charged residues (glutamate and aspartate); under appropriate conditions (50 mM ammonium bicarbonate), it cleaves only glutamyl bonds. Chymotrypsin, a less specific protease, preferentially catalyzes the hydrolysis of peptide bonds on the carbonyl side of uncharged residues with aromatic or bulky hydrophobic side chains, such as phenylalanine, tyrosine, and tryptophan (Figure 3.20b).

By judicious application of cyanogen bromide, trypsin, *S. aureus* V8 protease, and chymotrypsin to individual samples of a large protein one can generate many peptide fragments of various sizes. These fragments can then be separated and sequenced by Edman degradation. In the final stage of sequence determination the amino acid sequence of a large polypeptide chain can be deduced by lining up matching sequences of overlapping peptide fragments as illustrated in Figure 3.20c. When referring to an amino acid residue whose position in the sequence is known it is customary to follow the residue abbreviation with its sequence number. For example, the third residue of the peptide shown in Figure 3.20 is called Ala-3.

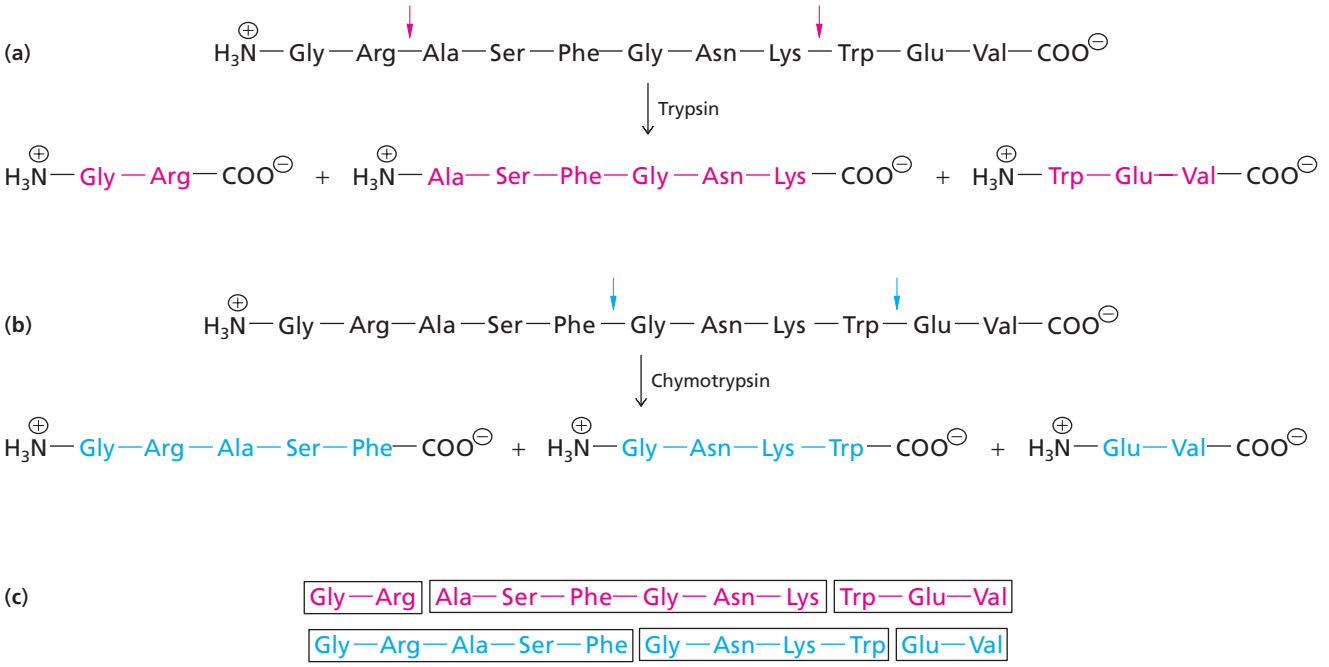
The process of generating and sequencing peptide fragments is especially important in obtaining information about the sequences of proteins whose N-termini are blocked. For example, the N-terminal α-amino groups of many bacterial proteins are formylated and do not react at all when subjected to the Edman degradation procedure. Peptide fragments with unblocked N-termini can be produced by selective cleavage and then separated and sequenced so that at least some of the internal sequence of the protein can be obtained.

For proteins that contain disulfide bonds, the complete covalent structure is not fully resolved until the positions of the disulfide bonds have been established. The positions of the disulfide cross-links can be determined by fragmenting the intact protein, isolating the peptide fragments, and determining which fragments contain cystine residues. The task of determining the positions of the cross-links becomes quite complicated when the protein contains several disulfide bonds.



▼ **Figure 3.19**
Protein cleavage by cyanogen bromide (CNBr). Cyanogen bromide cleaves polypeptide chains at the C-terminal side of methionine residues. The reaction produces a peptidyl homoserine lactone and generates a new N-terminus.

Downloaded From Vet-e-Books.com



▲ **Figure 3.20**
Cleavage and sequencing of an oligopeptide. **(a)** Trypsin catalyzes cleavage of peptides on the carbonyl side of the basic residues arginine and lysine. **(b)** Chymotrypsin catalyzes cleavage of peptides on the carbonyl side of uncharged residues with aromatic or bulky hydrophobic side chains, including phenylalanine, tyrosine, and tryptophan. **(c)** By using the Edman degradation procedure to determine the sequence of each fragment (highlighted in boxes) and then lining up the matching sequences of overlapping fragments, one can determine the order of the fragments and thus deduce the sequence of the entire oligopeptide.

Deducing the amino acid sequence of a particular protein from the sequence of its gene (Figure 3.21) overcomes some of the technical limitations of direct analytical techniques. For example, the amount of tryptophan can be determined and aspartate and asparagine residues can be distinguished because they are encoded by different codons. However, direct sequencing of proteins is still important since it is the only way of determining whether modified amino acids are present or whether amino acid residues have been removed after protein synthesis is complete.

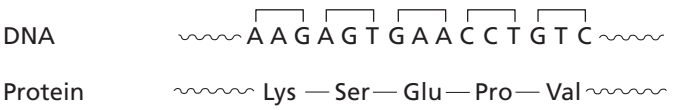
Researchers frequently want to identify a particular unknown protein. Let's say you have displayed human serum proteins on an SDS gel and you note the presence of a protein band at 67 KDa. What is that protein? Two recent developments have made the job of identifying unknown proteins much easier—sensitive mass spectrometry and genome sequences. Let's see how they work.

First, you isolate the protein by cutting out the unknown protein band and eluting the 67 KD protein. The next step is to digest the protein with a protease that cuts at specific sites. Let's say you choose trypsin, an enzyme that cleaves the peptide bond following arginine (R) or lysine (K) residues. After digestion with trypsin you end up with several dozen peptide fragments all of which end with arginine or lysine.

Next, you subject the peptide mixture to mass spectrometry choosing a method such as MALDI–TOF where the precise molecular weights of the peptides can be determined. The resulting spectrum is shown in Figure 3.22. You now have a "fingerprint" of the unknown protein corresponding to the molecular weights of all the trypsin digestion products.

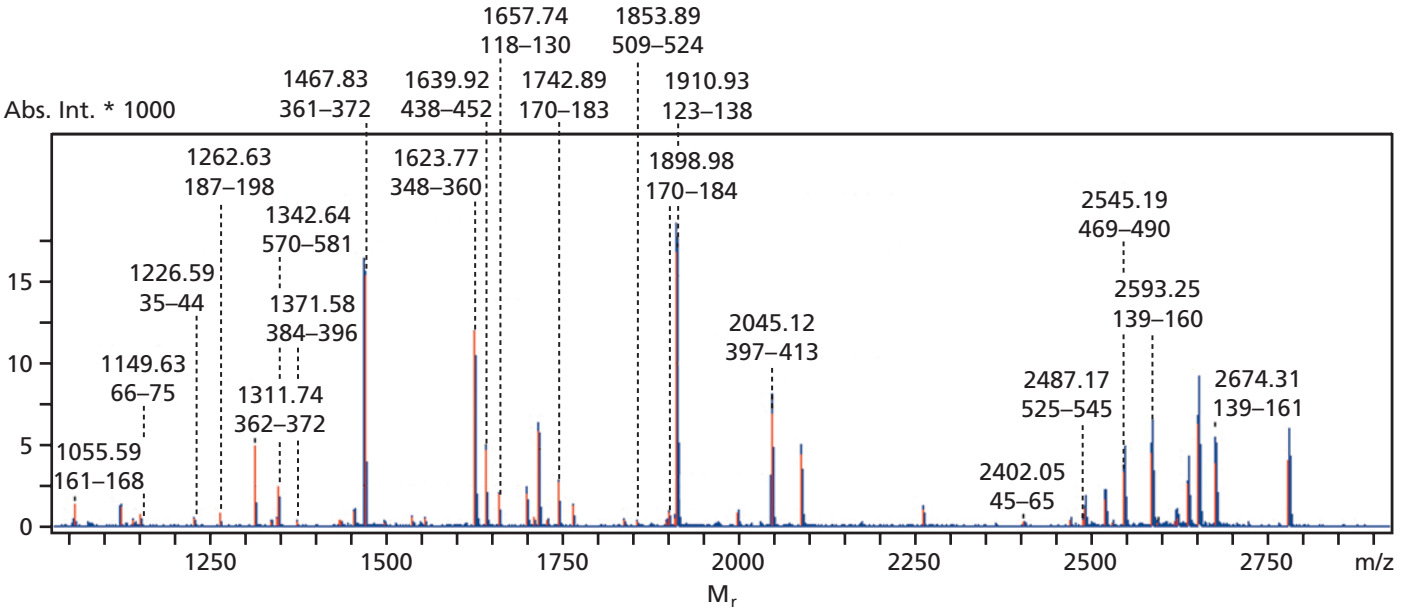
In many labs the technique of chemical sequencing using Edman degradation has been replaced by methods using the mass spectrometer. If you wanted to determine the sequences of each peptide shown in Figure 3.22 your next step would be to fragment each peptide into various sized pieces and measure the precise molecular weight of each fragment in the mass spectrometer.

The data can be used to determine the sequence of the peptide. For example, take the tryptic peptide of M_r = 1226.59 shown in Figure 3.22. One of the large pieces produced by fragmenting this peptide has a molecular weight of 1079.5. The difference



◀ **Figure 3.21**
Sequences of DNA and protein. The amino acid sequence of a protein can be deduced from the sequence of nucleotides in the corresponding gene. A sequence of three nucleotides specifies one amino acid. A, C, G, and T represent the nucleotide residues of DNA.

Downloaded From Vet-e-Books.com



▲ **Figure 3.22**
Tryptic fingerprint of a 67 kDa serum protein. The numbers over each peak are the mass of the fragment. The number below each mass refer to the residues in Figure 3.23 (Adapted from Detlevuvkaw, Wikipedia entry on peptide mass fingerprinting)

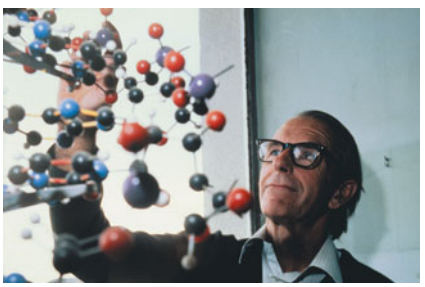
corresponds to a Phe (F) residue (1226.6 − 1079.5 = 147.1), meaning that Phe (F) is the residue at one end of the tryptic peptide. Another large fragment might have a molecular weight of 1098.5 and the difference (1226.6 − 1098.1) is the exact molecular weight of a Lys (K) residue. Thus, Lys (K) is the residue at the other end of the peptide. This has to be the C-terminal end since you know that trypsin cleaves after lysine or arginine residues. You can get the exact sequence of the peptide by analyzing the masses of all fragments in this manner. One of them will have a molecular weight of 258.0 and that is almost certainly the dipeptide Glu-Glu (EE). (The actual analysis is a bit more complicated than this but the principle is the same.)

But it's often not necessary to do the second mass spectrometry analysis in order to identify an unknown protein. Since your unkown protein is from a species whose genome has been sequenced you can simply compare the tryptic fingerprint to the predicted fingerprints of all the proteins encoded by all the genes in the genome. The database consists of a collection of hypothetical peptides produced by analyzing the amino acid sequence of each protein including proteins of unknown function that are known only from their sequence. In most cases your collection of peptide masses from the unknown protein will match only one protein from one of the genes in the database.

In this case, the match is to human serum albumin, a well known serum protein (Figure 3.23). The masses of several of the peptides correspond to the predicted masses of the peptides identified in red in the sequence. Take, for example, the peptide of M_r = 1226.59 in the output from the tryptic fingerprint. This is exactly the predicted mass of the peptide from residues 35–44 (FKDLGEENFK). (Note that the first trypsin cleavage site follows the arginine residue at position 34 and the second cleavage site is after the lysine residue at position 44.)

A single match is not sufficient to identify an unknown protein. In the example shown here there are 21 peptide fragments that match the amino acid sequence of human serum albumin and this is more than sufficient to uniquely identify the protein.

In 1953, Frederick Sanger was the first scientist to determine the complete sequence of a protein (insulin). In 1958, he was awarded a Nobel Prize for this work. Twenty-two years later, Sanger won a second Nobel Prize for pioneering the sequencing of nucleic acids. Today we know the amino acid sequences of thousands of different proteins. These sequences not only reveal details of the structure of individual proteins but also allow researchers to identify families of related proteins and to predict the three-dimensional structure, and sometimes the function, of newly discovered proteins.



▲ **Frederick Sanger (1918–)** Sanger won the Nobel Prize in Chemistry in 1958 for his work on sequencing proteins. He was awarded a second Nobel Prize in Chemistry in 1980 for developing methods of sequencing DNA.

Downloaded From Vet-e-Books.com



▲ **Figure 3.23**
The sequence of human serum albumin. Red residues highlight predicted tryptic peptides and the ones identified in the tryptic fingerprint (Figure 3.22) are underlined.

3.11 Comparisons of the Primary Structures of Proteins Reveal Evolutionary Relationships

In many cases workers have obtained sequences of the same protein from a number of different species. The results show that closely related species contain proteins with very similar amino acid sequences and that proteins from distantly related species are much less similar in sequence. The differences reflect evolutionary change from a common ancestral protein sequence. As more and more sequences were determined it soon became clear that one could construct a tree of similarities and this tree closely resembled the phylogenetic trees constructed from morphological comparisons and the fossil record. The evidence from molecular data was producing independent confirmation of the history of life.

The first sequence-based trees were published almost 50 years ago. One of the earliest examples was the tree for cytochrome *c*—a single polypeptide chain of approximately 104 residues. It provides us with an excellent example of evolution at the molecular level. Cytochrome *c* is found in all aerobic organisms and the protein sequences from distantly related species, such as mammals and bacteria, are similar enough to confidently conclude that the proteins are **homologous**. (Different proteins and genes are defined as homologues if they have descended from a common ancestor. The evidence for homology is based on sequence similarity.)

The first step in revealing evolutionary relationships is to align the amino acid sequences of proteins from a number of species. Figure 3.24 shows an example of such an alignment for cytochrome *c*. The alignment reveals a remarkable conservation of residues at certain positions. For example, every sequence contains a proline at position 30 and a methionine at position 80. In general, conserved residues contribute to the structural stability of the protein or are essential for its function.

There is selection against any amino acid substitutions at these invariant positions. A limited number of substitutions are observed at other sites. In most cases, the allowed substitutions are amino acid residues with similar properties. For example, position 20 can be occupied by leucine, isoleucine, or valine—these are all hydrophobic residues. Similarly, many sites can be occupied by a number of different polar residues. Some positions are highly variable—residues at these sites contribute very little to the structure and function of the protein. The majority of observed amino acid substitutions in homologous proteins are neutral with respect to natural selection. The fixation of substitutions at such positions during evolution is due to random genetic drift and the phylogenetic tree represents proteins that have the same fuction even though they have different amino acid sequences.

The function of cytochrome *c* is described in Section 14.7.

KEY CONCEPT

***Homology* is a conclusion that is based on evidence such as sequence similarity. Homologous proteins descend from a common ancestor. There are degrees of sequence similarity (e.g., 75% identity), but homology is an all-or-nothing conclusion. Something is either homologous or it isn't.**

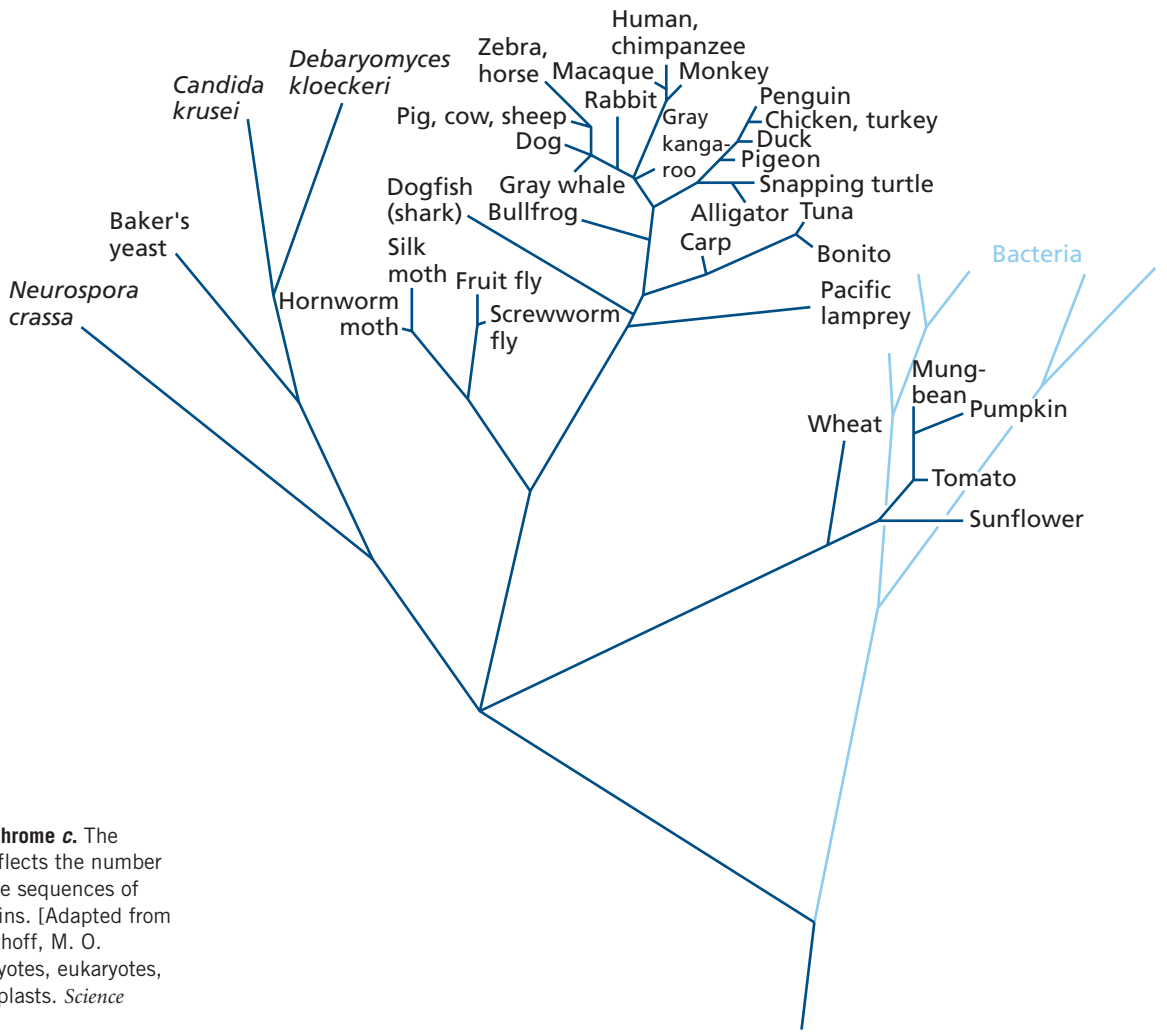
Downloaded From Vet-e-Books.com

Figure 3.24 ►
Cytochrome *c* sequences. The sequences of cytochrome *c* proteins from various species are aligned to show their similarities. In some cases, gaps (signified by hyphens) have been introduced to improve the alignment. The gaps represent deletions and insertions in the genes that encode these proteins. For some species, additional residues at the ends of the sequence have been omitted. Hydrophobic residues are blue and polar residues are red.

The cytochrome *c* sequences of humans and chimpanzees are identical. This is a reflection of their close evolutionary relationship. The monkey and macaque sequences are very similar to the human and chimpanzee sequences as expected since all four species are primates. Similarly, the sequences of the plant cytochrome *c* molecules resemble each other much more than they resemble any of the other sequences.

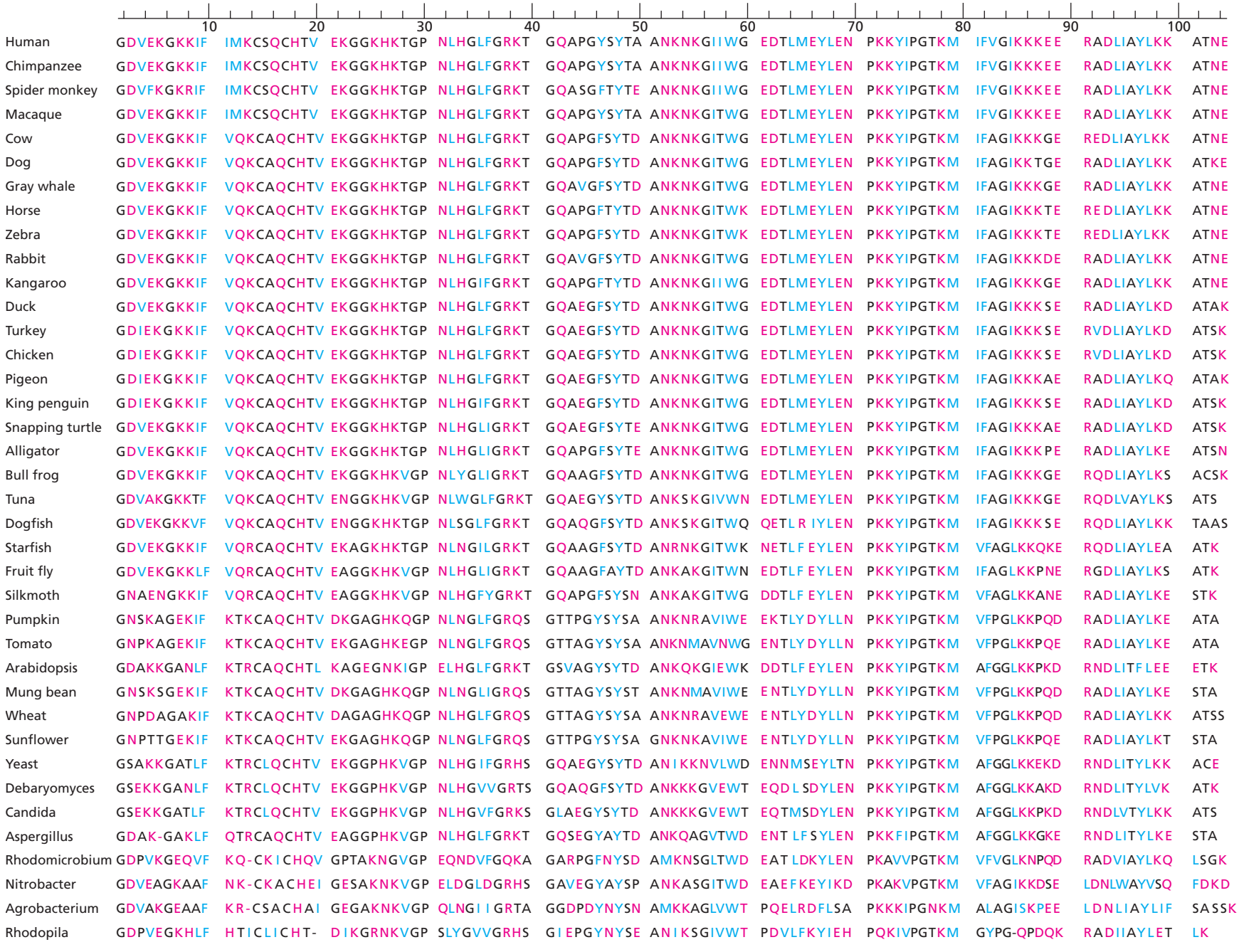
Figure 3.25 illustrates the similarities between cytochrome *c* sequences in different species by depicting them as a tree whose branches are proportional in length to the number of differences in the amino acid sequences of the protein. Species that are closely related cluster together on the same branches of the tree because their proteins are very similar. At great evolutionary distances the number of differences may be very large. For example, the bacterial sequences differ substantially from the eukaryotic sequences reflecting divergence from a common ancestor that lived several billion years ago. The tree clearly reveals the three main kingdoms of eukaryotes—fungi, animals, and plants. (Protist sequences are not included in this tree in order to make it less complicated.)

Note that every species has changed since divurging from their common ancastor.



► **Figure 3.25**
Phylogenetic tree for cytochrome *c*. The length of the branches reflects the number of differences between the sequences of many cytochrome *c* proteins. [Adapted from Schwartz, R. M., and Dayhoff, M. O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* 199:395–403.]

Downloaded From Vet-e-Books.com



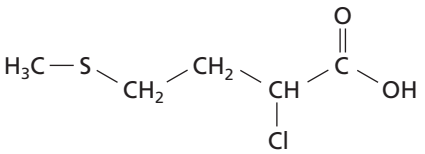
Downloaded From Vet-eBooks.com

Summary

1. Proteins are made from 20 standard amino acids each of which contains an amino group, a carboxyl group, and a side chain, or R group. Except for glycine, which has no chiral carbon, all amino acids in proteins are of the L configuration.
2. The side chains of amino acids can be classified according to their chemical structures—aliphatic, aromatic, sulfur containing, alcohols, bases, acids, and amides. Some amino acids are further classified as having highly hydrophobic or highly hydrophilic side chains. The properties of the side chains of amino acids are important determinants of protein structure and function.
3. Cells contain additional amino acids that are not used in protein synthesis. Some amino acids can be chemically modified to produce compounds that act as hormones or neurotransmitters. Some amino acids are modified after incorporation into polypeptides.
4. At pH 7, the α-carboxyl group of an amino acid is negatively charged ($—COO^{\ominus}$) and the α-amino group is positively charged ($—NH_3^{\oplus}$). The charges of ionizable side chains depend on both the pH and their pK_a values.
5. Amino acid residues in proteins are linked by peptide bonds. The sequence of residues is called the primary structure of the protein.
6. Proteins are purified by methods that take advantage of the differences in solubility, net charge, size, and binding properties of individual proteins.
7. Analytical techniques such as SDS–PAGE and mass spectrometry reveal properties of proteins such as molecular weight.
8. The amino acid composition of a protein can be determined quantitatively by hydrolyzing the peptide bonds and analyzing the hydrolysate chromatographically.
9. The sequence of a polypeptide chain can be determined by the Edman degradation procedure in which the N-terminal residues are successively cleaved and identified.
10. Proteins with very similar amino acid sequences are homologous—they descend from a common ancestor.
11. A comparison of sequences from different species reveals evolutionary relationships.

Problems

1. Draw and label the stereochemical structure of L-cysteine. Indicate whether it is *R* or *S* by referring to Box 3.2 on page 61.
2. Show that the Fischer projection of the common form of threonine (page 60) corresponds to 2*S*, 3*R*-threonine. Draw and name the three other isomers of threonine.
3. Histamine dihydrochloride is administered to melanoma (skin cancer) patients in combination with anticancer drugs because it makes the cancer cells more receptive to the drugs. Draw the chemical structure of histamine dihydrochloride.
4. Dried fish treated with salt and nitrite has been found to contain the mutagen 2-chloro-4-methylthiobutanoic acid (CMBA). From what amino acid is CMBA derived?



5. For each of the following modified amino acid side chains, identify the amino acid from which it was derived and the type of chemical modification that has occurred.
   (a) $—CH_2OPO_3^{2\ominus}$
   (b) $—CH_2CH(COO^{\ominus})_2$
   (c) $—(CH_2)_4—NH—C(O)CH_3$
6. The tripeptide glutathione (GSH) (γ-Glu-Cys-Gly) serves a protective function in animals by destroying toxic peroxides that are generated during aerobic metabolic processes. Draw the chemical structure of glutathione. Note: The γ symbol indicates that the peptide bond between Glu and Cys is formed between the γ-carboxyl of Glu and the amino group of Cys.
7. Melittin is a 26-residue polypeptide found in bee venom. In its monomeric form, melittin is thought to insert into lipid-rich membrane structures. Explain how the amino acid sequence of melittin accounts for this property.

$H_3\overset{\oplus}{N}$-Gly-Ile-Gly-Ala-Val-Leu-Lys-Val-Leu-Thr-Gly-Leu

Pro-Ala-Leu-Ile-Ser-Trp-Ile-Lys-Arg-Lys-Arg-Gln-Gln-NH_2

8. Calculate the isoelectric points of (a) arginine and (b) glutamate.
9. Oxytocin is a nonapeptide (a nine-residue peptide) hormone involved in the milk-releasing response in lactating mammals. The sequence of a synthetic version of oxytocin is shown below. What is the net charge of this peptide at (a) pH 2.0, (b) pH 8.5, and (c) pH 10.7? Assume that the ionizable groups have the pK_a values listed in Table 3.2. The disulfide bond is stable at pH 2.0, pH 8.5, and pH 10.7. Note that the C-terminus is amidated.

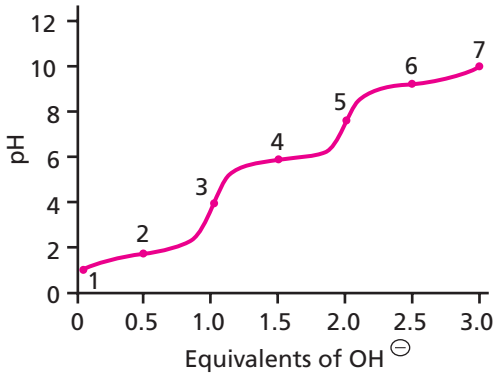
Cys—Phe—Ile—Glu—Asn—Cys—Pro—His—Gly—NH_2 (S—S bridge between Cys 1 and Cys 6)

10. Draw the following structures for compounds that would occur during the Edman degradation procedure: (a) PTC-Leu-Ala, (b) PTH-Ser, (c) PTH-Pro.
11. Predict the fragments that will be generated from the treatment of the following peptide with (a) trypsin, (b) chymotrypsin, and (c) *S. aureus* V8 protease.

Gly-Ala-Trp-Arg-Asp-Ala-Lys-Glu-Phe-Gly-Gln

Downloaded From Vet-e-Books.com

12. The titration curve for histidine is shown below. The pK_a values are 1.8 (—COOH), 6.0 (side chain), and 9.3 (—$NH_3^{\oplus}$).



(a) Draw the structure of histidine at each stage of ionization.
(b) Identify the points on the titration curve that correspond to the four ionic species.
(c) Identify the points at which the average net charge is +2, +0.5 and −1.
(d) Identify the point at which the pH equals the pK_a of the side chain.
(e) Identify the point that indicates complete titration of the side chain.
(f) In what pH ranges would histidine be a good buffer?

13. You have isolated a decapeptide (a 10-residue peptide) called FP, which has anticancer activity. Determine the sequence of the peptide from the following information. (Note that amino acids are separated by commas when their sequence is not known.)

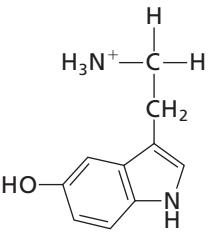
(a) One cycle of Edman degradation of intact FP yields 2 mol of PTH-aspartate per mole of FP.
(b) Treatment of a solution of FP with 2-mercaptoethanol followed by the addition of trypsin yields three peptides with the composition (Ala, Cys, Phe), (Arg, Asp), and (Asp, Cys, Gly, Met, Phe). The intact (Ala, Cys, Phe) peptide yields PTH-cysteine in the first cycle of Edman degradation.
(c) Treatment of 1 mol of FP with carboxypeptidase (which cleaves the C-terminal residue from peptides) yields 2 mol of phenylalanine.
(d) Treatment of the intact pentapeptide (Asp, Cys, Gly, Met, Phe) with CNBr yields two peptides with the composition (homoserine lactone, Asp) and (Cys, Gly, Phe). The (Cys, Gly, Phe) peptide yields PTH-glycine in the first cycle of Edman degradation.

14. A portion of the amino acid sequences for cytochrome *c* from the alligator and bullfrog are given (from Figure 3.24).

	Amino acids 31-50	
Alligator:	NLHGLIGRKT	GQAPGFSYTE
Bullfrog:	NLYGLIGRKT	GQAAGFSYTD

(a) Give an example of a substitution involving similar amino acids.
(b) Give an example of a more radical substitution.

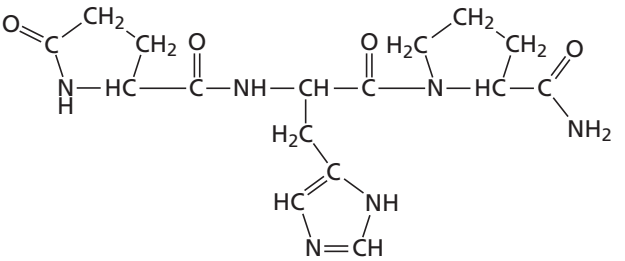
15. Several common amino acids are modified to produce biologically important amines. Serotonin is a biologically important neurotransmitter synthesized in the brain. Low levels of serotonin in the brain have been linked to conditions such as depression, aggression, and hyperactivity. From what amino acid is serotonin derived? Identify the differences in structure between the amino acid and serotonin.



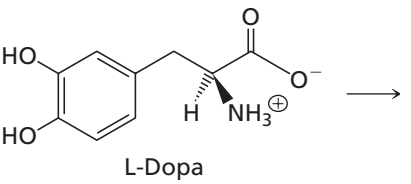
Serotonin

16. The structure of thyrotropin-releasing hormone (TRH) is shown below. TRH is a peptide hormone originally isolated from the extracts of hypothalamus.

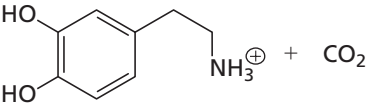
(a) How many peptide bonds are present in TRH?
(b) From what tripeptide is TRH derived?
(c) What result do the modifications have on the charges of the amino and carboxyl-terminal groups?



17. Chirality plays a major role in the development of new pharmaceuticals. People with Parkinson's disease have depleted amounts of dopamine in their brains. In an effort to increase the amount of dopamine in patients, they are given the drug L-dopa which is converted to dopamine in the brain. L-Dopa is marketed in an enantiomerically pure form. (a) Give the *RS* designation for L-dopa. (b) From which amino acid are both L-dopa and dopamine derived?



L-Dopa

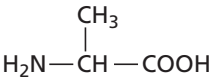


Dopamine

Downloaded From Vet-e-Books.com

18. Generations of biochemistry students have encountered a question like the one below on their final exam.

Calculate the approximate concentration of the uncharged form of alanine (see below) in a 0.01M solution of alanine at (a) pH 2.4 (b) pH 6.15 and (c) pH 9.9.



Can you answer the question without peeking at the solution?

19. A solution of 0.01M alanine is adjusted to pH 2.4 by adding NaOH. What is the concentration of the zwitterion in this solution? What would it be if the pH was 4.0?

Selected Readings

General

Creighton, T. E. (1993). *Proteins: Structures and Molecular Principles*, 2nd ed. (New York: W. H. Freeman), pp. 1–48.

Greenstein, J. P., and Winitz, M. (1961). *Chemistry of the Amino Acids* (New York: John Wiley & Sons).

Kreil, G. (1997). D-Amino Acids in Animal Peptides. *Annu. Rev. Biochem.* 66:337–345.

Meister, A. (1965). *Biochemistry of the Amino Acids*, 2nd ed. (New York: Academic Press).

Protein Purification and Analysis

Hearn, M. T. W. (1987). General strategies in the separation of proteins by high-performance liquid chromatographic methods. *J. Chromatogr.* 418:3–26.

Mann, M., Hendrickson, R.C., and Pandry, A. (2001) Analysis of Proteins and Proteomes by Mass Spectrometry. *Annu. Rev. Biochem.* 70:437–473.

Sherman, L. S., and Goodrich, J. A. (1985). The historical development of sodium dodecyl sulphate–polyacrylamide gel electrophoresis. *Chem. Soc. Rev.* 14:225–236.

Stellwagen, E. (1990). Gel filtration. *Methods Enzymol.* 182:317–328.

Amino Acid Analysis and Sequencing

Doolittle, R. F. (1989). Similar amino acid sequences revisited. *Trends Biochem. Sci.* 14:244–245.

Han, K. -K., Belaiche, D., Moreau, O., and Briand, G. (1985). Current developments in stepwise Edman degradation of peptides and proteins. *Int. J. Biochem.* 17:429–445.

Hunkapiller, M. W., Strickler, J. E., and Wilson, K. J. (1984). Contemporary methodology for protein structure determination. *Science* 226:304–311.

Ozols, J. (1990). Amino acid analysis. *Methods Enzymol.* 182:587–601.

Sanger, F. (1988). Sequences, sequences, and sequences. *Annu. Rev. Biochem.* 57:1–28.

Downloaded From Vet-e-Books.com





Proteins: Three-Dimensional Structure and Function

We saw in the previous chapter that a protein can be described as a chain of amino acids joined by peptide bonds in a specific sequence. However, polypeptide chains are not simply linear but are also folded into compact shapes that contain coils, zigzags, turns, and loops. Over the last 50 years the three-dimensional shapes, or conformations, of thousands of proteins have been determined. A **conformation** is a spatial arrangement of atoms that depends on the rotation of a bond or bonds. The *conformation* of a molecule, such as a protein, can change without breaking covalent bonds whereas the various **configurations** of a molecule can be changed only by breaking and re-forming covalent bonds. (Recall that the L and D forms of amino acids represent different configurations.) Each protein has an astronomical number of potential conformations. Since every amino acid residue has a number of possible conformations and since there are many residues in a protein. Nevertheless, under physiological conditions most proteins fold into a single stable shape known as its native conformation. A number of factors constrain rotation around the covalent bonds in a polypeptide chain in its native conformation. These include the presence of hydrogen bonds and other weak interactions between amino acid residues. The biological function of a protein depends on its native three-dimensional conformation.

A protein may be a single polypeptide chain or it may be composed of several polypeptide chains bound to each other by weak interactions. As a general rule, each polypeptide chain is encoded by a single gene although there are some interesting exceptions to this rule. The size of genes and the polypeptides they encode can vary by more than an order of magnitude. Some polypeptides contain only 100 amino acid residues with a relative molecular mass of about 11,000 (M_r = 11,000) (Recall that the average relative molecular mass of an amino acid residue of a protein is 110.) On the other hand, some very large polypeptide chains contain more than 2000 amino acid residues (M_r = 220,000).

From the intensity of the spots near the centre, we can infer that the protein molecules are relatively dense globular bodies, perhaps joined together by valency bridges, but in any event separated by relatively large spaces which contain water. From the intensity of the more distant spots, it can be inferred that the arrangement of atoms inside the protein molecule is also of a perfectly definite kind, although without the periodicities characterising the fibrous proteins. The observations are compatible with oblate spheroidal molecules of diameters about 25 A. and 35 A., arranged in hexagonal screw-axis. . . . At this stage, such ideas are merely speculative, but now that a crystalline protein has been made to give X-ray photographs, it is clear that we have the means of checking them and, by examining the structure of all crystalline proteins, arriving at a far more detailed conclusion about protein structure than previous physical or chemical methods have been able to give.

—Dorothy Crowfoot Hodgkin (1934)

Top: Bighorn sheep. The skin, wool, and horns are composed largely of fibrous proteins.

Downloaded From Vet-eBooks.com

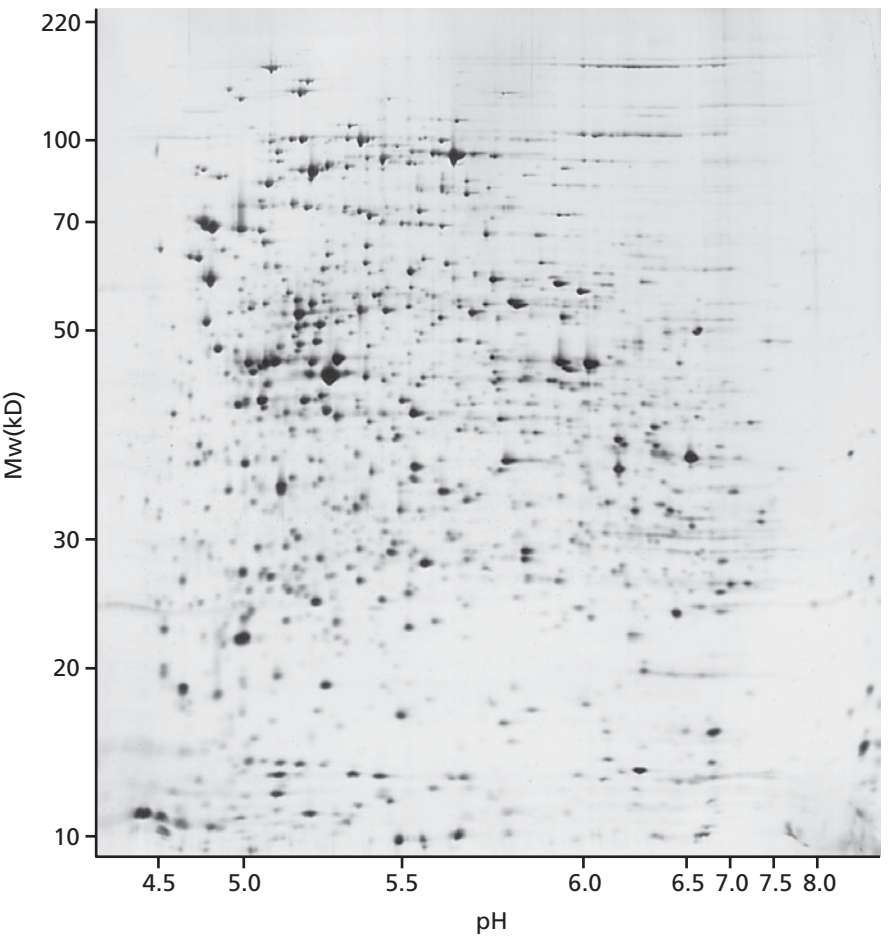
In some species, the size and sequence of every polypeptide can be determined from the sequence of the genome. There are about 4000 different polypeptides in the bacterium *Escherichia coli* with an average size of about 300 amino acid residues (M_r = 33,000). The fruit fly *Drosophila melanogaster* contains about 14,000 different polypeptides with an average size about the same as that in bacteria. Humans and other mammals have about 20,000 different polypeptides. The study of large sets of proteins, such as the entire complement of proteins produced by a cell, is part of a field of study called **proteomics**.

Classes of proteins are described in the introduction to Chapter 3, and the various classes of enzymes are described in Section 5.1.

Proteins come in a variety of shapes. Many are water-soluble, compact, roughly spherical macromolecules whose polypeptide chains are tightly folded. Such proteins—traditionally called **globular proteins**—characteristically have a hydrophobic interior and a hydrophilic surface. They possess indentations or clefts that specifically recognize and transiently bind other compounds. By selectively binding other molecules these proteins serve as dynamic agents of biological action. Many globular proteins are enzymes—the biochemical catalysts of cells. About 31% of the polypeptides in *E. coli* are classical metabolic enzymes such as those described in the next few chapters. Other proteins include various factors, carrier proteins, and regulatory proteins; 12% of the known proteins in *E. coli* fall into these categories.

The terms *globular proteins* and *fibrous proteins* are rarely used in modern scientific publications. There are many proteins that don't fit into either category.

Polypeptides can also be components of large subcellular or extracellular structures such as ribosomes, flagella and cilia, muscle, and chromatin. **Fibrous proteins** are a particular class of structural proteins that provide mechanical support to cells or organisms. Fibrous proteins are typically assembled into large cables or threads. Examples of fibrous proteins are α-keratin, the major component of hair and nails, and collagen, the major protein component of tendons, skin, bones, and teeth. Other examples of structural proteins include the protein components of viruses, bacteriophages, spores, and pollen.



▶ ***Escherichia* coli proteins.** Proteins from *E. coli* cells are separated by two-dimensional gel electrophoresis. In the first dimension, the proteins are separated by a pH gradient where each protein migrates to its isoelectric point. The second dimension separates proteins by size on an SDS–polyacrylamide gel. Each spot corresponds to a single polypeptide. There are about 4000 different proteins in *E. coli*, but some of them are present in very small quantities and can't be seen on this 2-D gel. This figure is from the Swiss-2D PAGE database. You can visit this site and click on any one of the spots to find out more about a particular protein.

Downloaded From Vet-e-Books.com

Many proteins are either integral components of membranes or membrane-associated proteins. Membrane proteins account for at least 16% of the polypeptides in *E. coli* and a much higher percentage in eukaryotic cells.

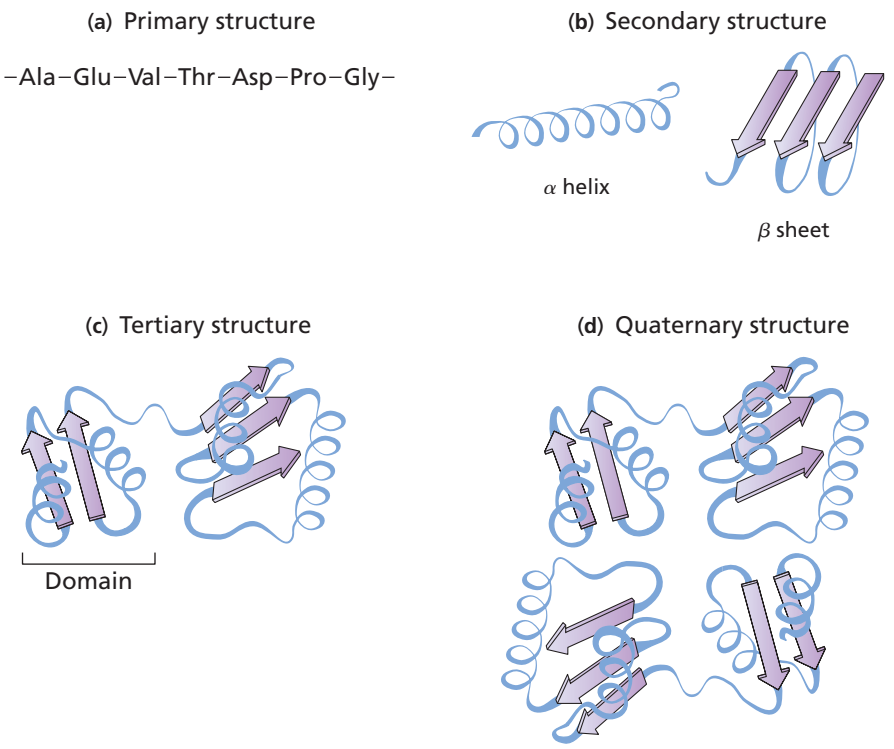
This chapter describes the molecular architecture of proteins. We will explore the conformation of the peptide bond and see that two simple shapes, the α helix and the β sheet, are common structural elements in all classes of proteins. We will describe higher levels of protein structure and discuss protein folding and stabilization. Finally, we will examine how protein structure is related to function using collagen, hemoglobin, and antibodies as examples. Above all, we will learn that proteins have properties beyond those of free amino acids. Chapters 5 and 6 describe the role of proteins as enzymes. The structures of membrane proteins are examined in more detail in Chapter 9 and proteins that bind nucleic acids are covered in Chapters 20 to 22.

4.1 There Are Four Levels of Protein Structure

Individual protein molecules have up to four levels of structure (Figure 4.1). As noted in Chapter 3, **primary structure** describes the linear sequence of amino acid residues in a protein. The three-dimensional structure of a protein is described by three additional levels: secondary structure, tertiary structure, and quaternary structure. The forces responsible for maintaining, or stabilizing, these three levels are primarily noncovalent.

Secondary structure refers to regularities in local conformations maintained by hydrogen bonds between amide hydrogens and carbonyl oxygens of the peptide backbone. The major secondary structures are α helices, β strands, and turns. Cartoons showing the structures of folded proteins usually represent α-helical regions by helices and β strands by broad arrows pointing in the N-terminal to C-terminal direction.

Tertiary structure describes the completely folded and compacted polypeptide chain. Many folded polypeptides consist of several distinct globular units linked by a short stretch of amino acid residues as shown in Figure 4.1c. Such units are called **domains**. Tertiary structures are stabilized by the interactions of amino acid side chains in nonneighboring regions of the polypeptide chain. The formation of tertiary structure brings distant portions of the primary and secondary structures close together.



◀ **Figure 4.1**
Levels of protein structure. (a) The linear sequence of amino acid residues defines the primary structure. **(b)** Secondary structure consists of regions of regularly repeating conformations of the peptide chain such as α helices and β sheets. **(c)** Tertiary structure describes the shape of the fully folded polypeptide chain. The example shown has two domains. **(d)** Quaternary structure refers to the arrangement of two or more polypeptide chains into a multisubunit molecule.

Downloaded From Vet-eBooks.com

Some proteins possess **quaternary structure**—the association of two or more polypeptide chains into a multisubunit, or oligomeric, protein. The polypeptide chains of an oligomeric protein may be identical or different.

4.2 Methods for Determining Protein Structure

As we saw in Chapter 3, the amino acid sequence of polypeptides (i.e., primary structure) can be determined directly by sequencing the protein or indirectly by sequencing the gene. The usual technique for determining the three-dimensional conformation of a protein is X-ray crystallography. In this technique, a beam of collimated (parallel) X rays is aimed at a crystal of protein molecules. Electrons in the crystal diffract the X rays that are then recorded on film or by an electronic detector (Figure 4.2). Mathematical analysis of the diffraction pattern produces an image of the electron clouds surrounding atoms in the crystal. This electron density map reveals the overall shape of the molecule and the positions of each of the atoms in three-dimensional space. By combining these data with the principles of chemical bonding it is possible to deduce the location of all the bonds in a molecule and hence its overall structure. The technique of X-ray crystallography has developed to the point where it is possible to determine the structure of a protein without precise knowledge of the amino acid sequence. In practice, knowledge of the primary structure makes fitting of the electron density map much easier at the stage where chemical bonds between atoms are determined.

Initially, X-ray crystallography was used to study the simple repeating units of fibrous proteins and the structures of small biological molecules. Dorothy Crowfoot Hodgkin was one of the early pioneers in the application of X-ray crystallography to biological molecules. She solved the structure of penicillin in 1947 and developed many of the techniques used in the study of large proteins. Hodgkin received the Nobel Prize in 1964 for determining the structure of vitamin B_{12} and she later published the structure of insulin.

The chief impediment to determining the three-dimensional structure of an entire protein was the difficulty of calculating atomic positions from the positions and intensities of diffracted X-ray beams. Not surprisingly, the development of X-ray crystallography of macromolecules closely followed the development of computers. By 1962, John C. Kendrew and Max Perutz had elucidated the structures of the proteins myoglobin and hemoglobin, respectively, using large and very expensive computers at Cambridge University in the United Kingdom. Their results provided the first insights into the nature of the tertiary structures of proteins and earned them a Nobel Prize in 1962. Since then, the structures of many proteins have been revealed by X-ray crystallography. In recent years, there have been significant advances in the technology due to the availability of inexpensive high-speed computers and improvements in producing focused beams of X rays. The determination of protein structures is now limited mainly

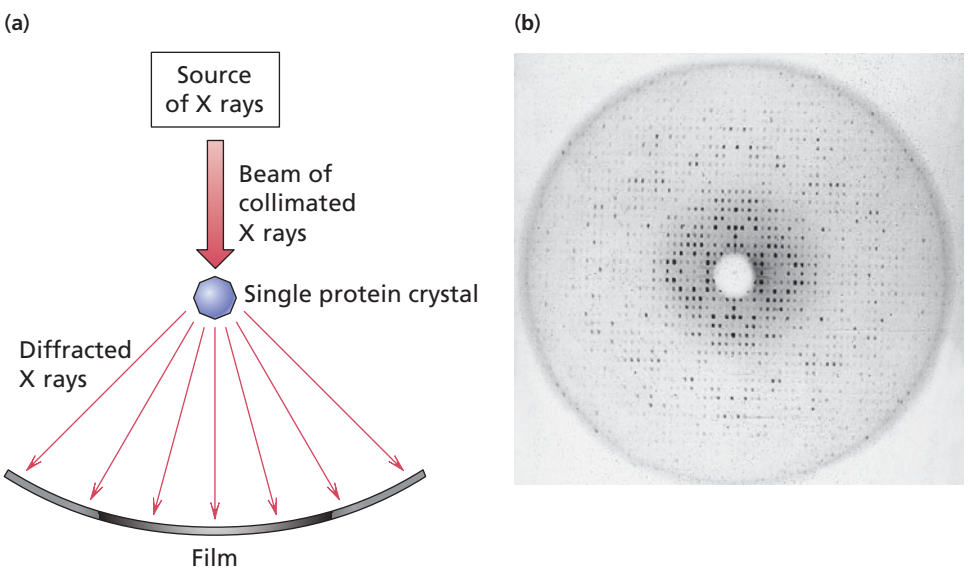


Figure 4.2 ►
X-ray crystallography. (a) Diagram of X rays diffracted by a protein crystal. **(b)** X-ray diffraction pattern of a crystal of adult human deoxyhemoglobin. The location and intensity of the spots are used to determine the three-dimensional structure of the protein.

Downloaded From Vet-e-Books.com



◀ **Bioinformatics in the 1950s.** Bror Strandberg (left) and Dick Dickerson (right) carrying computer tapes from the EDSAC II computer center in Cambridge, UK. The tapes contain X-ray diffraction data from crystals of myoglobin.

by the difficulty of preparing crystals of a quality suitable for X-ray diffraction and even that step is mostly carried out by computer-driven robots.

A protein crystal contains a large number of water molecules and it is often possible to diffuse small ligands such as substrate or inhibitor molecules into the crystal. In many cases, the proteins within the crystal retain their ability to bind these ligands and they often exhibit catalytic activity. The catalytic activity of enzymes in the crystalline state demonstrates that the proteins crystallize in their *in vivo* native conformations. Thus, the protein structures solved by X-ray crystallography are accurate representations of the structures that exist inside cells.

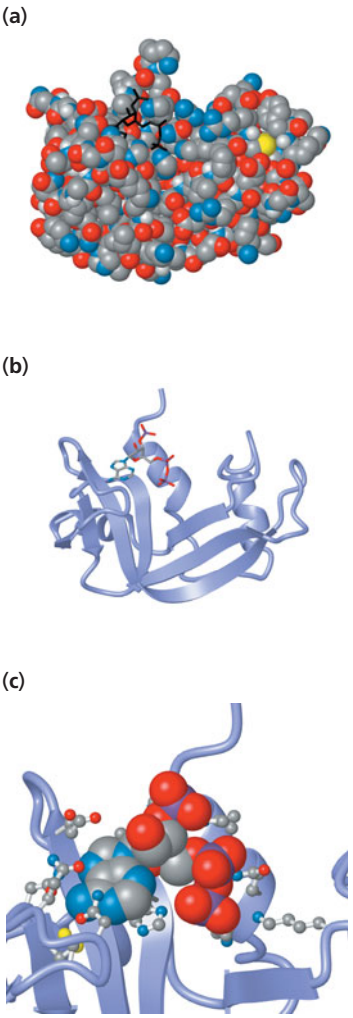
Once the three-dimensional coordinates of the atoms of a macromolecule have been determined, they are deposited in a data bank where they are available to other scientists. Biochemists were among the early pioneers in exploiting the Internet to share data with researchers around the world—the first public domain databases of biomolecular structures and sequences were established in the late 1970s. Many of the images in this text were created using data files from the Protein Data Bank (PDB).

Visit the website for information on how to view three-dimensional structures and retrieve data files.



◀ Max Perutz (1914–2002) (left) and John C. Kendrew (1917–1997) (right). Kendrew determined the structure of myoglobin and Perutz determined the structure of hemoglobin. They shared the Nobel Prize in 1962.

Downloaded From Vet-e-Books.com



▲ **Figure 4.3**
Bovine (*Bos taurus*) ribonuclease A. Ribonuclease A is a secreted enzyme that hydrolyzes RNA during digestion. **(a)** Space-filling model showing a bound substrate analog in black. **(b)** Cartoon ribbon model of the polypeptide chain showing secondary structure. **(c)** View of the substrate-binding site. The substrate analog (5′-diphosphoadenine-3′-phosphate) is depicted as a space-filling model, and the side chains of amino acid residues are shown as ball-and-stick models. [PDB 1AFK]

We will list the PDB filename, or accession number, for every protein structure shown in this text so that you can view the three-dimensional structure on your own computer.

There are many ways of depicting the three-dimensional structure of proteins. Space-filling models (Figure 4.3a) depict each atom as a solid sphere. Such images reveal the dense, closely packed nature of folded polypeptide chains. Space-filling models of structures are used to illustrate the overall shape of a protein and the surface exposed to aqueous solvent. One can easily appreciate that the interior of folded proteins is nearly impenetrable, even by small molecules such as water.

The structure of a protein can also be depicted as a simplified cartoon that emphasizes the backbone of the polypeptide chain (Figure 4.3b). In these models, the amino acid side chains have been eliminated, making it easier to see how the polypeptide folds into a three-dimensional shape. Such models have the advantage of allowing us to see into the interior of the protein, and they also reveal elements of secondary structure such as α helices and β strands. By comparing the structures of different proteins, it is possible to recognize common folds and patterns that can't be seen in space-filling models.

The most detailed models are those that emphasize the structures of the amino acid side chains and the various covalent bonds and weak interactions between atoms (Figure 4.3c). Such detailed models are especially important in understanding how a substrate binds in the active site of an enzyme. In Figure 4.3c, the backbone is shown in the same orientation as in Figure 4.3b.

Another technique for analyzing the macromolecular structure of proteins is nuclear magnetic resonance (NMR) spectroscopy. This method permits the study of proteins in solution and therefore does not require the painstaking preparation of crystals. In NMR spectroscopy, a sample of protein is placed in a magnetic field. Certain atomic nuclei absorb electromagnetic radiation as the applied magnetic field is varied. Because absorbance is influenced by neighboring atoms, interactions between atoms that are close together can be recorded. By combining these results with the amino acid sequence and known structural constraints it is possible to calculate a number of structures that satisfy the observed interactions.

Figure 4.4 depicts the complete set of structures for bovine ribonuclease A—the same protein whose X-ray crystal structure is shown in Figure 4.3. Note that the possible structures are very similar and the overall shape of the molecule is easily seen. In some cases, the set of NMR structures may represent fluctuations, or "breathing," of the protein in solution. The similarity of the NMR and X-ray crystal structures indicates that the protein structures found in crystals accurately represent the structure of the protein in solution but in some cases the structures do not agree. Often this is due to disordered regions that do not show up in the X-ray crystal structure (Section 4.7D). On very rare occasions the protein crystallyzes in a conformation that is not the true native form. The NMR structure is thought to be more accurate.

In general, the NMR spectra for small proteins such as ribonuclease A can be easily solved but the spectrum of a large molecule can be extremely complex. For this reason, it is very difficult to determine the structure of larger proteins but the technique is very powerful for smaller proteins.

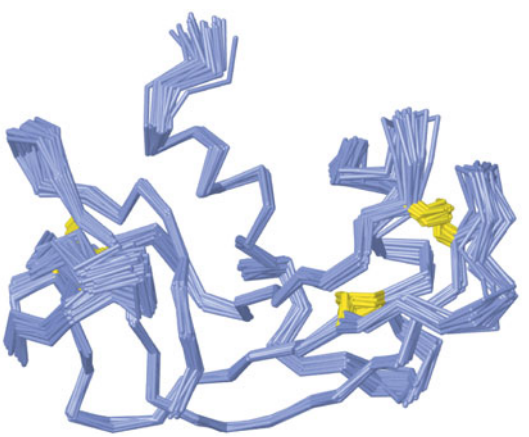


Figure 4.4 ▶
Bovine ribonuclease A NMR structure. The figure combines a set of very similar structures that satisfy the data on atomic interactions. Only the backbone of the polypeptide chain is shown. Compare this structure with that in Figure 4.3b. Note the presence of disulfide bridges (yellow), which are not shown in the images derived from the X-ray crystal structure. [PDB 2AAS].

Downloaded From Vet-e-Books.com

4.3 The Conformation of the Peptide Group

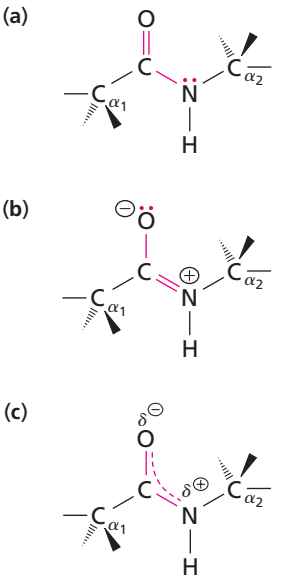
Our detailed study of protein structure begins with the structure of the peptide bonds that link amino acids in a polypeptide chain. The two atoms involved in the peptide bond, along with their four substituents (the carbonyl oxygen atom, the amide hydrogen atom, and the two adjacent α-carbon atoms), constitute the peptide group. X-ray crystallographic analyses of small peptides reveal that the bond between the carbonyl carbon and the nitrogen is shorter than typical C—N single bonds but longer than typical C═N double bonds. In addition, the bond between the carbonyl carbon and the oxygen is slightly longer than typical C═O double bonds. These measurements reveal that peptide bonds have some double-bond properties and can best be represented as a resonance hybrid (Figure 4.5).

Note that the peptide group is polar. The carbonyl oxygen has a partial negative charge and can serve as a hydrogen acceptor in hydrogen bonds. The nitrogen has a partial positive charge, and the —NH group can serve as a hydrogen donor in hydrogen bonds. Electron delocalization and the partial double-bond character of the peptide bond prevent unrestricted free rotation around the C—N bond. As a result, the atoms of the peptide group lie in the same plane (Figure 4.6). Rotation is still possible around each N—Cα bond and each Cα—C bond in the repeating N—Cα—C backbone of proteins. As we will see, restrictions on free rotation around these two additional bonds ultimately determine the three-dimensional conformation of a protein.

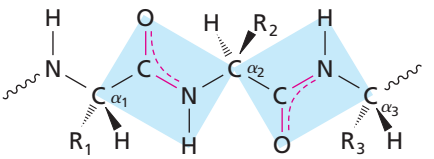
Because of the double-bond nature of the peptide bond, the conformation of the peptide group is restricted to one of two possible conformations, either *trans* or *cis* (Figure 4.7). In the *trans* conformation, the two α-carbons of adjacent amino acid residues are on opposite sides of the peptide bond and at opposite corners of the rectangle formed by the planar peptide group. In the *cis* conformation, the two α-carbons are on the same side of the peptide bond and are closer together. The *cis* and *trans* conformations arise during protein synthesis when the peptide bond is formed by joining amino acids to the growing polypeptide chain. The two conformations are not easily interconverted by free rotation around the peptide bond once it has formed.

The *cis* conformation is less favorable than the extended *trans* conformation because of steric interference between the side chains attached to the two α-carbon atoms. Consequently, nearly all peptide groups in proteins are in the *trans* conformation. Rare exceptions occur, usually at bonds involving the amide nitrogen of proline. Because of the unusual ring structure of proline, the *cis* conformation creates only slightly more steric interference than the *trans* conformation.

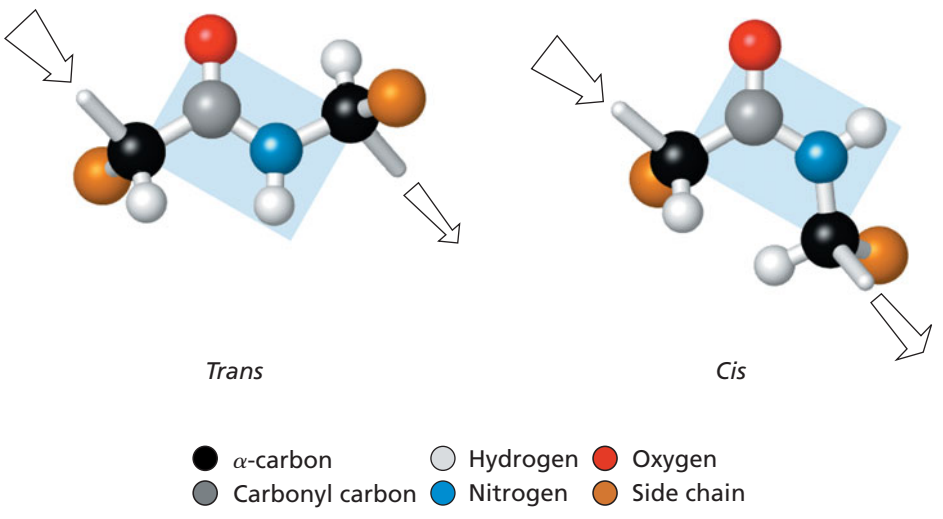
Remember that even though the atoms of the peptide group lie in a plane, rotation is still possible about the N—Cα and Cα—C bonds in the repeating N—Cα—C backbone. This rotation is restricted by steric interference between main-chain and side-chain atoms of adjacent residues. One of the most important restrictions on free rotation is steric interference between carbonyl oxygens on adjacent amino acid residues in the polypeptide



▲ **Figure 4.5**
Resonance structure of the peptide bond. **(a)** In this resonance form, the peptide bond is shown as a single C—N bond. **(b)** In this resonance form, the peptide bond is shown as a double bond. **(c)** The actual structure is best represented as a hybrid of the two resonance forms in which electrons are delocalized over the carbonyl oxygen, the carbonyl carbon, and the amide nitrogen. Rotation around the C—N bond is restricted due to the double-bond nature of the resonance hybrid form.



▲ **Figure 4.6**
Planar peptide groups in a polypeptide chain. A peptide group consists of the N—H and C═O groups involved in formation of the peptide bond, as well as the α-carbons on each side of the peptide bond. Two peptide groups are highlighted in this diagram.



◀ **Figure 4.7**
***Trans* and *cis* conformations of a peptide group.** Nearly all peptide groups in proteins are in the *trans* conformation, which minimizes steric interference between adjacent side chains. The arrows indicate the direction from the N- to the C-terminus.

Downloaded From Vet-e-Books.com

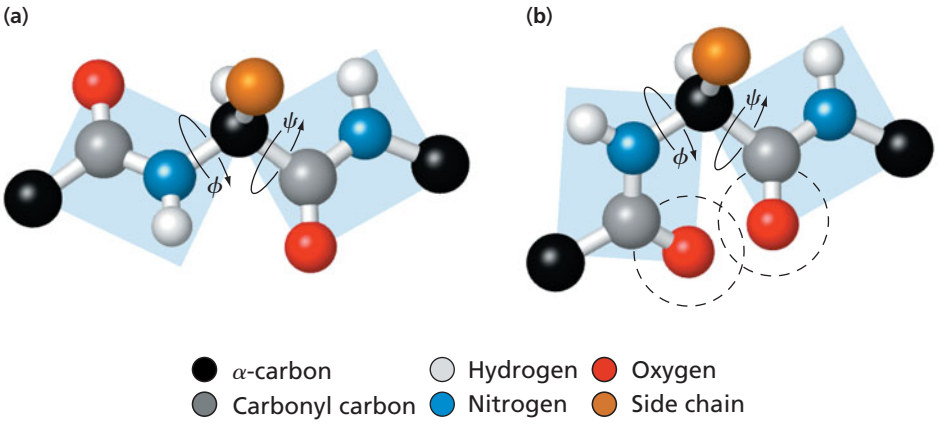
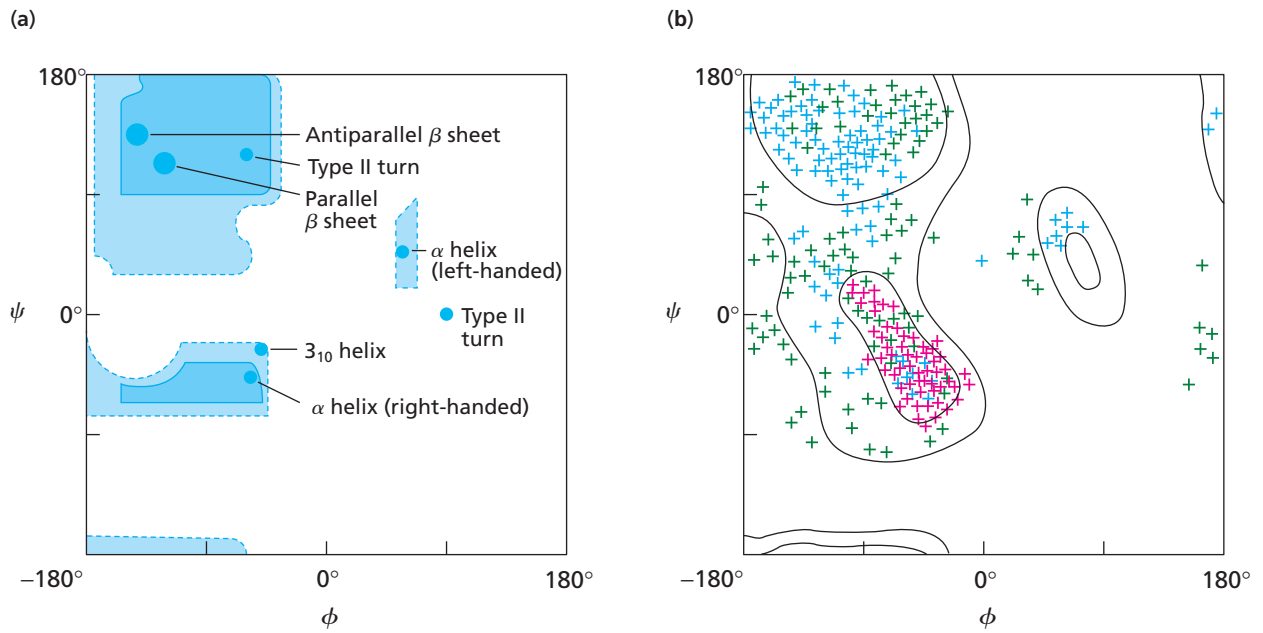


Figure 4.8 ▸
Rotation around the N—Cα and Cα—C bonds that link peptide groups in a polypeptide chain. **(a)** Peptide groups in an extended conformation. **(b)** Peptide groups in an unstable conformation caused by steric interference between carbonyl oxygens of adjacent residues. The van der Waals radii of the carbonyl oxygen atoms are shown by the dashed lines. The rotation angle around the N—Cα bond is called φ (phi), and that around the Cα—C bond is called ψ (psi). The substituents of the outer α-carbons have been omitted for clarity.

chain (Figure 4.8). The presence of bulky side chains also restricts free rotation around the N—Cα and Cα—C bonds. Proline is a special case—rotation around the N—Cα bond is constrained because it is part of the pyrrolidine ring structure of proline.

The rotation angle around the N—Cα bond of a peptide group is designated φ (phi), and that around the Cα—C bond is designated ψ (psi). The peptide bond angle is ω (omega). Because rotation around peptide bonds is hindered by their double-bond character, most of the conformation of the backbone of a polypeptide can be described by φ and ψ. Each of these angles is defined by the relative positions of four atoms of the backbone. Clockwise angles are positive, and counterclockwise angles are negative, with each having a 180° sweep. Thus, each of the rotation angles can range from −180° to +180°.

The biophysicist G. N. Ramachandran and his colleagues constructed space-filling models of peptides and made calculations to determine which values of φ and ψ are sterically permitted in a polypeptide chain. Permissible angles are shown as shaded regions in **Ramachandran plots** of φ versus ψ. Figure 4.9a shows the results of theoretical calculations—the dark, shaded regions represent permissible angles for most residues, and the lighter areas cover the φ and ψ values for smaller amino acid residues where the



▲ **Figure 4.9**
Ramachandran plot. **(a)** Solid lines indicate the range of permissible φ and ψ values based on molecular models. Dashed lines give the outer limits for an alanine residue. Large blue dots correspond to values of φ and ψ that produce recognizable conformations such as the α helix and β sheets. The positions shown for the type II turn are for the second and third residues. The white portions of the plot correspond to values of φ and ψ that were predicted to occur rarely. **(b)** Observed φ and ψ values in known structures. Crosses indicate values for typical residues in a single protein. Residues in an α helix are shown in red, β-strand residues are blue, and others are green.

Downloaded From Vet-e-Books.com

R groups don't restrict rotation. Blank areas on a Ramachandran plot are nonpermissible areas, due largely to steric hindrance. The conformations of several types of ideal secondary structure fall within the shaded areas, as expected.

Another version of a Ramachandran plot is shown in Figure 4.9b. This plot is based on the observed φ and ψ angles of hundreds of proteins whose structures are known. The enclosed inner regions represent angles that are found very frequently, and the outer enclosed regions represent angles that are less frequent. Typical observed angles for α helices, β sheets, and other structures in a protein are plotted. The most important difference between the theoretical and observed Ramachandran plots is in the region around 0°φ and −90°ψ. This region should not be permitted according to the modeling studies but there are many examples of residues with these angles. It turns out that steric clashes are prevented in these regions by allowing a small amount of rotation around the peptide bond. The peptide group does not have to be exactly planar—a little bit of wiggle is permitted!

Some bulky amino acid residues have smaller permitted areas. Proline is restricted to a φ value of about −60° to −77° because its N—Cα bond is constrained by inclusion in the pyrrolidine ring of the side chain. In contrast, glycine is exempt from many steric restrictions because it lacks a β-carbon. Thus, glycine residues have greater conformational freedom than other residues and have φ and ψ values that often fall outside the shaded regions of the Ramachandran plot.

KEY CONCEPT

The three-dimensional conformation of a polypeptide backbone is defined by the φ (phi) and ψ (psi) angles of rotation around each peptide group.

BOX 4.1 FLOWERING IS CONTROLLED BY *CIS/TRANS* SWITCHES

Almost all peptide groups adopt the *trans* conformation since that is the one favored during protein synthesis. It is much more stable than the *cis* conformation (with one exception). Spontaneous switching to the *cis* conformation is very rare and it is almost always accompanied by loss of function since the structure of the protein is severely affected.

However, the activity of some proteins is actually *regulated* by conformation changes due to *cis/trans* isomerization. The change in peptide group conformation invariably takes place at proline residues because the *cis* conformation is almost as stable as the *trans* conformation. This is the one exception to the rule.

Specific enzymes, called peptidyl prolyl *cis/trans* isomerases, catalyze the interconversion of *cis* and *trans* conformation at proline residues by transiently destabilizing the resonance hybrid structure of the peptide bond and allowing rotation. One important class of these enzymes recognizes Ser-Pro and Thr-Pro bonds whenever the serine and threonine residues are phosphorylated. Phosphorylation of amino acid residues is an important mechanism of regulation by covalent modification (see Section 5.9D). The gene for this type of peptidyl prolyl *cis/trans* isomerase is called *Pin1* and it is present in all eukaryotes.

In the small flowering plant, *Arabidopsis thalianna*, *Pin1* protein acts on some transcription factors that control the timing of flowering. When threonine residues are phosphorylated, the transcription factors are recognized by *Pin1* and the conformation of the Thr-Pro bond is switched from *trans* to *cis*. The resulting conformational change in the structure of the protein leads to activation of the transcription factors and transcription of the genes required for producing flowers. Flowering is considerably delayed when the synthesis of peptidyl prolyl *cis/trans* isomerase is inhibited by mutations in the *Pin1* gene.

In humans the *cis/trans* isomerase encoded by *Pin1* plays a role in regulating gene expression by modifying RNA polymerase, transcription factors, and other proteins. Mutations in this gene have been implicated in several hereditary diseases. The structure of human peptidyl prolyl *cis/trans* isomerase is shown in Figure 4.23e.



▲ *Arabidopsis thalianna*, also known as thale cress or mouse-ear cress, is a relative of mustard. It is a favorite model organism in plant biology because it is easy to grow in the laboratory.

Downloaded From Vet-e-Books.com

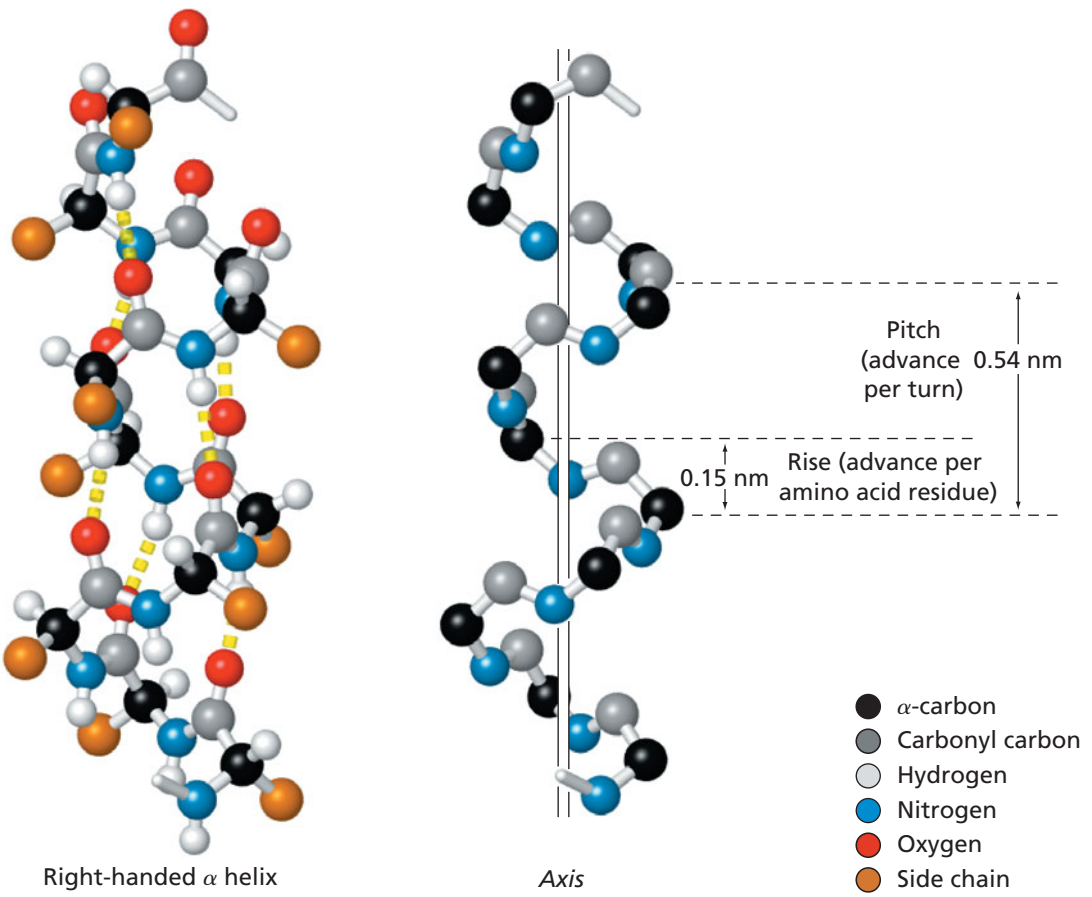


▲ Linus Pauling (1901–1994), winner of the Nobel Prize in Chemistry in 1954 and the Nobel Peace Prize in 1962.

4.4 The α Helix

The α-helical conformation was proposed in 1950 by Linus Pauling and Robert Corey. They considered the dimensions of peptide groups, possible steric constraints, and opportunities for stabilization by formation of hydrogen bonds. Their model accounted for the major repeat observed in the structure of the fibrous protein α-keratin. This repeat of 0.50 to 0.55 nm turned out to be the pitch (the axial distance per turn) of the α helix. Max Perutz added additional support for the structure when he observed a secondary repeating unit of 0.15 nm in the X-ray diffraction pattern of α-keratin. The 0.15 nm repeat corresponds to the rise of the α helix (the distance each residue advances the helix along its axis). Perutz also showed that the α helix was present in hemoglobin, confirming that this conformation was present in more complex globular proteins.

In theory, an α helix can be either a right- or a left-handed screw. The α helices found in proteins are almost always right-handed, as shown in Figure 4.10. In an ideal α helix, the pitch is 0.54 nm, the rise is 0.15 nm, and the number of amino acid residues required for one complete turn is 3.6 (i.e., approximately 3 2/3 residues: one carbonyl group, three N—Cα—C units, and one nitrogen). Most α helices are slightly distorted in proteins but they generally have between 3.5 and 3.7 residues per turn.



▲ **Figure 4.10**
α Helix. A region of α-helical secondary structure is shown with the N-terminus at the bottom and the C-terminus at the top of the figure. Each carbonyl oxygen forms a hydrogen bond with the amide hydrogen of the fourth residue further toward the C-terminus of the polypeptide chain. The hydrogen bonds are approximately parallel to the long axis of the helix. Note that all the carbonyl groups point toward the C-terminus. In an ideal α helix, equivalent positions recur every 0.54 nm (the pitch of the helix), each amino acid residue advances the helix by 0.15 nm along the long axis of the helix (the rise), and there are 3.6 amino acid residues per turn. In a right-handed helix the backbone turns in a clockwise direction when viewed along the axis from its N-terminus. If you imagine that the right-handed helix is a spiral staircase, you will be turning to the right as you walk *down* the staircase.

Downloaded From Vet-eBooks.com

Within an α helix, each carbonyl oxygen (residue *n*) of the polypeptide backbone is hydrogen-bonded to the backbone amide hydrogen of the fourth residue further toward the C-terminus (residue *n* + 4). (The three amino groups at one end of the helix and the three carbonyl groups at the other end lack hydrogen-bonding partners within the helix.) Each hydrogen bond closes a loop containing 13 atoms—the carbonyl oxygen, 11 backbone atoms, and the amide hydrogen. Thus, an α helix can also be called a 3.6_{13} helix based on its pitch and hydrogen-bonded loop size. The hydrogen bonds that stabilize the helix are nearly parallel to the long axis of the helix.

The φ and ψ angles of each residue in an α helix are similar. They cluster around a stable region of the Ramachandran plot centered at a φ value of −57° and a ψ value of −47° (Figure 4.9). The similarity of these values is what gives the α helix a regular, repeating structure. The intramolecular hydrogen bonds between residues *n* and *n* + 4 tend to "lock in" rotation around the N—Cα and Cα—C bonds restricting the φ and ψ angles to a relatively narrow range.

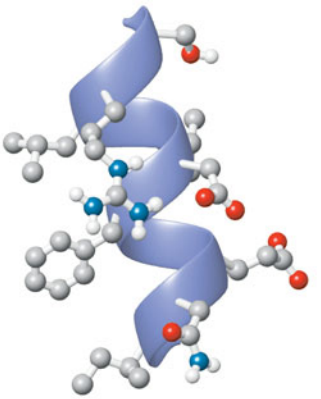
A single intrahelical hydrogen bond would not provide appreciable structural stability but the cumulative effect of many hydrogen bonds within an α helix stabilizes this conformation. Hydrogen bonds between amino acid residues are especially stable in the hydrophobic interior of a protein where water molecules do not enter and therefore cannot compete for hydrogen bonding. In an α helix, all the carbonyl groups point toward the C-terminus. The entire helix is a dipole with a positive N-terminus and a negative C-terminus since each peptide group is polar and all the hydrogen bonds point in the same direction.

The side chains of the amino acids in an α helix point outward from the cylinder of the helix and they are not involved in the hydrogen bonds that stabilize the α helix (Figure 4.11). However, the identity of the side chains affects the stability in other ways. Because of this, some amino acid residues are found in α-helical conformations more often than others. For example, alanine has a small, uncharged side chain and fits well into the α-helical conformation. Alanine residues are prevalent in the α helices of all classes of proteins. In contrast, tyrosine and asparagine with their bulky side chains are less common in α helices. Glycine, whose side chain is a single hydrogen atom, destabilizes α-helical structures since rotation around its α-carbon is so unconstrained. For this reason, many α helices begin or end with glycine residues. Proline is the least common residue in an α helix because its rigid cyclic side chain disrupts the right-handed helical conformation by occupying space that a neighboring residue of the helix would otherwise occupy. In addition, because it lacks a hydrogen atom on its amide nitrogen, proline cannot fully participate in intrahelical hydrogen bonding. For these reasons, proline residues are found more often at the ends of α helices than in the interior.

Proteins vary in their α-helical content. In some proteins most of the residues are in α helices, whereas other proteins contain very little α-helical structure. The average content of α helix in the proteins that have been examined is 26%. The length of a helix in a protein can range from about 4 or 5 residues to more than 40—the average is about 12.

Many α helices have hydrophilic amino acids on one face of the helix cylinder and hydrophobic amino acids on the opposite face. The amphipathic nature of the helix is easy to see when the amino acid sequence is drawn as a spiral called a helical wheel. The α helix shown in Figure 4.11 can be drawn as a helical wheel representing the helix viewed along its axis. Because there are 3.6 residues per turn of the helix, the residues are plotted every 100° along the spiral (Figure 4.12). Note that the helix is a right-handed screw and it is terminated by a glycine residue at the C-terminal end. The hydrophilic residues (asparagine, glutamate, aspartate, and arginine) tend to cluster on one side of the helical wheel.

Amphipathic helices are often located on the surface of a protein with the hydrophilic side chains facing outward (toward the aqueous solvent) and the hydrophobic side chains facing inward (toward the hydrophobic interior). For example, the helix shown in Figures 4.11 and 4.12 is on the surface of the water-soluble liver enzyme alcohol dehydrogenase with the side chains of the first, fifth, and eighth residues



▲ **Figure 4.11**
View of a right-handed α helix. The blue ribbon indicates the shape of the polypeptide backbone. All the side chains, shown as ball-and-stick models, project outward from the helix axis. This example is from residues Ile-355 (bottom) to Gly-365 (top) of horse liver alcohol dehydrogenase. Some hydrogen atoms are not shown. [PDB 1ADF].



▲ **A right-handed α helix.** This helix was created by Julian Voss-Andreae. It stands outside Linus Panling's childhood home in Portland, Oregon, United States.

Downloaded From Vet-eBooks.com

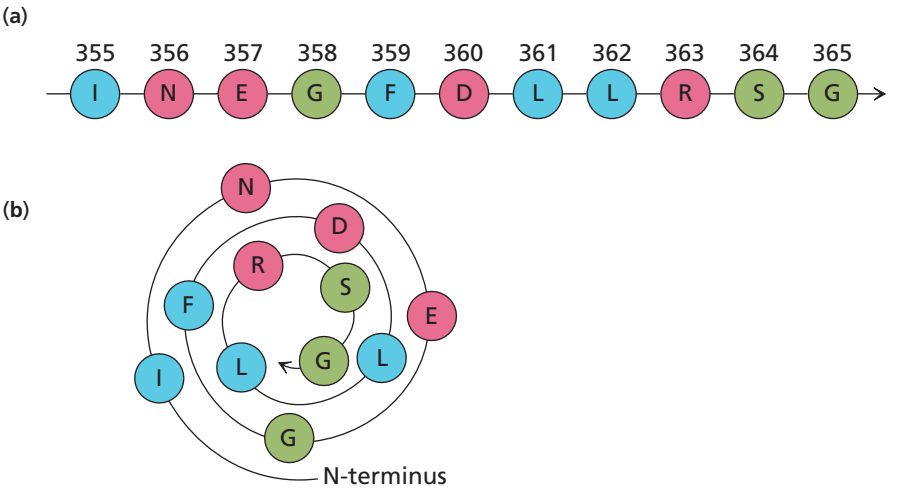


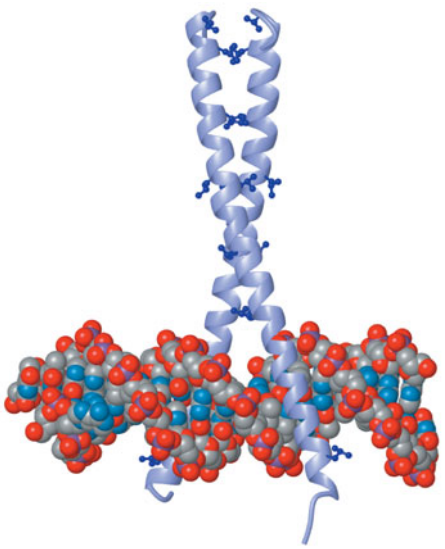
Figure 4.12 ►
α helix in horse liver alcohol dehydrogenase. Highly hydrophobic residues are blue, less hydrophobic residues are green, and highly hydrophilic residues are red. **(a)** Sequence of amino acids. **(b)** Helical wheel diagram.

The known frequencies of various amino acid residues in α helices are used to predict the secondary structure based on the primary sequence alone.

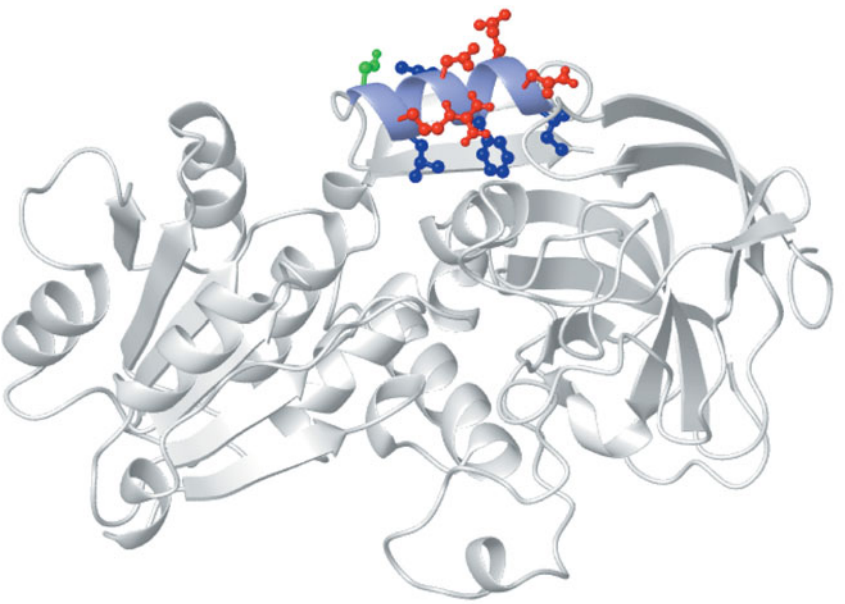
(isoleucine, phenylalanine, and leucine, respectively) buried in the protein interior (Figure 4.13).

There are many examples of two amphipathic α helices that interact to produce an extended coiled-coil structure where the two α helices wrap around each other with their hydrophobic faces in contact and their hydrophilic faces exposed to solvent. A common structure in DNA-binding proteins is called a leucine zipper (Figure 4.14). The name refers to the fact that two α helices are "zippered" together by the hydrophobic interactions of leucine residues (and other hydrophobic residues) on one side of an amphipathic helix. The ends of the helices form the DNA-binding region of the protein.

Some proteins contain a few short regions of a 3_{10} helix. Like the α helix, the 3_{10} helix is right-handed. The carbonyl oxygen of a 3_{10} helix forms a hydrogen bond with the amide hydrogen of residue $n + 3$ (as opposed to residue $n + 4$ in an α helix) so the 3_{10} helix has a tighter hydrogen-bonded ring structure than the α helix—10 atoms rather than 13—and has fewer residues per turn (3.0) and a longer pitch (0.60 nm) (Figure 4.15).



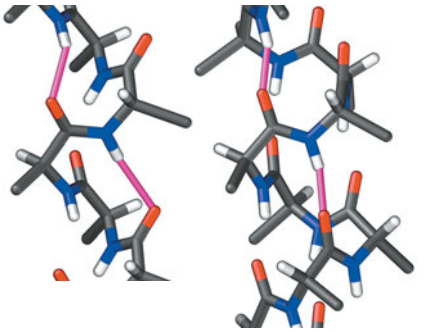
▲ Figure 4.14
Leucine zipper region of yeast (*Saccharomyces cerevisiae*). GCN4 protein bound to DNA. GCN4 is a transcription regulatory protein that binds to specific DNA sequences. The DNA-binding region consists of two amphipathic α helices, one from each of the two subunits of the protein. The side chains of leucine residues are shown in a darker blue than the ribbon. Only the leucine zipper region of the protein is shown in the figure. [PDB 1YSA].



▲ Figure 4.13
Horse (*Equus ferus*) liver alcohol dehydrogenase. The amphipathic α helix is highlighted. The side chains of highly hydrophobic residues are shown in blue, less hydrophobic residues are green, and charged residues are shown in red. Note that the side chains of the hydrophobic residues are directed toward the interior of the protein and that the side chains of charged residues are exposed to the surface. [PDB 1ADF].

Downloaded From Vet-e-Books.com

The 3_{10} helix is slightly less stable than the α helix because of steric hindrances and the awkward geometry of its hydrogen bonds. When a 3_{10} helix occurs, it is usually only a few residues in length and often is the last turn at the C-terminal end of an α helix. Because of its different geometry, the φ and ψ angles of residues in a 3_{10} helix occupy a different region of the Ramachandran plot than the residues of an α helix (Figure 4.9).



▲ **Figure 4.15**
The 3_{10} helix. In the 3_{10} helix (left) hydrogen bonds (pink) form between the amide group of one residue and the carbonyl oxygen of a residue three positions away. In an α helix (right) the carbonyl group bonds to an amino acid residue four positions away.

4.5 β Strands and β Sheets

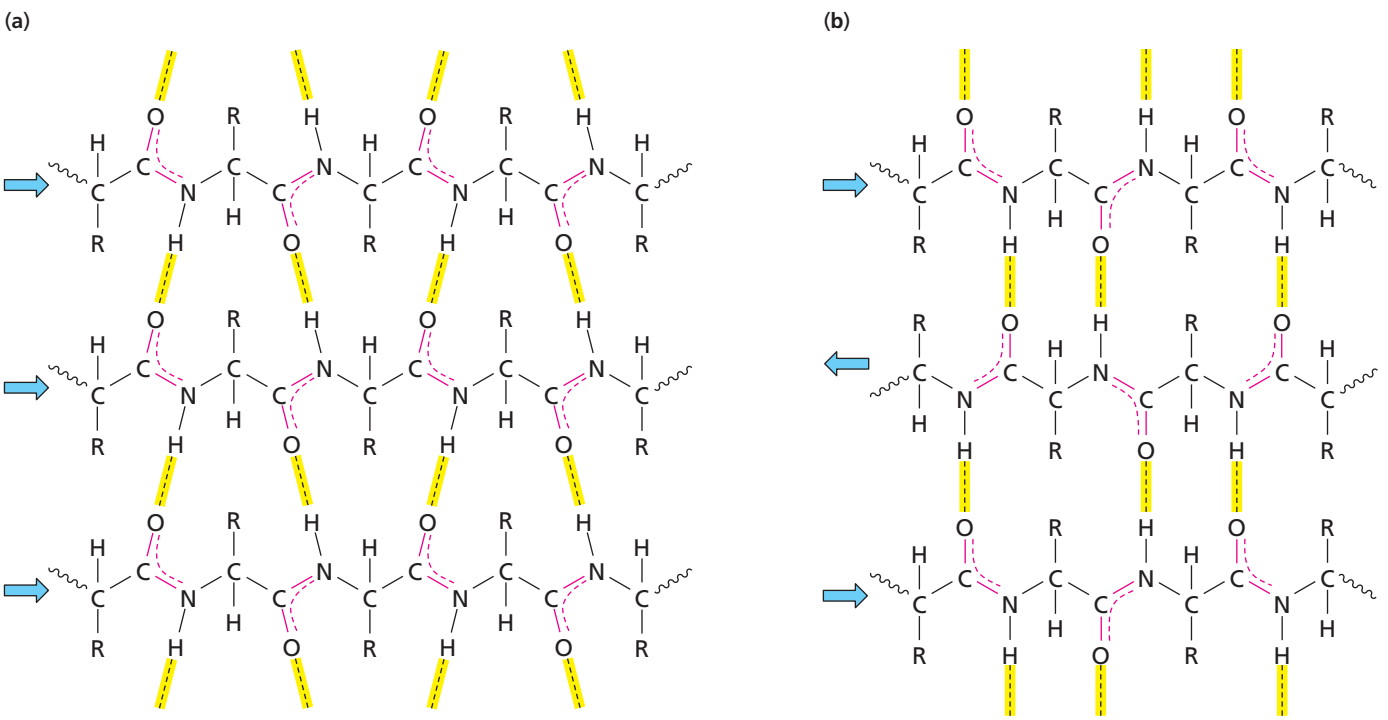
The other common secondary structure is called β structure, a class that includes β strands and β sheets. **β Strands** are portions of the polypeptide chain that are almost fully extended. Each residue in a β strand accounts for about 0.32 to 0.34 nm of the overall length in contrast to the compact coil of an α helix where each residue corresponds to 0.15 nm of the overall length. When multiple β strands are arranged side-by-side they form **β sheets**, a structure originally proposed by Pauling and Corey at the same time they developed a theoretical model of the α helix.

Proteins rarely contain isolated β strands because the structure by itself is not significantly more stable than other conformations. However, β sheets are stabilized by hydrogen bonds between carbonyl oxygens and amide hydrogens on adjacent β strands. Thus, in proteins, the regions of β structure are almost always found in sheets.

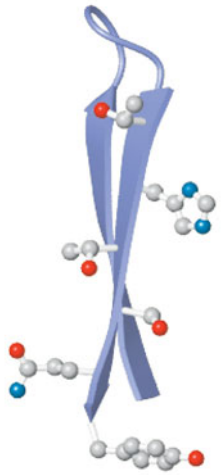
The hydrogen-bonded β strands can be on separate polypeptide chains or on different segments of the same chain. The β strands in a sheet can be either parallel (running in the same N- to C-terminal direction) (Figure 4.16a) or antiparallel (running in opposite N- to C-terminal directions) (Figure 4.16b). When the β strands are antiparallel, the hydrogen bonds are nearly perpendicular to the extended polypeptide chains. Note that in the antiparallel β sheet, the carbonyl oxygen and the amide hydrogen atoms of one residue form hydrogen bonds with the amide hydrogen and carbonyl oxygen of a single residue in the other strand. In the parallel arrangement, the hydrogen bonds are not perpendicular to the extended chains and each residue forms hydrogen bonds with the carbonyl and amide groups of two different residues on the adjacent strand.

Parallel sheets are less stable than antiparallel sheets, possibly because the hydrogen bonds are distorted in the parallel arrangement. The β sheet is sometimes called a **β pleated sheet** since the planar peptide groups meet each other at angles, like the folds of an accordion. As a result of the bond angles between peptide groups, the amino acid

▼ **Figure 4.16**
β Sheets. Arrows indicate the N- to C-terminal direction of the peptide chain. **(a)** Parallel β sheet. The hydrogen bonds are evenly spaced but slanted. **(b)** Antiparallel β sheet. The hydrogen bonds are essentially perpendicular to the β strands, and the space between hydrogen-bonded pairs is alternately wide and narrow.



Downloaded From Vet-eBooks.com



▲ **Figure 4.17**
View of two strands of an antiparallel β sheet from influenza virus A neuraminidase. Only the side chains of the front β strand are shown. The side chains alternate from one side of the β strand to the other side. Both strands have a right-handed twist. [PDB 1BJI]

side chains point alternately above and below the plane of the sheet. A typical β sheet contains from two to as many as 15 individual β strands. Each strand has an average of six amino acid residues.

The β strands that make up β sheets are often twisted and the sheet is usually distorted and buckled. The three-dimensional view of the β sheet of ribonuclease A (Figure 4.3) shows a more realistic view of β sheets than the idealized structures in Figure 4.16.

A view of two strands of a small β sheet is shown in Figure 4.17. The side chains of the amino acid residues in the front strand alternately project to the left and to the right of (i.e., above and below) the β strand, as described above. Typically, β strands twist slightly in a right-hand direction; that is, they twist clockwise as you look along one strand.

The φ and ψ angles of the bonds in a β strand are restricted to a broad range of values occupying a large, stable region in the upper left-hand corner of the Ramachandran plot. The typical angles for residues in parallel and antiparallel strands are not identical (see Figure 4.9). Because most β strands are twisted, the φ and ψ angles exhibit a broader range of values than those seen in the more regular α helix.

Although we usually think of β sheets as examples of secondary structure this is not, strictly speaking, correct. In many cases, the individual β strands are located in different regions of the protein and only come together to form the β sheet when the protein adopts its final tertiary conformation. Sometimes the quaternary structure of a protein gives rise to a large β sheet. Some proteins are almost entirely β sheets but most proteins have a much lower β-strand content.

In the previous section we noted that amphipathic α helices have hydrophobic side chains that project outward on one side of the helix. This is the side that interacts with the rest of the protein creating a series of hydrophobic interactions that help stabilize the tertiary structure. The side chains of β sheets project alternately above and below the plane of the β strands. One surface may consist of hydrophobic side chains that allow the β sheet to lie on top of other hydrophobic residues in the interior of the protein.

An example of such hydrophobic interactions between two β sheets is seen in the structure of the coat protein of grass pollen grains (Figure 4.18a). This protein is the major allergen affecting people who are allergic to grass pollen. One surface of each β sheet contains hydrophobic side chains and the opposite surface has hydrophilic side chains. The two hydrophobic surfaces interact to form the hydrophobic core of the protein and the hydrophilic surfaces are exposed to solvent as shown in Figure 4.18b. This is an example of a β sandwich, one of several arrangements of secondary structural elements that are covered in more detail in the section on tertiary structure (Section 4.7).

KEY CONCEPT

There are only three different kinds of common secondary structure: α helix, β strand, and turns.

4.6 Loops and Turns



▲ **U-turns are allowed in proteins.**

In both an α helix and a β strand there are consecutive residues with a similar conformation that is repeated throughout the structure. Proteins also contain stretches of nonrepeating three-dimensional structure. Most of these non-repeating regions of secondary structure can be characterized as loops or turns since they cause directional changes in the polypeptide backbone. The conformations of peptide groups in nonrepetitive regions are constrained just as they are in repetitive regions. They have φ and ψ values that are usually well within the permitted regions of the Ramachandran plot and often close to the values of residues that form α helices or β strands.

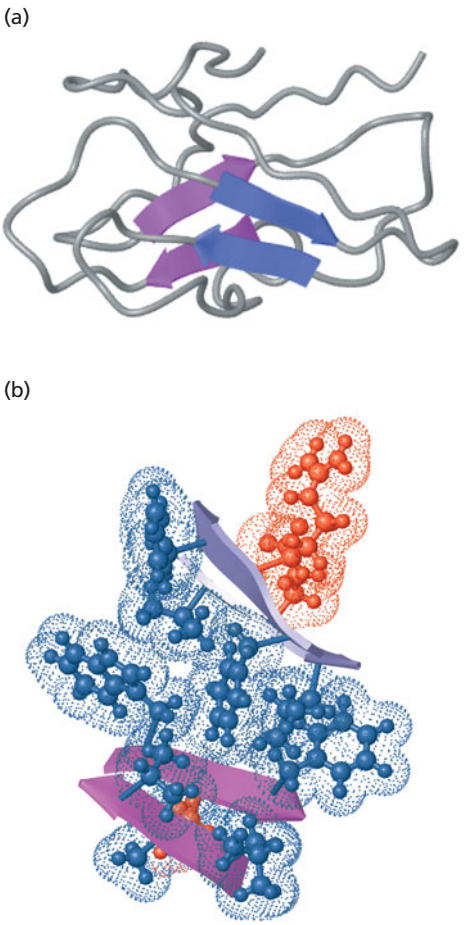
Loops and turns connect α helices and β strands and allow the polypeptide chain to fold back on itself producing the compact three-dimensional shape seen in the native structure. As much as one-third of the amino acid residues in a typical protein are found in such nonrepetitive structures. **Loops** often contain hydrophilic residues and are usually found on the surfaces of proteins where they are exposed to solvent and form hydrogen bonds with water. Some loops consist of many residues of extended nonrepetitive structure. About 10% of the residues can be found in such regions.

Downloaded From Vet-e-Books.com

Loops containing only a few (up to five) residues are referred to as **turns** if they cause an abrupt change in the direction of a polypeptide chain. The most common types of tight turns are called **reverse turns**. They are also called **β turns** because they often connect different antiparallel β strands. (Recall that in order to create a β sheet the polypeptide must fold so that two or more regions of β strand are adjacent to one another as shown in Figure 4.17.) This terminology is misleading since β turns can also connect α helices or an α helix and a β strand.

There are two common types of β turn, designated type I and type II. Both types of turn contain four amino acid residues and are stabilized by hydrogen bonding between the carbonyl oxygen of the first residue and the amide hydrogen of the fourth residue (Figure 4.19). Both type I and type II turns produce an abrupt (usually about 180°) change in the direction of the polypeptide chain. In type II turns, the third residue is glycine about 60% of the time. Proline is often the second residue in both types of turns.

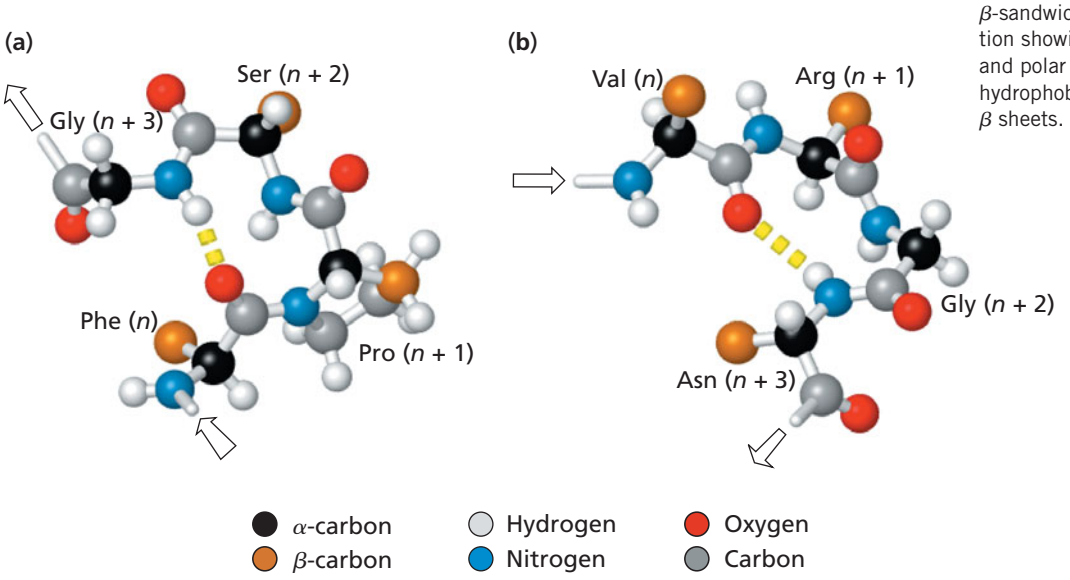
Proteins contain many turn structures. They all have internal hydrogen bonds that stabilize the structure and that's why they can be considered a form of secondary structure. Turns make up a significant proportion of the structure in many proteins. Some of the bonds in turn residues have φ and ψ angles that lie outside the "permitted" regions of a typical Ramachandran plot (Figure 4.9). This is especially true of residues in the third position of type II turns where there is an abrupt change in the direction of the backbone. This residue is often glycine so the bond angles can adopt a wider range of values without causing steric clashes between the side-chain atoms and the backbone atoms. Ramachandran plots usually show only the permitted regions for all residues except glycine—this is why the rotation angles of type II turns appear to lie in a restricted area.



▲ **Figure 4.18**
Structure of PHL P2 from Timothy grass (*Phleum pratense*) pollen. (a) The two short, two-stranded, antiparallel β sheets are highlighted in blue and purple to show their orientation within the protein. **(b)** View of the β-sandwich structure in a different orientation showing hydrophobic residues (blue) and polar residues (red). A number of hydrophobic interactions connect the two β sheets. [PDB 1BMW].

4.7 Tertiary Structure of Proteins

Tertiary structure results from the folding of a polypeptide (which may already possess some regions of α helix and β structure) into a closely packed three-dimensional structure. An important feature of tertiary structure is that amino acid residues that are far apart in the primary structure are brought together permitting interactions among their side chains. Whereas secondary structure is stabilized by hydrogen bonding between amide hydrogens and carbonyl oxygens of the polypeptide backbone, tertiary



▲ **Figure 4.19**
Reverse turns. (a) Type I β turn. The structure is stabilized by a hydrogen bond between the carbonyl oxygen of the first N-terminal residue (Phe) and the amide hydrogen of the fourth residue (Gly). Note the proline residue at position $n+1$. **(b)** Type II β turn. This turn is also stabilized by a hydrogen bond between the carbonyl oxygen of the first N-terminal residue (Val) and the amide hydrogen of the fourth residue (Asn). Note the glycine residue at position $n+2$. [PDB 1AHL (giant sea anemone neurotoxin)].

Downloaded From Vet-e-Books.com

structure is stabilized primarily by noncovalent interactions (mostly the hydrophobic effect) between the side chains of amino acid residues. Disulfide bridges, though covalent, are also elements of tertiary structure they are not part of the primary structure since they form only after the protein folds.

A. Supersecondary Structures

Supersecondary structures, or **motifs**, are recognizable combinations of α helices, β strands, and loops that appear in a number of different proteins. Sometimes motifs are associated with a particular function although structurally similar motifs may have different functions in different proteins. Some common motifs are shown in Figure 4.20.

One of the simplest motifs is the helix–loop–helix (Figure 4.20a). This structure occurs in a number of calcium-binding proteins. Glutamate and aspartate residues in the loop of these proteins form part of the calcium-binding site. In certain DNA-binding proteins a version of this supersecondary structure is called a helix–turn–helix motif since the residues that connect the helices form a reverse turn. In these proteins, the residues of the α helices bind DNA.

The coiled-coil motif consists of two amphipathic α helices that interact through their hydrophobic edges (Figure 4.20b) as in the leucine zipper example (Figure 4.14). Several α helices can associate to form a helix bundle (Figure 4.20c). In this case, the individual α helices have opposite orientations, whereas they are parallel in the coiled-coil motif.

The $\beta\alpha\beta$ unit consists of two parallel β strands linked to an intervening α helix by two loops (Figure 4.20d). The helix connects the C-terminal end of one β strand to the N-terminal end of the next and often runs parallel to the two strands. A hairpin consists of two adjacent antiparallel β strands connected by a β turn (Figure 4.20e). (One example of a hairpin motif is shown in Figure 4.16.)

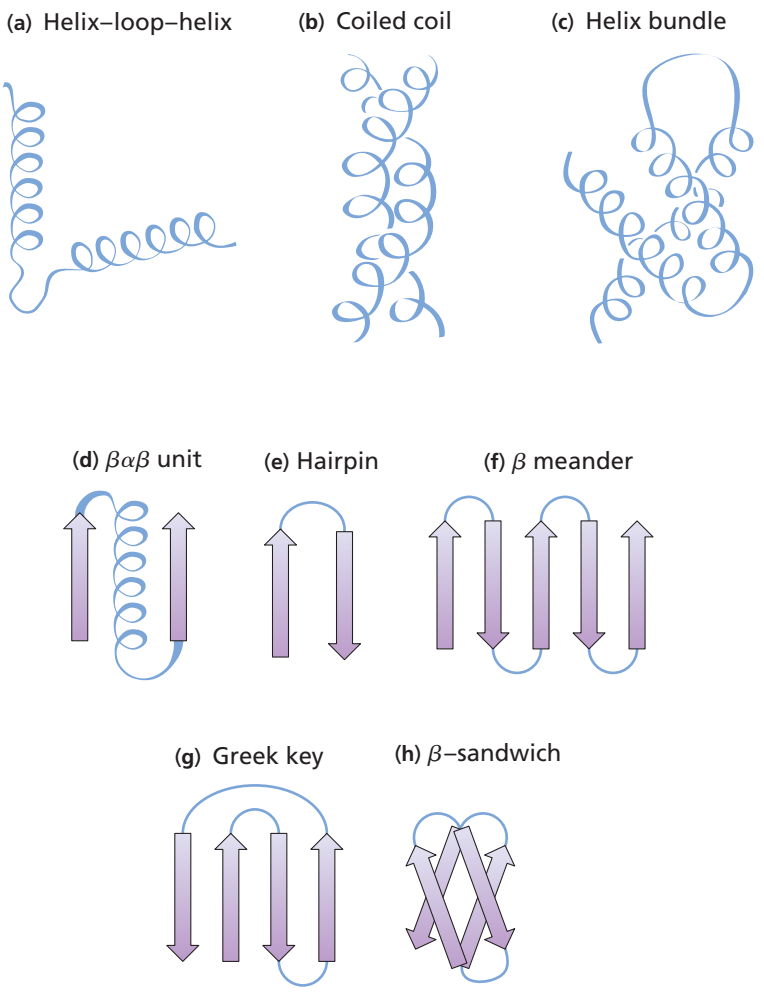


Figure 4.20 ▸
Common motifs. In folded proteins α helices and strands are commonly connected by loops and turns to form supersecondary structures, shown here as two-dimensional representations. Arrows indicate the N- to C-terminal direction of the peptide chain.

Downloaded From Vet-e-Books.com

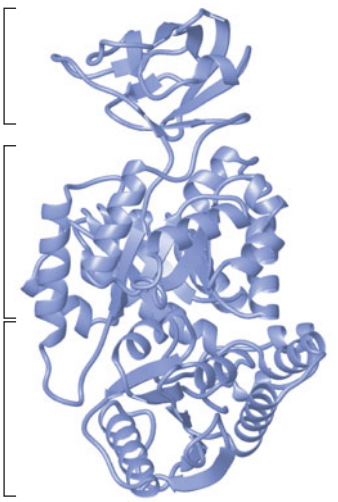
The β meander motif (Figure 4.20f) is an antiparallel β sheet composed of sequential β strands connected by loops or turns. The order of strands in the β sheet is the same as their order in the sequence of the polypeptide chain. The β meander sheet may contain one or more hairpins but, more typically, the strands are joined by larger loops. The Greek key motif takes its name from a design found on classical Greek pottery. This is a β sheet motif linking four antiparallel β strands such that strands 3 and 4 form the outer edges of the sheet and strands 1 and 2 are in the middle of the sheet. The β sandwich motif is formed when β strands or sheets stack on top of one another (Figure 4.20h). The figure shows an example of a β sandwich where the β strands are connected by short loops and turns, but β sandwiches can also be formed by the interaction of two β sheets in different regions of the polypeptide chain, as seen in Figure 4.18.

B. Domains

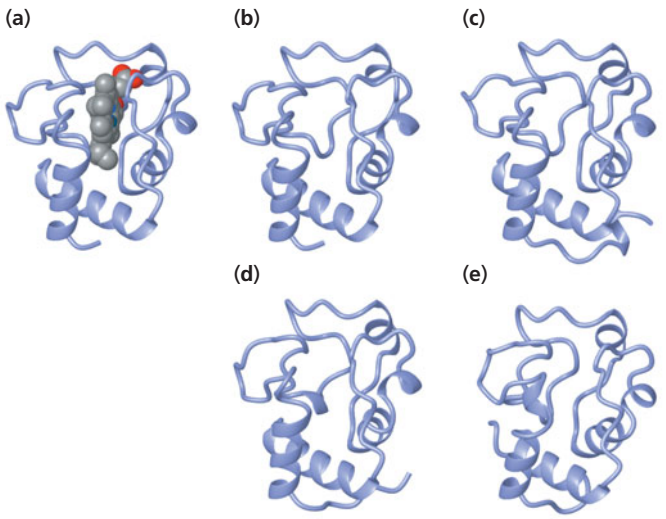
Many proteins are composed of several discrete, independently folded, compact units called **domains**. Domains may consist of combinations of motifs. The size of a domain varies from as few as 25 to 30 amino acid residues to more than 300. An example of a protein with multiple domains is shown in Figure 4.21. Note that each domain is a distinct compact unit consisting of various elements of secondary structure. Domains are usually connected by loops but they are also bound to each other through weak interactions formed by the amino acid side chains on the surface of each domain. The top domain of pyruvate kinase in Figure 4.21 contains residues 116 to 219, the central domain contains residues 1 to 115 plus 220 to 388, and the bottom domain contains residues 389 to 530. In general, domains consist of a contiguous stretch of amino acid residues as in the top and bottom domains of pyruvate kinase but in some cases a single domain may contain two or more different regions of the polypeptide chain as in the middle domain.

The evolutionary conservation of protein structure is one of the most important observations that has emerged from the study of proteins in the past few decades. This conservation is most easily seen in the case of single-domain homologous proteins from different species. For example, in Chapter 3 we examined the sequence similarity of cytochrome *c* and showed that the similarities in primary structure could be used to construct a phylogenetic tree that reveals the evolutionary relationships of the proteins from different species (Section 3.11). As you might expect, the tertiary structures of cytochrome *c* proteins are also highly conserved (Figure 4.22). Cytochrome *c* is an example of a protein that contains a heme prosthetic group. The conservation of protein structure is a reflection of its interaction with heme and its conserved function as an electron transport protein in diverse species.

Some domain structures occur in many different proteins whereas others are unique. In general, proteins can be grouped into families according to similarities in domain structures and amino acid sequence. All of the members of a family have descended from a common ancestral protein. Some biochemists believe that there may be only a few thousand families

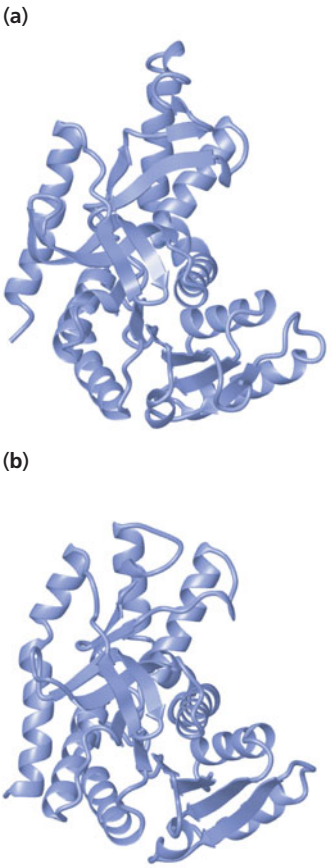


▲ **Figure 4.21**
Pyruvate kinase from cat (*Felis domesticus*). The main polypeptide chain of this common enzyme folds into three distinct domains as indicated by brackets. [PDB 1PKM].



◄ **Figure 4.22**
Conservation of cytochrome *c* structure. **(a)** Tuna (*Thunnus alalunga*) cytochrome *c* bound to heme [PDB 5CYT]. **(b)** Tuna cytochrome *c* polypeptide chain. **(c)** Rice (*Oryza sativa*) cytochrome *c* [PDB 1CCR]. **(d)** Yeast (*Saccharomyces cerevisiae*) cytochrome *c* [PDB 1YCC]. **(e)** Bacterial (*Rhodopila globiformis*) cytochrome *c* [PDB 1HRO].

Downloaded From Vet-e-Books.com



▲ **Figure 4.23**
Structural similarity of lactate and malate dehydrogenase. (a) *Bacillus stereothermophilus* lactate dehydrogenase [PDB 1LDN]. **(b)** *Escherichia coli* malate dehydrogenase [PDB 1EMD].

The enzymatic activities of lactate dehydrogenase and malate dehydrogenase are compared in Box 7.1.

suggesting that all modern proteins are descended from only a few thousand proteins that were present in the most primitive organisms living 3 billion years ago.

Lactate dehydrogenase and malate dehydrogenase are different enzymes that belong to the same family of proteins. Their structures are very similar as shown in Figure 4.23. The sequences of the proteins are only 23% identical. In spite of the obvious similarity in structure, Nevertheless, this level of sequence similarity is significant enough to conclude that the two proteins are homologous. They descend from a common ancestral gene that duplicated billions of years ago before the last common ancestor of all extant species of bacteria. Both lactate dehydrogenase and malate dehydrogenase are present in the same species which is why they are members of a family of related proteins. Protein families contain related proteins that are present in the same species. The cytochrome *c* proteins shown in Figure 4.22 are evolutionarily related but strictly speaking they are not members of a protein family because there is only one of them in each species. Protein familes arise from gene duplication events.

Protein domains can be classified by their structures. One commonly used classification scheme groups these domains into four categories. The "all-α" category contains domains that consist almost entirely of α helices and loops. "All-β" domains contain only β sheets and nonrepetitive structures that link β strands. The other two categories contain domains that have a mixture of α helices and β strands. Domains in the "α/β" class have supersecondary structures such as the $\beta\alpha\beta$ motif and others in which regions of α helix and β strand alternate in the polypeptide chain. In the "$\alpha + \beta$" category, the domains consist of local clusters of α helices and β sheet where each type of secondary structure arises from separate contiguous regions of the polypeptide chain.

Protein domains can be further classified by the presence of characteristic folds within each of the four main structural categories. A fold is a combination of secondary structures that form the core of a domain. Figure 4.24 on pages 103–104 shows selected examples of proteins from each of the main categories and illustrates a number of common domain folds. Some domains have easily recognizable folds, such as the β meander that contains antiparallel β strands connected by hairpin loops (Figure 4.20f), or helix bundles (Figure 4.19c). Other folds are more complex (Figure 4.25).

The important point about Figure 4.24 is not to memorize the structures of common proteins and folds. The key concept is that proteins can adopt an amazing variety of different sizes and shapes (tertiary structure) even though they contain only three basic forms of secondary structure.

C. Domain Structure, Function, and Evolution

The relationship between domain structure and function is complex. Often a single domain has a particular function such as binding small molecules or catalyzing a single reaction. In multifunctional enzymes, each catalytic activity can be associated with one of several domains found in a single polypeptide chain (Figure 4.24j). However, in many cases the binding of small molecules and the formation of the active site of an enzyme take place at the interface between two separate domains. These interfaces often form crevices, grooves, and pockets that are accessible on the surface of the protein. The extent of contact between domains varies from protein to protein.

The unique shapes of proteins, with their indentations, interdomain interfaces, and other crevices, allow them to fulfill dynamic functions by selectively and transiently binding other molecules. This property is best illustrated by the highly specific binding of reactants (substrates) to substrate-binding sites, or active sites, of enzymes. Because many binding sites are positioned toward the interior of a protein, they are relatively free of water. When substrates bind, they fit so well that some of the few remaining water molecules in the binding site are displaced.

D. Intrinsically Disordered Proteins

This section on tertiary structure wouldn't be complete without mentioning those proteins and domains that have no stable three-dimensional structure. These intrinsically disordered proteins (and domains) are quite common and the lack of secondary and tertiary structure is encoded in the amino acid sequences. There has been selection for

Downloaded From Vet-e-Books.com

clusters of charged residues (positive or negative) and proline residues that maintain the polypeptide chain in a disordered state.

Many of these proteins interact with other proteins. They contain short amino acid sequences that serve as binding sites and these binding sites are within the intrinsically disordered regions. This allows easy access to the binding site. If a protein contains two different binding sites for other proteins then the disordered polypeptide chain acts as a tether to bring the two binding proteins closer together. Several transcription factors also contain disordered regions when they are not bound to DNA. These regions become ordered when the proteins interact with DNA.

KEY CONCEPT

There are only three basic types of secondary structure but thousands of tertiary folds and domains.

Speculations on the possible relationship between protein domains and gene organization will be presented in Chapter 21.

4.8 Quaternary Structure

Many proteins exhibit an additional level of organization called *quaternary structure*. Quaternary structure refers to the organization and arrangement of subunits in a protein with multiple subunits. Each subunit is a separate polypeptide chain. A multisubunit protein is referred to as an oligomer (proteins with only one polypeptide chain are monomers). The subunits of a multisubunit protein may be identical or different. When the subunits are identical, dimers and tetramers predominate. When the subunits differ, each type often has a different function. A common shorthand method for describing oligomeric proteins uses Greek letters to identify types of subunits and subscript numerals to indicate numbers of subunits. For example, an $\alpha_2\beta\gamma$ protein contains two subunits designated α and one each of subunits designated β and γ.

The subunits within an oligomeric protein always have a defined stoichiometry and the arrangement of the subunits gives rise to a stable structure where subunits are usually held together by weak noncovalent interactions. Hydrophobic interactions are the principal forces involved although electrostatic forces may contribute to the proper alignment of the subunits. Because intersubunit forces are usually rather weak, the subunits of an oligomeric protein can often be separated in the laboratory. *In vivo,* however, the subunits usually remain tightly associated.

Examples of several multisubunit proteins are shown in Figure 4.26. In the case of triose phosphate isomerase (Figure 4.26a) and HIV protease (Figure 4.26b), the identical subunits associate through weak interactions between the side chains found mainly in loop regions. Similar interactions are responsible for the formation of the MS2 capsid protein that consists of a trimer of identical subunits (Figure 4.26d). In this case, the trimer units assemble into a more complex structure—the bacteriophage particle. The enzyme HGPRT (Figure 4.26e) is a tetramer formed from the association of two pairs of nonidentical subunits. Each of the subunits is a recognizable domain.

The potassium channel protein (Figure 4.26c) is an example of a tetramer of identical subunits where the subunits interact to form a membrane-spanning region consisting of an eight-helix bundle. The subunits do not form separate domains within the protein but instead come together to form a single channel. The bacterial photosystem shown in Figure 4.26f is a complex example of quaternary structure. Three of the subunits contribute to a large membrane-bound helix bundle while a fourth subunit (a cytochrome) sits on the exterior surface of the membrane.

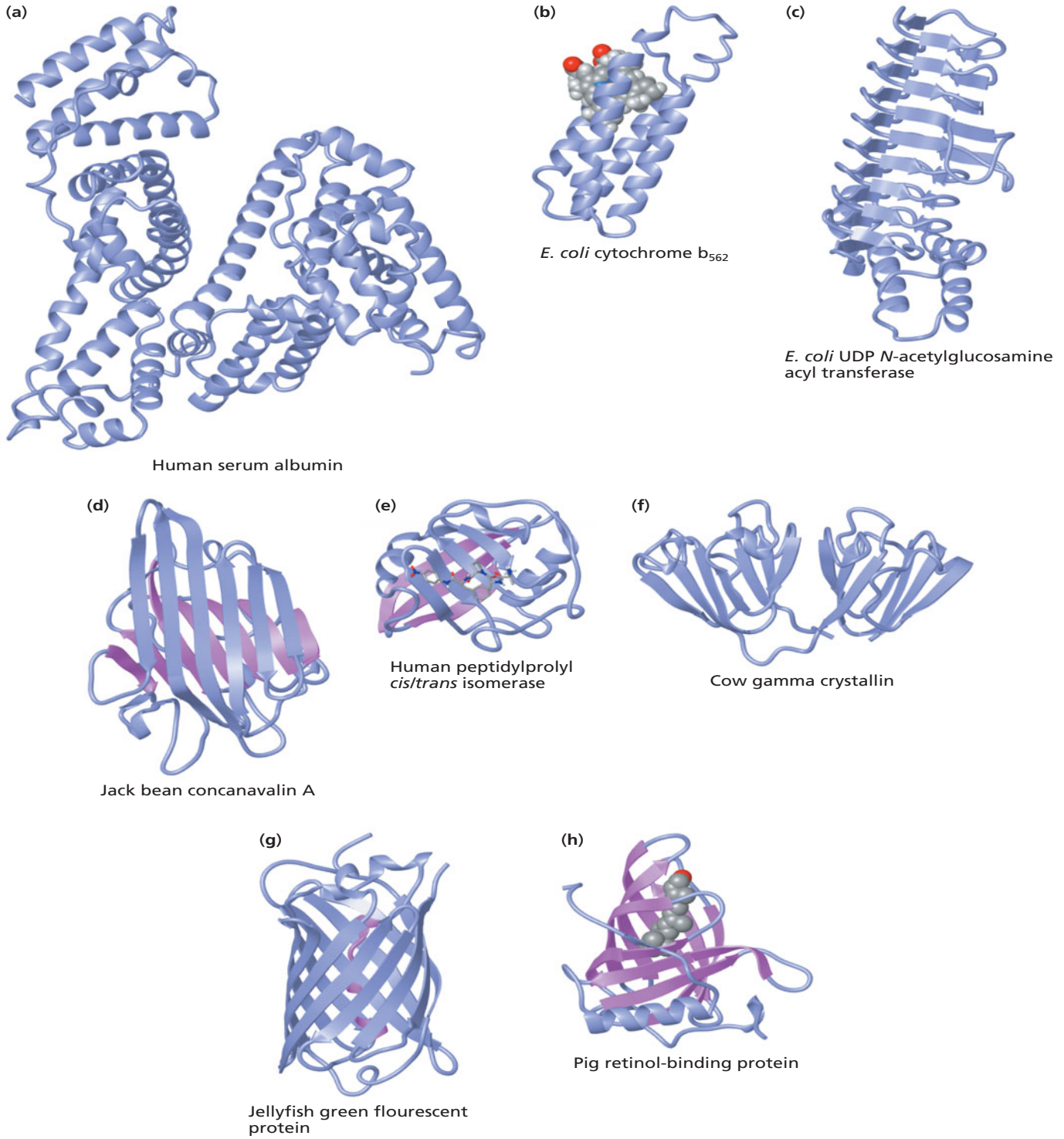
The structures and functions of bacterial and plant photosystems are described in Chapter 15.

Determination of the subunit composition of an oligomeric protein is an essential step in the physical description of a protein. Typically, the molecular weight of the native oligomer is estimated by gel-filtration chromatography and then the molecular weight of each chain is determined by SDS–polyacrylamide gel electrophoresis (Section 3.6). For a protein having only one type of chain, the ratio of the two values provides the number of chains per oligomer.

The fact that a large proportion of proteins consist of multiple subunits is probably related to several factors:

1. Oligomers are usually more stable than their dissociated subunits suggesting that quaternary structure prolongs the life of a protein *in vivo.*
2. The active sites of some oligomeric enzymes are formed by residues from adjacent polypeptide chains.

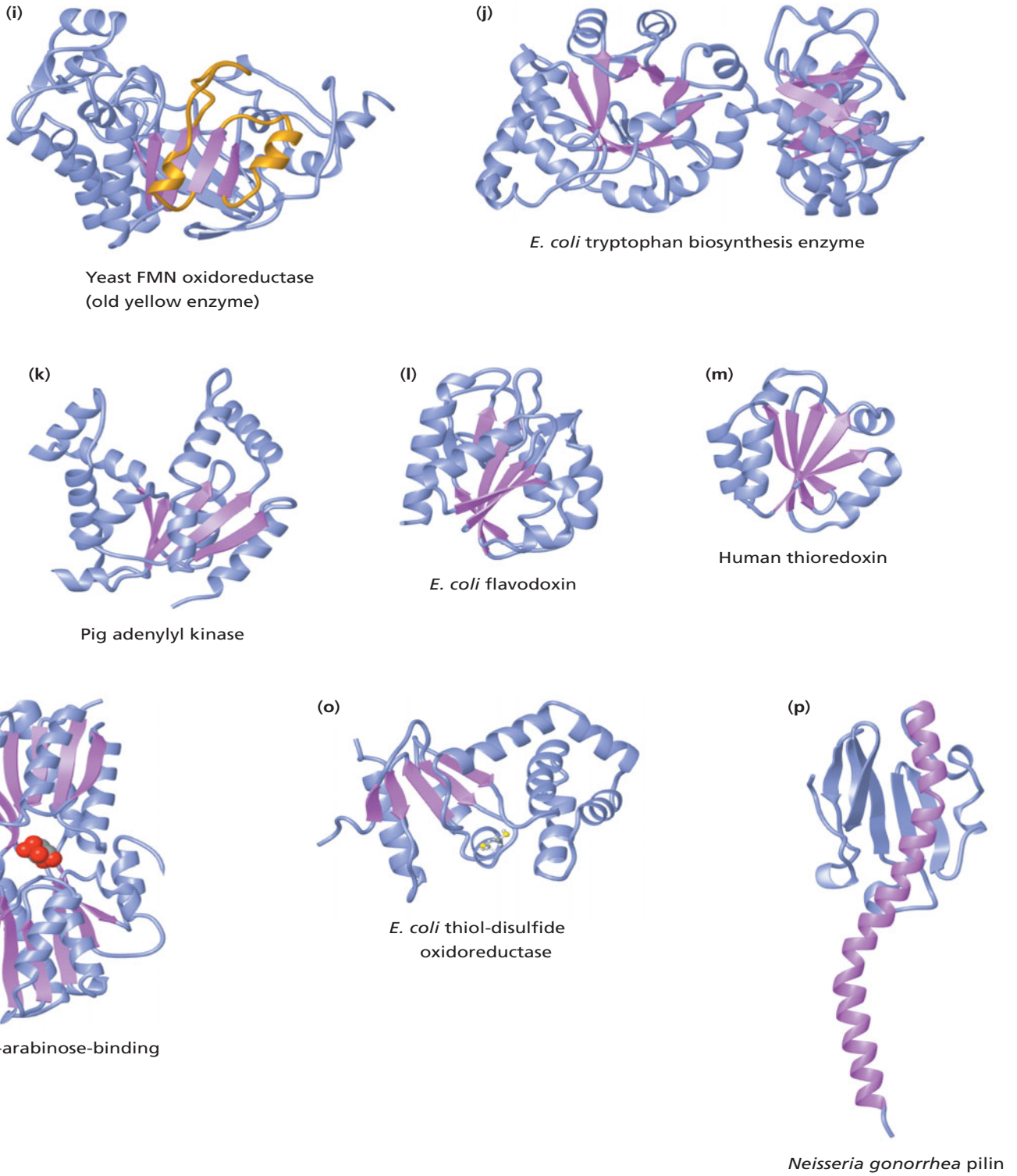
Downloaded From Vet-e-Books.com



▲ Figure 4.24

Examples of tertiary structure in selected proteins. (a) Human (*Homo sapiens*) serum albumin [PDB 1BJ5] (class: all-α). This protein has several domains consisting of layered α helices and helix bundles. **(b)** *Escherichia coli* cytochrome $\mathbf{b}_{562}$ [PDB 1QPU] (class: all-α). This is a heme-binding protein consisting of a single four-helix bundle domain. **(c)** *Escherichia coli* UDP N-acetylglucosamine acyl transferase [PDB 1LXA] (class: all-β). The structure of this enzyme shows a classic example of a β helix domain. **(d)** Jack bean (*Canavalia ensiformis*) concanavalin A [PDB 1CON] (class: all-β). This carbohydrate-binding protein (lectin) is a single-domain protein made up of a large β sandwich fold. **(e)** Human (*Homo sapiens*) peptidylprolyl *cis/trans* isomerase [PDB 1VBS] (class: all-β). The dominant feature of the structure is a β sandwich fold. **(f)** Cow (*Bos taurus*) γ-crystallin [PDB 1A45] (class: all-β) This protein contains two β barrel domains. **(g)** Jellyfish (*Aequorea victoria*) green fluorescent protein [PDB 1GFL] (class: all-β). This is a β barrel structure with a central α helix. The strands of the sheet are antiparallel. **(h)** Pig (*Sus scrofa*) retinol-binding protein [PDB 1AQB] (class: all-β). Retinol binds in the interior of a β barrel fold. **(I)** Brewer's yeast (*Saccharomyces carlsburgensis*) old yellow enzyme (FMN oxidoreductase) [PDB 1OYA] (class: α/β). The central fold is an α/β barrel with parallel β strands connected by α helices. Two of the connecting α helical regions are highlighted in yellow. **(j)** *Escherichia coli* enzyme required for tryptophan biosynthesis [PDB 1PII] (class: α/β). This is a bifunctional enzyme containing two distinct domains. Each domain is an example of an α/β barrel. The left-hand domain contains the indolglycerol phosphate

Downloaded From Vet-e-Books.com



▲ **Figure 4.24** (***continued***)
synthetase activity, and the right-hand domain contains the phosphoribosylanthranilate isomerase activity. **(k)** Pig (*Sus scrofa*) adenylyl kinase [PDB 3ADK] (class: α/β). This single-domain protein consists of a five-stranded parallel β sheet with layers of α helices above and below the sheet. The substrate binds in the prominent groove between α helices. **(l)** *Escherichia coli* flavodoxin [PDB 1AHN] (class: α/β). The fold is a five-stranded parallel twisted sheet surrounded by α helices. **(m)** Human (*Homo sapiens*) thioredoxin [PDB 1ERU] (class: α/β). The structure of this protein is very similar to that of *E. coli* flavodoxin except that the five-stranded twisted sheet in the thioredoxin fold contains a single antiparallel strand. **(n)** *Escherichia coli* L-arabinose-binding protein [PDB 1ABE] (class: α/β). This is a two-domain protein where each domain is similar to that in *E. coli* flavodoxin. The sugar L-arabinose binds in the cavity between the two domains. **(o)** *Escherichia coli* DsbA (thiol-disulfide oxidoreductase/disulfide isomerase) [PDB 1A23] (class: α/β). The predominant feature of this structure is a (mostly) antiparallel β sheet sandwiched between α helices. Cysteine side chains at the end of one of the α helices are shown (sulfur atoms are yellow). **(p)** *Neisseria gonorrhea* pilin [PDB 2PIL] (class: $\alpha + \beta$). This polypeptide is one of the subunits of the pili on the surface of the bacteria responsible for gonorrhea. There are two distinct regions of the structure: a β sheet and a long α helix.

Downloaded From Vet-e-Books.com

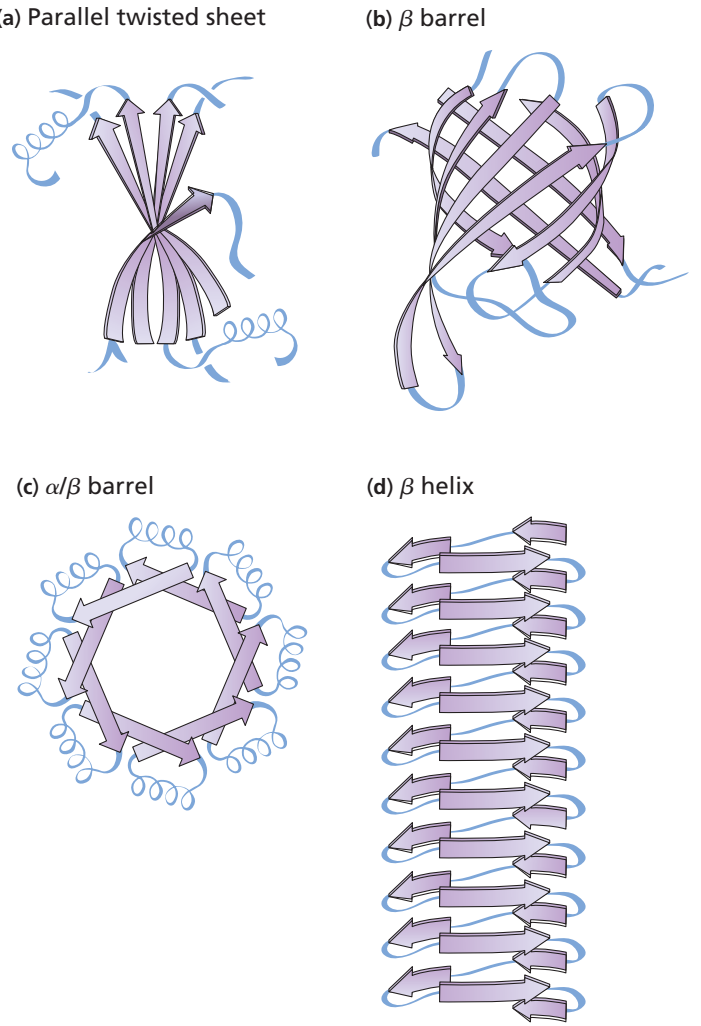


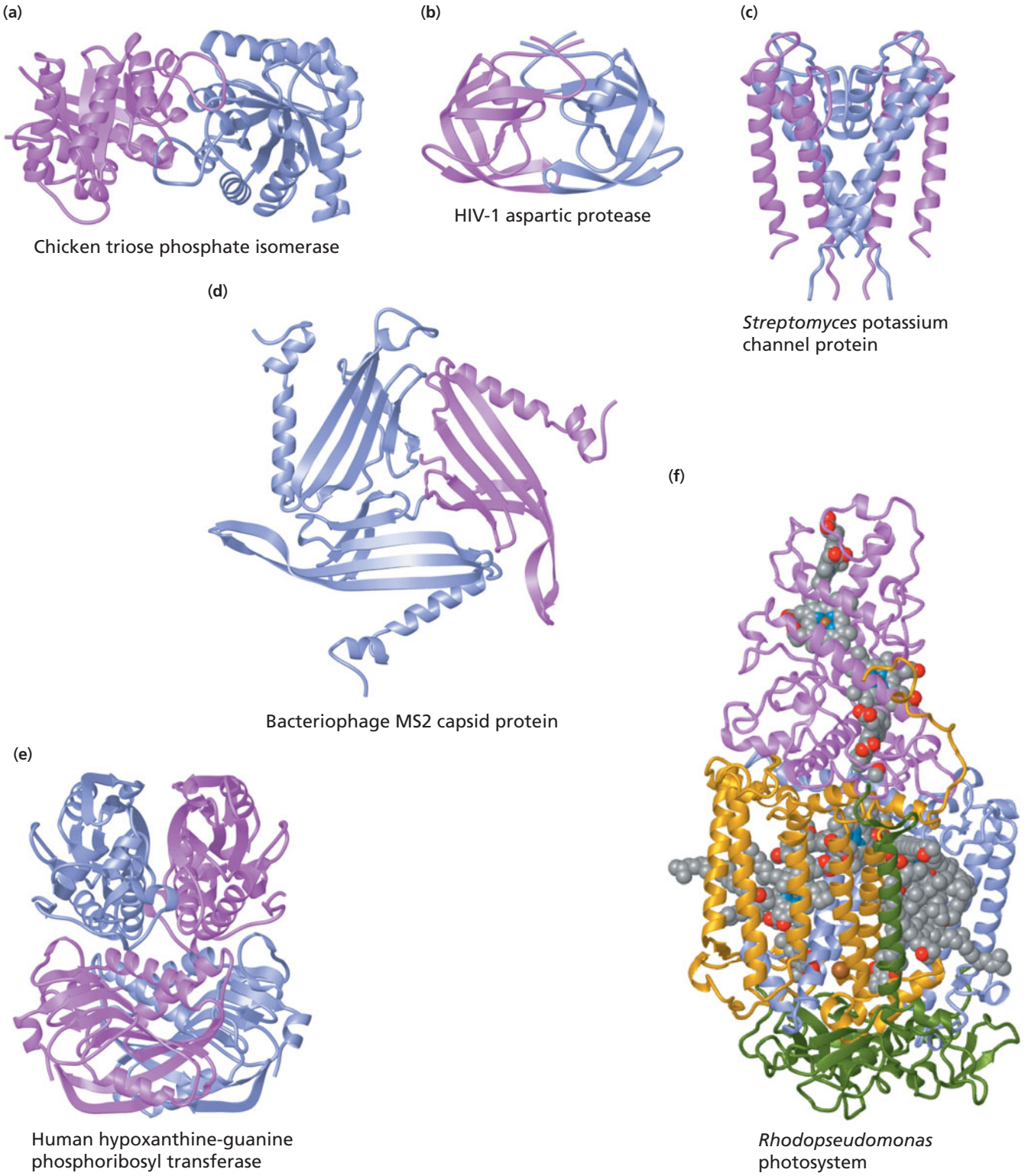
Figure 4.25 ►
Common domain folds.

3. The three-dimensional structures of many oligomeric proteins change when the proteins bind ligands. Both the tertiary structures of the subunits and the quaternary structures (i.e., the contacts between subunits) may be altered. Such changes are key elements in the regulation of the biological activity of certain oligomeric proteins.
4. Different proteins can share the same subunits. Since many subunits have a defined function (e.g., ligand binding), evolution has favored selection for different combinations of subunits to carry out related functions. This is more efficient than selection for an entirely new monomeric protein that duplicates part of the function.
5. A multisubunit protein may bring together two sequential enzymatic steps where the product of the first reaction becomes the substrate of the second reaction. This gives rise to an effect known as channeling (Section 5.11).

As shown in Figure 4.26, the variety of multisubunit proteins ranges from simple homodimers such as triose phosphate isomerase to large complexes such as the photosystems in bacteria and plants. We would like to know how many proteins are monomers and how many are oligomers but studies of cell proteomes—the complete complement of proteins —have only begun.

Table 4.1 on page 108 shows the results of a survey of *E. coli* proteins in the SWISS-PROT database. Of those polypeptides that have been analyzed, only about 19% are in monomers. Dimers are the largest class among the oligomers, and homodimers—where the two subunits are identical—represent 31% of all proteins. The next largest class is tetramers of identical subunits. Note that trimers are relatively rare. Most proteins exhibit dyad symmetry meaning that you can usually draw a line through a protein dividing it into two halves that are symmetrical about this axis. This dyad symmetry is seen even in

Downloaded From Vet-e-Books.com



▲ **Figure 4.26**

Quaternary structure. (a) Chicken (*Gallus gallus*) triose phosphate isomerase [PDB 1TIM]. This protein has two identical subunits with α/β barrel folds. **(b)** HIV-1 aspartic protease [PDB 1DIF]. This protein has two identical all-β subunits that bind symmetrically. HIV protease is the target of many new drugs designed to treat AIDS patients. **(c)** *Streptomyces lividans* potassium channel protein [PDB 1BL8]. This membrane-bound protein has four identical subunits, each of which contributes to a membrane-spanning eight-helix bundle. **(d)** Bacteriophage MS2 capsid protein [PDB 2MS2]. The basic unit of the MS2 capsid is a trimer of identical subunits with a large β sheet. **(e)** Human (*Homo sapiens*) hypoxanthine-guanine phosphoribosyl transferase (HGPRT) [PDB 1BZY]. HGPRT is a tetrameric protein containing two different types of subunit. **(f)** *Rhodopseudomonas viridis* photosystem [PDB 1PRC]. This complex, membrane-bound protein has two identical subunits (orange, blue) and two other subunits (purple, green) bound to several molecules of photosynthetic pigments.

Downloaded From Vet-eBooks.com

Table 4.1 Natural occurrence of oligomeric proteins in *Escherichia coli*

Oligomeric state	Number of homooligomers	Number of heterooligomers	Percent
Monomer	72		19.4
Dimer	115	27	38.2
Trimer	15	5	5.4
Tetramer	62	16	21.0
Pentamer	1	1	0.1
Hexamer	20	1	5.6
Heptamer	1	1	0.1
Octamer	3	6	2.4
Nonamer	0	0	0.0
Decamer	1	0	0.0
Undecamer	0	1	0.0
Dodecamer	4	2	1.6
Higher oligomers	8		2.2
Polymers	10		2.7

heterooligomers such as hypoxanthine-guanine phosphoribosyl transferase (HGPRT, Figure 4.26e) and hemoglobin (Section 4.14). Of course, there are many exceptions, especially when the oligomers are large complexes.

We will encounter many other examples of multisubunit proteins throughout this textbook, especially in the chapters on information flow (Chapters 20–22). DNA polymerase, RNA polymerase, and the ribosome are excellent examples. Other examples include GroEL (Section 4.11D) and pyruvate dehydrogenase (Section 13.1). Many of these large proteins are easily seen in electron micrographs, as illustrated in Figure 4.27.

Large complexes are referred to, metaphorically, as *protein machines* since the various polypeptide components work together to carry out a complex reaction. The term

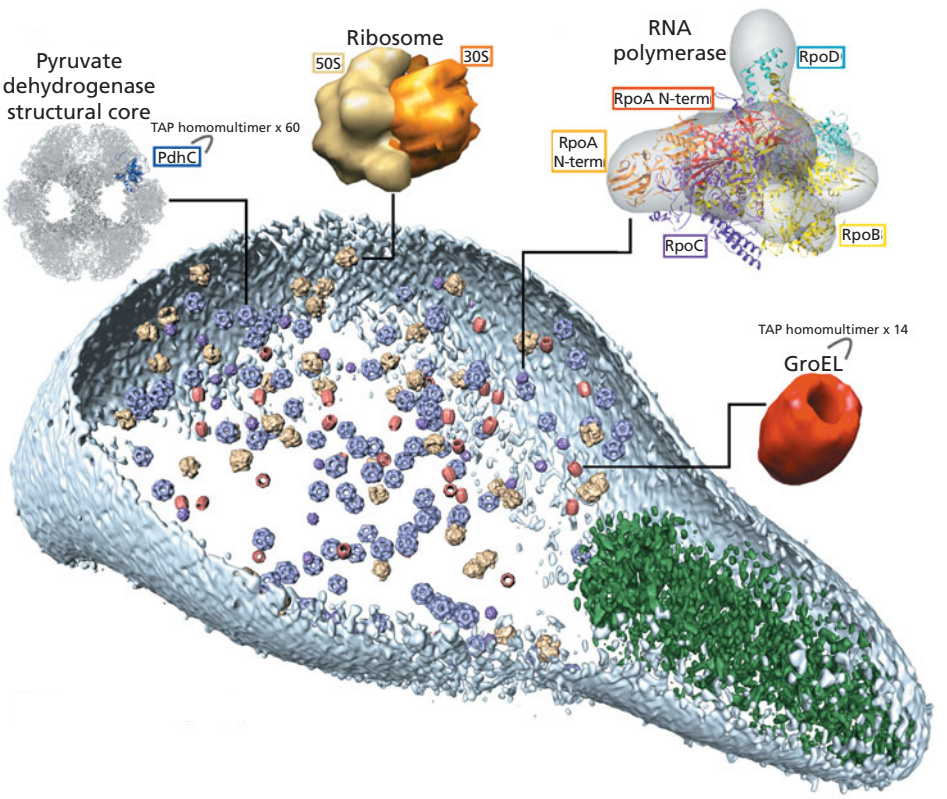
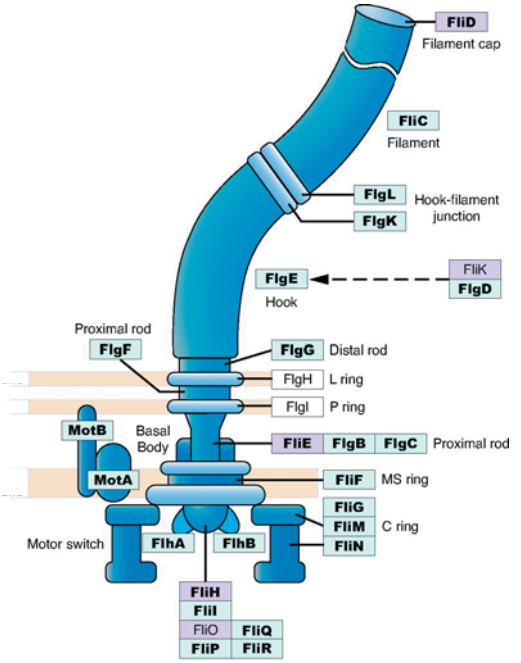


Figure 4.27 ►
Large protein complexes in the bacterium *Mycoplasma pneumoniae*. *M. pneumoniae* causes some forms of pneumonia in humans. This species has one of the smallest genomes known (689 protein-encoding genes). Most of those genes are likely to represent the minimum proteome of a living cell. The cell contains several large complexes found in all cells: pyruvate dehygrogenase (purple), ribosome (yellow), GroEL (red), and RNA polymerase (orange). It also contains a rod (green) found only in some bacteria. [Adapted from Kühner et al. (2009). Proteome organization in a genome-reduced bacterium. *Science* 326:1235–1240]

Downloaded From Vet-e-Books.com

was originally coined to describe complexes such as the replisome (Figure 20.15) but there are many other examples, including those shown in Figure 4.27.

The bacterial flagellum (Figure 4.28) is a spectacular example of a protein machine. The complex drives the rotation of a long flagellum using protonmotive force as an energy source (Section 14.3). More than 50 genes are required to build the flagellum in *E. coli* but surveys of other bacteria reveal that there are only about 21 core proteins required to build a functional flagellum. The evolutionary history of this protein machine is being actively investigated and it appears that it was built up by combining simpler components involved in ATP synthesis and membrane secretion.



▲ **Figure 4.28**
Bacterial flagellum. The bacterial flagellum is a protein machine composed of 21 core subunits found in all species (blue boxes). Two additional subunits are missing in Firmicutes (white boxes) and five others are sporadically distributed. The flagellum (hook + filament + cap) spins as the motor complex rotates. The three layers represent the outer membrane (top), the peptidoglycan layer (middle), and the cytoplasmic membrane (bottom). (Courtesy of Howard Ochman.)

4.9 Protein–Protein Interactions

The various subunits in multisubunit proteins bind to each other so strongly that they rarely dissociate inside the cell. These protein–protein contacts are characterized by a number of weak interactions. We have already become familiar with the type of interactions involved: hydrogen bonds, charge–charge interactions, van der Waals forces, and hydrophobic interactions (Section 2.5). In some cases the contact areas between two subunits are localized to small patches on the surface of the polypeptides but while in other cases there can be extensive contact spread over large portions of the polypeptides. The distinguishing feature of subunit contacts is the cumulative effect of a large number of individual weak interactions giving a binding strength that is sufficient to keep the subunits together.

In addition to subunit–subunit contacts, there are many other types of protein–protein interactions that are less stable. These range from transient contacts between external proteins and receptors on the cell surface to weak interactions between various enzymes in metabolic pathways. These weak interactions are much more difficult to detect but they are essential components of many biochemical reactions.

Consider a simple interaction between two proteins, P1 and P2, to give a complex P1:P2. The equilibrium between the free and bound molecules can be described by either an association constant (K_a) or a dissociation constant (K_d) ($K_a = 1/K_d$).

$$\text{P1} + \text{P2} \rightleftharpoons \text{P1:P2} \qquad K_a = \frac{[\text{P1:P2}]}{[\text{P1}][\text{P2}]} \tag{4.1}$$

$$\text{P1:P2} \rightleftharpoons \text{P1} + \text{P2} \qquad K_d = \frac{[\text{P1}][\text{P2}]}{[\text{P1:P2}]} \tag{4.2}$$

Typical association constants for the binding of subunits in a multimeric protein are greater than $10^8\ M^{-1}$ ($K_a > 10^8\ M^{-1}$) and can range as high as $10^{14}\ M^{-1}$ for very tight interactions. At the other extreme are protein–protein interactions that are so weak they have no biological significance. These can be fortuitous interactions that arise from time to time because any two polypeptides will almost always form some kind of weak contact. The lower limit of relevant association constants is about $10^4\ M^{-1}$ ($K_a < 10^4\ M^{-1}$). The really interesting cases are those with association constants between these two values.

The binding of transcription factors to RNA polymerase is one example of weak protein–protein interactions that are very important. The association constants range from about $10^5\ M^{-1}$ to $10^7\ M^{-1}$. The interactions between proteins in signaling pathways also fall into this range as do the interactions between enzymes in metabolic pathways.

Let's look at what these association constants mean in terms of protein concentrations. As the concentrations of P1 and P2 increase it becomes more and more likely that they will interact and bind to each other. At some concentration, the rate of binding (a second-order reaction) becomes comparable to the rate of dissociation (a first-order reaction) and complexes will be present in appreciable amounts. Using the association constant, we can calculate the ratio of free polypeptide (P1 or P2) as a fraction of the total concentration of either one ($P1_T$ or $P2_T$). This ratio [free]/[total] tells us how much of the complex will be present at a given protein concentration.

Downloaded From Vet-e-Books.com

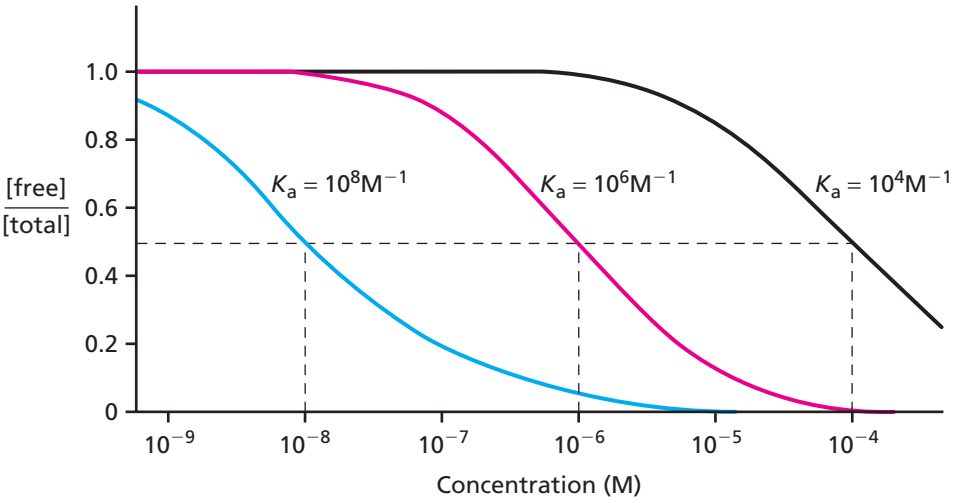


Figure 4.29 ►
Association constants and protein concentration. The ratio of free unbound protein to total protein is shown for a protein–protein interaction at three different association constants. Assuming that the concentration of the other component is in excess, the concentrations at which half the molecules are in complex and half are free corresponds to the reciprocal of the association constant. [Adapted from van Holde, Johnson, and Ho, *Principles of Physical Biochemistry*, Prentice Hall.]

The curves in Figure 4.29 show these ratios for three different association constants corresponding to very weak ($K_a = 10^4\ M^{-1}$), moderate ($K_a = 10^6\ M^{-1}$), and very strong ($K_a = 10^8\ M^{-1}$) protein–protein interactions. If we assume that one of the components is present in excess, then the curves represent the concentrations of only the rate-limiting polypeptide. One can demonstrate mathematically that for simple systems the point at which half of the polypeptide is free and half is in a complex corresponds to the reciprocal of the association constant. For example, if $K_a = 10^8\ M^{-1}$ then most of the polypeptide will be bound at any concentration over 10^{-8} M.

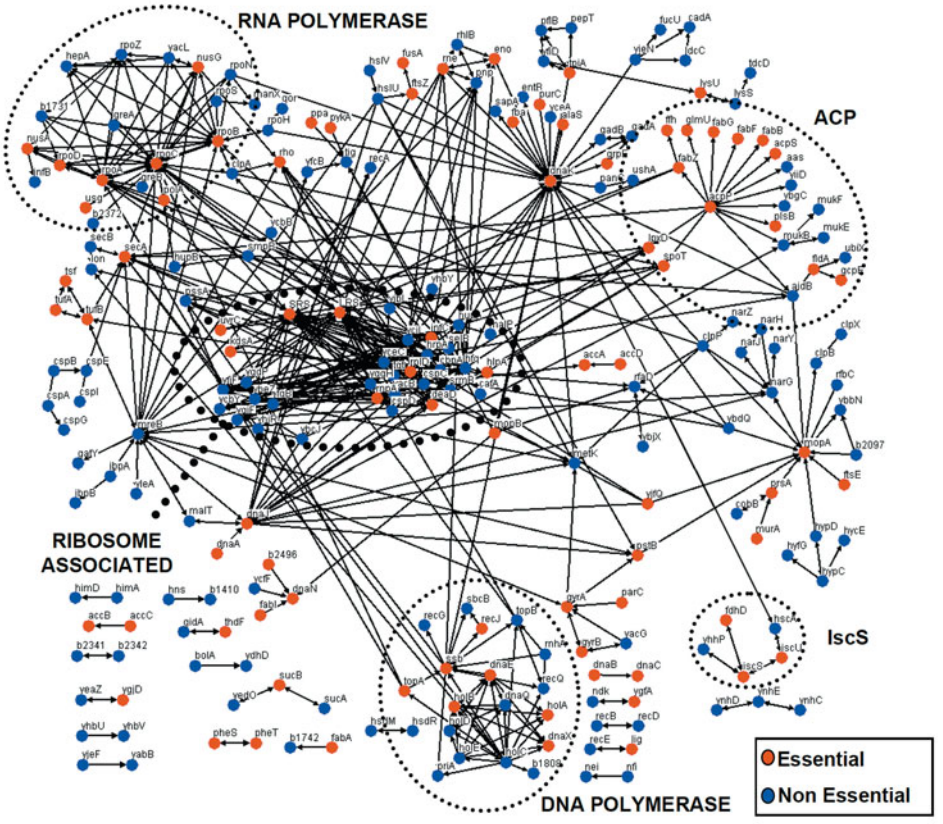
What does this mean in terms of molecules per cell? For an *E. coli* cell whose volume is about 2×10^{-15} l it means that as long as there are more than a dozen molecules per cell the complex will be stable if $K_a > 10^8\ M^{-1}$. This is why large oligomeric complexes can exist in *E. coli* even if there are only a few dozen per cell. Most eukaryotic cells are 1000 times larger and there must be 12,000 molecules in order to achieve a concentration of 10^{-8} M. Figure 4.29 also shows why it is impossible for weak interactions to produce significant numbers of P1:P2 complexes. The protein concentration has to be greater than 10^{-4} M in order for the complex to be present in significant quantity and this concentration corresponds to 120,000 molecules in an *E. coli* cell or 120 million molecules in a eukaryotic cell. There are no free polypeptides present at such concentrations so weak interactions of this magnitude are biologically meaningless.

There are many techniques for detecting moderate binding. These include direct techniques such as affinity chromatopraphy, immunoprecipitation, and chemical cross-linking. Newer techniques rely on more sophisticated manipulations such as phage display, two-hybrid analysis, and genetic methods. Many workers are attempting to map the interactions of every protein in the cell using these techniques. An example of such an "interactome" for many *E. coli* proteins is shown in Figure 4.30. Note that strong interactions between the subunits of oligomers are easily detected as shown by lines connecting the subunits of RNA polymerase, the ribosome, and DNA polymerase. Other lines connect RNA polymerase to various transcription factors—these represent moderate interactions. Further studies of the "interactome" in various species should give us a much better picture of the complex protein–protein interactions in living cells.

4.10 Protein Denaturation and Renaturation

Environmental changes or chemical treatments may disrupt the native conformation of a protein causing loss of biological activity. Such a disruption is called **denaturation**. The amount of energy needed to cause denaturation is often small, perhaps equivalent to that needed for the disruption of three or four hydrogen bonds. Some proteins may unfold completely when denatured to form a random coil (a fluctuating chain considered to be totally disordered) but most denatured proteins retain considerable internal structure. It is sometimes possible to find conditions under which small denatured proteins can spontaneously renature, or refold, following denaturation.

Downloaded From Vet-e-Books.com



◄ **Figure 4.30**
***E. coli* interactome.** Each point on the diagram represents a single *E. coli* protein. Red dots are essential proteins and blue dots are nonessential proteins. Lines joining the points indicate experimentally determined protein–protein interactions. Five large complexes are shown: RNA polymerase, DNA polymerase, ribosome and associated proteins, proteins interacting with cysteine desulfurase (IscS), and proteins associated with acyl carrier protein (ACP). (The role of ACP is described in Section 16.1.) [Adapted from Butland et al. (2005)]

Proteins are commonly denatured by heating. Under the appropriate conditions, a modest increase in temperature will result in unfolding and loss of secondary and tertiary structure. An example of thermal denaturation is shown in Figure 4.31. In this experiment, a solution containing bovine ribonuclease A is heated slowly and the structure of the protein is monitored by various techniques that measure changes in conformation. All of these techniques detect a change when denaturation occurs. In the case of bovine ribonuclease A, thermal denaturation also requires a reducing agent that disrupts internal disulfide bridges allowing the protein to unfold.

Denaturation takes place over a relatively small range of temperature. This indicates that unfolding is a cooperative process where the destabilization of just a few weak interactions leads to almost complete loss of native conformation. Most proteins have a characteristic "melting" temperature (T_m) that corresponds to the temperature at the midpoint of the transition between the native and denatured forms. The T_m depends on pH and the ionic strength of the solution.

Most proteins are stable at temperatures up to 50°C to 60°C under physiological conditions. Some species of bacteria, such as those that inhabit hot springs and the vicinity of deep ocean thermal vents, thrive at temperatures well above this range. Proteins in these species denature at much higher temperatures as expected. Biochemists are actively studying these proteins in order to determine how they resist denaturation.

Proteins can also be denatured by two types of chemicals—chaotropic agents and detergents (Section 2.4). High concentrations of chaotropic agents, such as urea and guanidinium salts (Figure 4.32), denature proteins by allowing water molecules to solvate nonpolar groups in the interior of proteins. The water molecules disrupt the hydrophobic interactions that normally stabilize the native conformation. The hydrophobic tails of

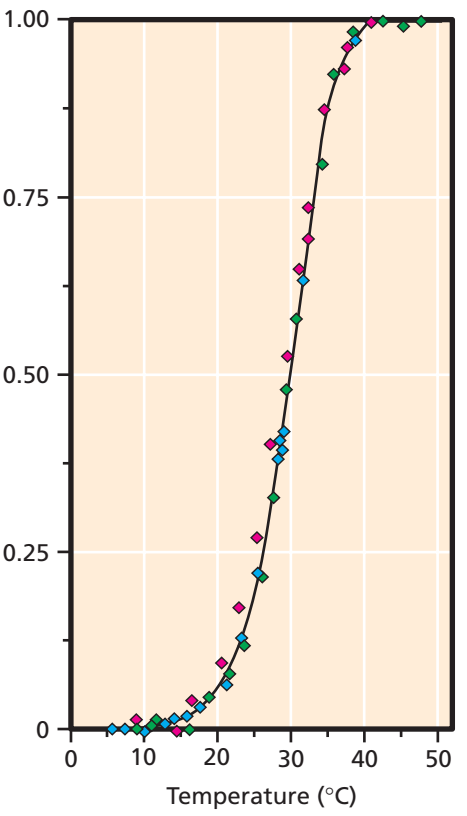
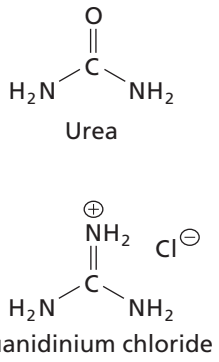
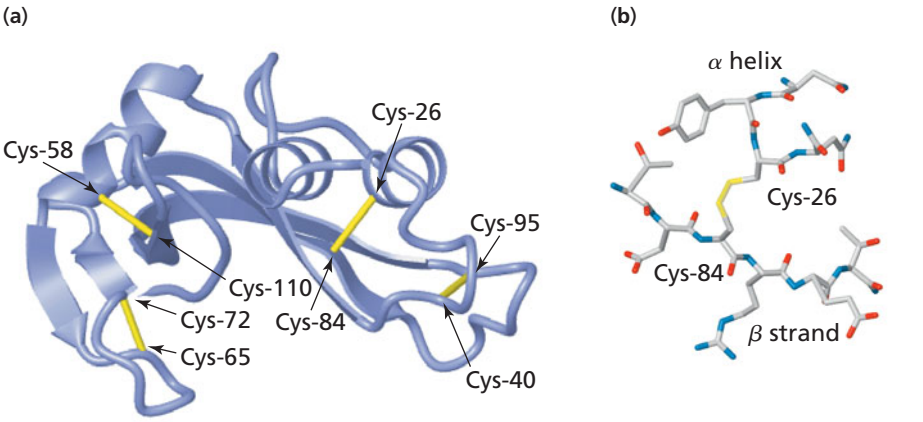


Figure 4.31 ►
Heat denaturation of ribonuclease A. A solution of ribonuclease A in 0.02 M KCl at pH 2.1 was heated. Unfolding was monitored by changes in ultraviolet absorbance (blue), viscosity (red), and optical rotation (green). The y-axis is the fraction of the molecule unfolded at each temperature. [Adapted from Ginsburg, A., and Carroll, W. R. (1965). Some specific ion effects on the conformation and thermal stability of ribonuclease. *Biochemistry* 4:2159–2174.

Downloaded From Vet-e-Books.com



▲ **Figure 4.32**
Urea and guanidinium chloride.



▲ **Figure 4.33**
Disulfide bridges in bovine ribonuclease A. (a) Location of disulfide bridges in the native protein. **(b)** View of the disulfide bridge between Cys-26 and Cys-84 [PDB 2AAS].

detergents, such as sodium dodecyl sulfate (Figure 2.8), also denature proteins by penetrating the protein interior and disrupting hydrophobic interactions.

The native conformation of some proteins (e.g., ribonuclease A) is stabilized by disulfide bonds. Disulfide bonds are not generally found in intracellular proteins but are sometimes found in proteins that are secreted from cells. The presence of disulfide bonds stabilizes proteins by making them less susceptible to unfolding and subsequent degradation when they are exposed to the external environment. Disulfide bond formation does not drive protein folding; instead, the bonds form where two cysteine residues are appropriately located once the protein has folded. Formation of a disulfide bond requires oxidation of the thiol groups of the cysteine residues (Figure 3.4), probably by disulfide-exchange reactions involving oxidized glutathione, a cysteine-containing tripeptide.

Figure 4.33a shows the locations of the disulfide bridges in ribonuclease A. (Compare this orientation of the protein with that shown in Figure 4.3.) There are four disulfide bridges. They can link adjacent β strands, β strands to α helices, or β strands to loops. Figure 4.33b is a view of the disulfide bridge between a cysteine residue in an α helix (Cys-26) and a cysteine residue in a β strand (Cys-84). Note that the S—S bond does not align with the cysteine side chains. Disulfide bridges will form whenever the two cysteine sulfhydryl groups are in close proximity in the native conformation.

The numbering convention for amino acid residues in a polypeptide starts at the N-terminal end (Section 3.5). Cys-26 is the 26th residue from the N-terminus.

Complete denaturation of proteins containing disulfide bonds requires cleavage of these bonds in addition to disruption of hydrophobic interactions and hydrogen bonds. 2-Mercaptoethanol or other thiol reagents can be added to a denaturing medium in order to reduce any disulfide bonds to sulfhydryl groups (Figure 4.34). Reduction of the disulfide bonds of a protein is accompanied by oxidation of the thiol reagent.

In a series of classic experiments, Christian B. Anfinsen and his coworkers studied the renaturation pathway of ribonuclease A that had been denatured in the presence of thiol reducing agents. Since ribonuclease A is a relatively small protein (124 amino acid

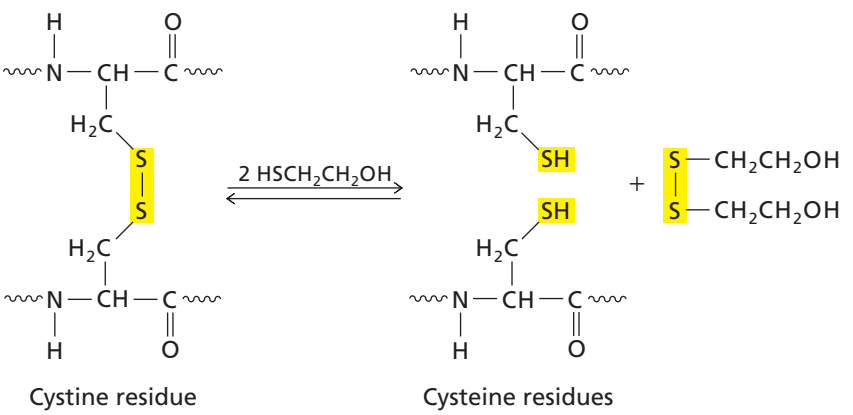


Figure 4.34 ►
Cleaving disulfide bonds. When a protein is treated with excess 2-mercaptoethanol ($HSCH_2CH_2OH$), a disulfide-exchange reaction occurs in which each cystine residue is reduced to two cysteine residues and 2-mercaptoethanol is oxidized to a disulfide.

Downloaded From Vet-e-Books.com

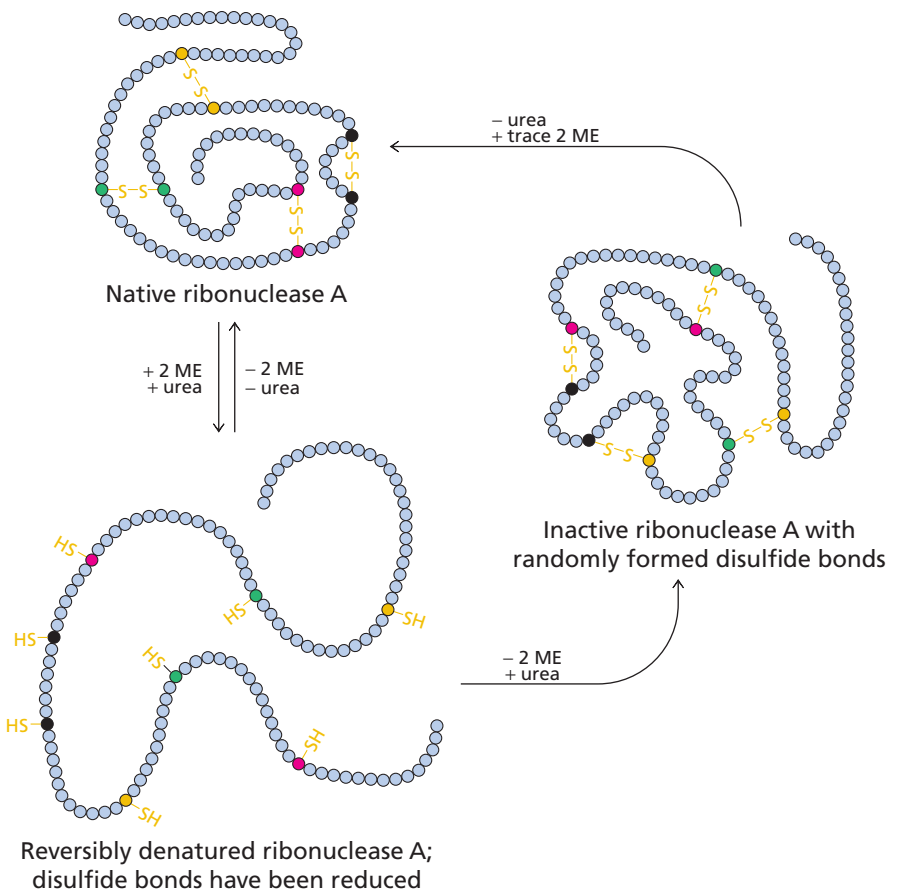
residues), it refolds (renatures) quickly once it is returned to conditions where the native form is stable (e.g., cooled below the melting temperature or removed from a solution containing chaotropic agents). Anfinsen was among the first to show that denatured proteins can refold spontaneously to their native conformation indicating that the information required for the native three-dimensional conformation is contained in the amino acid sequence of the polypeptide chain. In other words, the primary structure determines the tertiary structure.

Denaturation of ribonuclease A with 8 M urea containing 2-mercaptoethanol results in complete loss of tertiary structure and enzymatic activity and yields a polypeptide chain containing eight sulfhydryl groups (Figure 4.35). When 2-mercaptoethanol is removed and oxidation is allowed to occur in the presence of urea, the sulfhydryl groups pair randomly so that only about 1% of the protein population forms the correct four disulfide bonds recovering original enzymatic activity. (If the eight sulfhydryl groups pair randomly, 105 disulfide-bonded structures are possible—7 possible pairings for the first bond, 5 for the second, 3 for the third, and 1 for the fourth ($7 \times 5 \times 3 \times 1 = 105$)—but only one of these structures is correct.) However, when urea and 2-mercaptoethanol are removed simultaneously and dilute solutions of the reduced protein are then exposed to air, ribonuclease A spontaneously regains its native conformation, its correct set of disulfide bonds, and its full enzymatic activity. The inactive proteins containing randomly formed disulfide bonds can be renatured if urea is removed, a small amount of 2-mercaptoethanol is added, and the solution gently warmed. Anfinsen's experiments demonstrate that the correct disulfide bonds can form only after the protein folds into its native conformation. Anfinsen concluded that the renaturation of ribonuclease A is spontaneous, driven entirely by the free energy gained in changing to the stable physiological conformation. This conformation is determined by the primary structure.

Proteins occasionally adopt a nonnative conformation and form inappropriate disulfide bridges when they fold inside a cell. Anfinsen discovered an enzyme, called protein disulfide isomerase (PDI), that catalyzes reduction of these incorrect bonds. All



▲ **Christian B. Anfinsen (1916–1995).** Anfinsen was awarded the Nobel Prize in Chemistry in 1972 for his work on the refolding of proteins.



◄ **Figure 4.35**
Denaturation and renaturation of ribonuclease A. Treatment of native ribonuclease A (top) with urea in the presence of 2-mercaptoethanol unfolds the protein and disrupts disulfide bonds to produce reduced, reversibly denatured ribonuclease A (bottom). When the denatured protein is returned to physiological conditions in the absence of 2-mercaptoethanol, it refolds into its native conformation and the correct disulfide bonds form. However, when 2-mercaptoethanol alone is removed, ribonuclease A reoxidizes in the presence of air, but the disulfide bonds form randomly, producing inactive protein (such as the form shown on the right). When urea is removed, a trace of 2-mercaptoethanol is added to the randomly reoxidized protein, and the solution is warmed gently, the disulfide bonds break and re-form correctly to produce native ribonuclease A.

Downloaded From Vet-eBooks.com

living cells contain such an activity. The enzyme contains two reduced cysteine residues positioned in the active site. When the misfolded protein binds, the enzyme catalyzes a disulfide-exchange reaction whereby the disulfide in the misfolded protein is reduced and a new disulfide bridge is created between the two cysteine residues in the enzyme. The misfolded protein is then released and it can refold into the low-energy native conformation. The structure of the reduced form of *E. coli* disulfide isomerase (DsbA) is shown in Figure 4.24o.

4.11 Protein Folding and Stability

New polypeptides are synthesized in the cell by a translation complex that includes ribosomes, mRNA, and various factors (Chapter 21). As the newly synthesized polypeptide emerges from the ribosome, it folds into its characteristic three-dimensional shape. Folded proteins occupy a low-energy well that makes the native structure much more stable than alternative conformations (Figure 4.36). The *in vitro* experiments of Anfinsen and many other biochemists demonstrate that many proteins can fold spontaneously to reach this low-energy conformation. In this section we discuss the characteristics of those proteins that fold into a stable three-dimensional structure.

It is thought that as a protein folds the first few interactions trigger subsequent interactions. This is an example of cooperative effects in protein folding—the phenomenon whereby the formation of one part of a structure leads to the formation of the remaining parts of the structure. As the protein begins to fold, it adopts lower and lower energies and begins to fall into the energy well shown in Figure 4.36. The protein may become temporarily trapped in a local energy well (shown as small dips in the energy diagram) but eventually it reaches the energy minimum at the bottom of the well. In its final, stable, conformation, the native protein is much less sensitive to degradation than an extended, unfolded polypeptide chain. Thus, native proteins can have half-lives of many cell generations and some molecules may last for decades.

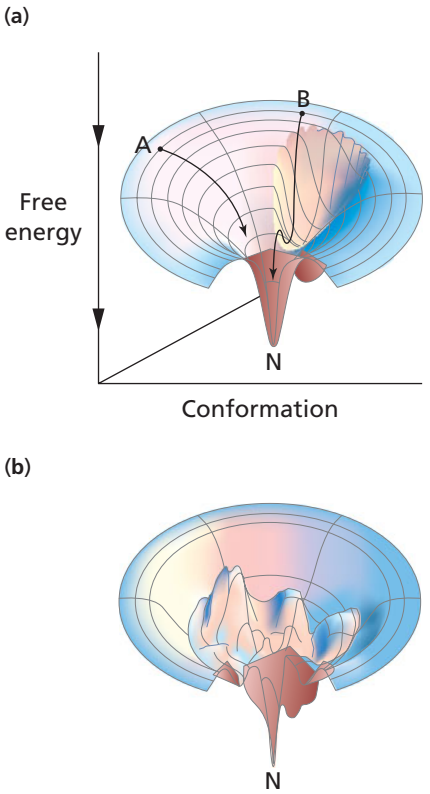
Folding is extremely rapid—in most cases the native conformation is reached in less than a second. Protein folding and stabilization depend on several noncovalent forces including the hydrophobic effect, hydrogen bonding, van der Waals interactions, and charge–charge interactions. Although noncovalent interactions are weak individually, collectively they account for the stability of the native conformations of proteins. The weakness of each noncovalent interaction gives proteins the resilience and flexibility to undergo small conformational changes. (Covalent disulfide bonds also contribute to the stability of certain proteins.)

In multidomain proteins the different domains fold independently of one another as much as possible. One of the reasons for limitations on the size of a domain (usually < 200 residues) is that large domains would fold too slowly if domains were larger than 300 residues. The rate of spontaneous folding would be too slow to be useful.

No actual protein-folding pathway has yet been described in detail but current research is focused on intermediates in the folding pathways of a number of proteins. Several hypothetical folding pathways are shown in Figure 4.37. During protein folding, the polypeptide collapses upon itself due to the hydrophobic effect and elements of secondary structure begin to form. This intermediate is called a molten globule. Subsequent steps involve rearrangement of the backbone chain to form characteristic motifs and, finally, the stable native conformation.

The mechanism of protein folding is one of the most challenging problems in biochemistry. The process is spontaneous and must be largely determined by the primary structure (sequence) of the polypeptide. It should be possible, therefore, to predict the structure of a protein from knowledge of its amino acid sequence. Much progress has been made in recent years by modeling the folding process using fast computers.

In the remainder of this section, we examine the forces that stabilize protein structure in more detail. We will also describe the role of chaperones in protein folding.



▲ **Figure 4.36**
Energy well of protein folding. The funnels represent the free-energy potential of folding proteins. **(a)** A simplified funnel showing two possible pathways to the low-energy native protein. In path B, the polypeptide enters a local low-energy minimum as it folds. **(b)** A more realistic version of the possible free-energy forms of a folding protein with many local peaks and dips.

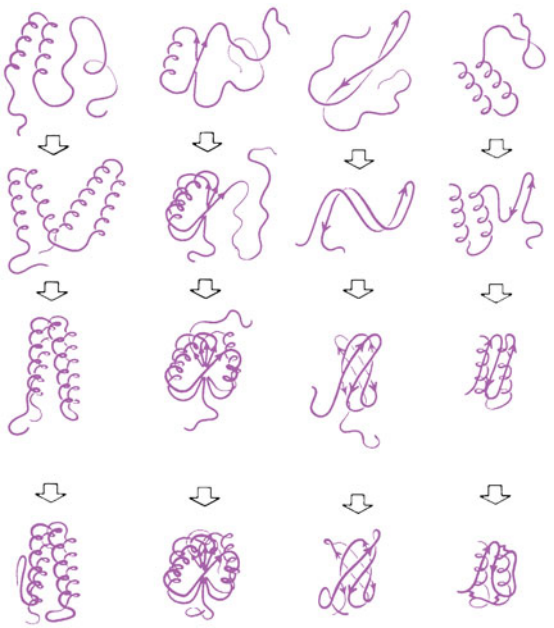
KEY CONCEPT

Most proteins fold spontaneously into a conformation with the lowest energy.

A. The Hydrophobic Effect

Proteins are more stable in water when their hydrophobic side chains are aggregated in the protein interior rather than exposed on the surface to the aqueous medium. Because

Downloaded From Vet-e-Books.com



◀ **Figure 4.37**
Hypothetical protein-folding pathways. The initially extended polypeptide chains form partial secondary structures, then approximate tertiary structures, and finally the unique native conformations. The arrows within the structures indicate the direction from the N- to the C-terminus.

water molecules interact more strongly with each other than with the nonpolar side chains of a protein, the side chains are forced to associate with one another causing the polypeptide chain to collapse into a more compact molten globule. The entropy of the polypeptide decreases as it becomes more ordered. This decrease is more than offset by the increase in solvent entropy as water molecules that were previously bound to the protein are released. (Folding also disrupts extended cages of water molecules surrounding hydrophobic groups.) This overall increase in the entropy of the system provides the major driving force for protein folding.

Whereas nonpolar side chains are driven into the interior of the protein, most polar side chains remain in contact with water on the surface of the protein. The sections of the polar backbone that are forced into the interior of a protein neutralize their polarity by hydrogen bonding to each other, often generating secondary structures. Thus, the hydrophobic nature of the interior not only accounts for the association of hydrophobic residues but also contributes to the stability of helices and sheets. Studies of folding pathways indicate that hydrophobic collapse and formation of secondary structures occur simultaneously

Localized examples of this hydrophobic effect are the interactions of the hydrophobic side of an amphipathic α helix with the protein core (Section 4.4) and the hydrophobic region between β sheets in the β-sandwich structure (Section 4.5). Most of the examples shown in Figures 4.25 and 4.26 contain juxtaposed regions of secondary structure that are stabilized by hydrophobic interactions between the side chains of hydrophobic amino acid residues.

B. Hydrogen Bonding

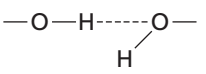
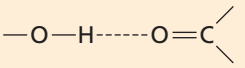
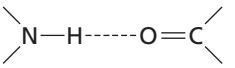
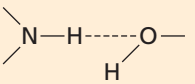
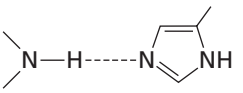
Hydrogen bonds contribute to the cooperativity of folding and help stabilize the native conformations of proteins. The hydrogen bonds in α helices, β sheets, and turns are the first to form, giving rise to defined regions of secondary structure. The final native structure also contains hydrogen bonds between the polypeptide backbone and water, between the polypeptide backbone and polar side chains, between two polar side chains, and between polar side chains and water. Table 4.2 shows some of the many types of hydrogen bonds found in proteins along with their typical bond lengths. Most hydrogen bonds in proteins are of the N—H—O type. The distance between the donor and acceptor atoms varies from 0.26 to 0.34 nm and the bonds may deviate from linearity by up to 40°. Recall that hydrogen bonds within the hydrophobic core of a protein are much more stable than those that form near the surface because the internal hydrogen bonds don't compete with water molecules.

KEY CONCEPT

Entropically driven reactions are reactions where the most important thermodynamic change is an increase in entropy of the system. We can say that the system is much more disordered at the end of the reaction than at the beginning. In the case of hydrophobic interactions, the change in entropy is mostly due to the release of ordered water molecules that shield hydrophobic groups (Section 2.5D).

Downloaded From Vet-eBooks.com

Table 4.2 Examples of hydrogen bonds in proteins

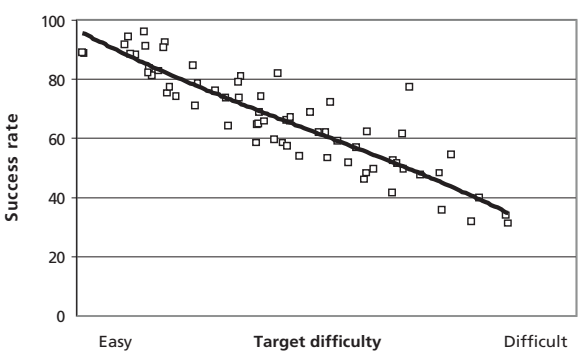
Type of hydrogen bond		Typical distance between donor and acceptor atom (nm)
Hydroxyl-hydroxyl		0.28
Hydroxyl-carbonyl		0.28
Amide-carbonyl		0.29
Amide-hydroxyl		0.30
Amide-imidazole nitrogen		0.31

BOX 4.2 CASP: THE PROTEIN FOLDING GAME

The basic principles of protein folding are reasonably well understood and it seems certain that if a protein has a stable three-dimensional structure it will be determined largely by the primary structure (sequence). This has led to efforts to predict tertiary structure from knowing the amino acid sequence. Biochemists have made huge advances in this theoretical work in the last 30 years.

The value of such work has to be assessed by making predictions of the structure of unknown proteins. This led in 1996 to the beginning of CASP–Critical Assessment of Methods of Protein Structure Prediction. This is a sort of game with no prizes other than the honor of being successful. Protein folding groups are given the amino acid sequences of a number of targets and asked to predict the three-dimensional structure. The targets are drawn from those proteins whose structures have just been determined but the data haven't yet been published. Contestants have only a few weeks to send in their predictions before the actual structures become known.

The results of the 2008 CASP round are shown in the figure. There were 121 targets and thousands of predictions were submitted. Success ranged from nearly 100% for easy proteins to only about 30% for difficult ones. ("Easy" targets are those where the Protein Data Bank (PDB) already contains the structures of several homologous proteins. "Difficult" targets are proteins with new folds that have never been solved.) The success rate for moderately difficult targets has climbed over the years as the prediction methods improved, but there's plenty of opportunity to make winning predictions at the very difficult end of the scale.



Downloaded From Vet-e-Books.com

C. Van der Waals Interactions and Charge–Charge Interactions

Van der Waals contacts between nonpolar side chains also contribute to the stability of proteins. The extent of stabilization due to optimized van der Waals interactions is difficult to determine. The cumulative effect of many van der Waals interactions probably makes a significant contribution to stability because nonpolar side chains in the interior of a protein are densely packed.

Charge–charge interactions between oppositely charged side chains may make a small contribution to the stability of proteins but most ionic side chains are found on the surfaces where they are solvated and can contribute only minimally to the overall stabilization of the protein. Nevertheless, two oppositely charged ions occasionally form an ion pair in the interior of a protein. Such ion pairs are much stronger than those exposed to water.

D. Protein Folding Is Assisted by Molecular Chaperones

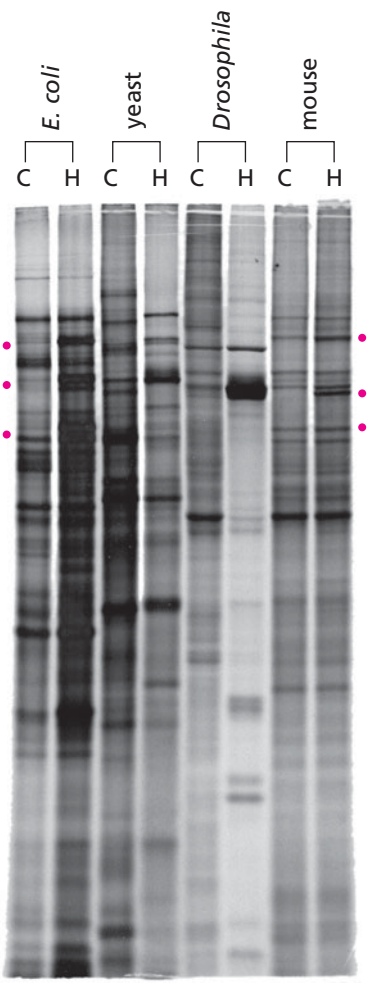
Studies of protein folding have led to two general observations regarding the folding of polypeptide chains into biologically active proteins. First, protein folding does not involve a random search in three-dimensional space for the native conformation. Instead, protein folding appears to be a cooperative, sequential process in which formation of the first few structural elements assists in the alignment of subsequent structural features. [The need for cooperativity is illustrated by a calculation made by Cyrus Levinthal. Consider a polypeptide of 100 residues. If each residue had three possible conformations that could interconvert on a picosecond time scale then a random search of all possible conformations for the complete polypeptide would take 10^{87} seconds—many times the estimated age of the universe (6×10^{17} seconds)!]

Second, to a first approximation the folding pattern and the final conformation of a protein depend on its primary structure. (Many proteins bind metal ions and coenzymes as described in Chapter 7. These external ligands are also required for proper folding.) As we saw in the case of ribonuclease A, simple proteins may fold spontaneously into their native conformations in a test tube without any energy input or assistance. Larger proteins will also fold spontaneously into their native structures since the final conformation represents the minimal free energy form. However, larger proteins are more likely to become temporarily trapped in a local energy well of the type illustrated in Figure 4.36b. The presence of such metastable incorrect conformations at best slows the rate of protein folding and at worst causes the folding intermediates to aggregate and fall out of solution. In order to overcome this problem inside the cell, the rate of correct protein folding is enhanced by a group of ubiquitous special proteins called **molecular chaperones**.

Chaperones increase the *rate* of correct folding of some proteins by binding newly synthesized polypeptides before they are completely folded. They prevent the formation of incorrectly folded intermediates that may trap the polypeptide in an aberrant form. Chaperones can also bind to unassembled protein subunits to prevent them from aggregating incorrectly and precipitating before they are assembled into a complete multisubunit protein.

There are many different chaperones. Most of them are heat shock proteins—proteins that are synthesized in response to temperature increases (heat shock) or other changes that cause protein denaturation *in vivo*. The role of heat shock proteins—now recognized as chaperones—is to repair the damage caused by temperature increases by binding to denatured proteins and helping them to refold rapidly into their native conformation.

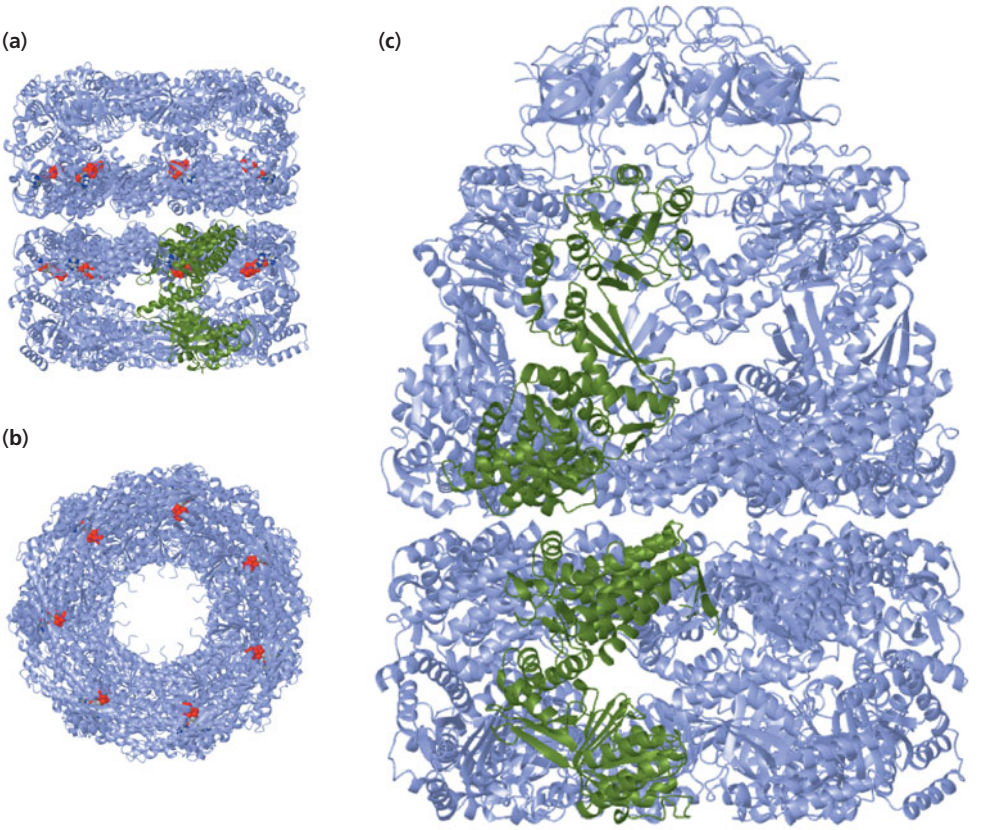
The major heat shock protein is Hsp70 (heat shock protein, M_r = 70,000). This protein is present in all species except for some species of archaebacteria. In bacteria, it is also called DnaK. The normal role of the chaperone Hsp70 is to bind to nascent



▶ **Heat shock proteins.** Proteins were synthesized for a short time in the presence of radioactive amino acids then run on an SDS–polyacrylamide gel. The gel was exposed to film to detect radioactive proteins. The resulting autoradiograph shows only those proteins that were labeled during the time of exposure to radioactive amino acids. Lanes "C" are proteins synthesized at normal growth temperatures, and lanes "H" are proteins synthesized during a short heat shock where cells are shifted to a temperature a few degrees above their normal growth temperature. The induction of heat shock proteins (chaperones) in four different species is shown. Red dots indicate major heat shock proteins: top = Hsp90, middle = Hsp70, bottom = Hsp60(GroEL).

Downloaded From Vet-e-Books.com

► **Figure 4.38**
***Escherichia coli* chaperonin (GroE).** The core structure consists of two identical rings composed of seven GroEL subunits. Unfolded proteins bind to the central cavity. Bound ATP molecules can be identified by their red oxygen atoms. **(a)** Side view. **(b)** Top view showing the central cavity. [PDB 1DER]. **(c)** During folding the size of the central cavity of one of the rings increases and the end is capped by a protein containing seven GroES subunits. [PDB 1AON].

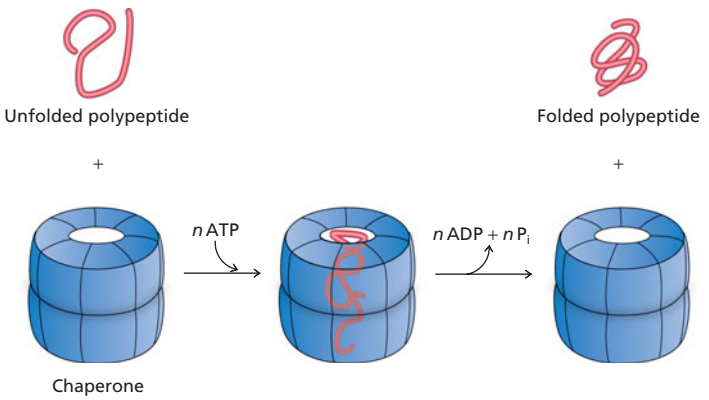


proteins while they are being synthesized in order to prevent aggregation or entrapment in a local low-energy well. The binding and release of nascent polypeptides is coupled to the hydrolysis of ATP and usually requires additional accessory proteins. Hsp70/DnaK is one of the most highly conserved proteins known in all of biology. This indicates that chaperone-assisted protein folding is an ancient and essential requirement for efficient synthesis of proteins with the correct three-dimensional structure.

Another important and ubiquitous chaperone is called chaperonin (also called GroE in bacteria). Chaperonin is also a heat shock protein (Hsp60) that plays an important and essential role in assisting normal protein folding inside the cell.

E. coli chaperonin is a complex multisubunit protein. The core structure consists of two rings containing seven identical GroEL subunits. Each subunit can bind a molecule of ATP (Figure 4.38a). A simplified version of chaperonin-assisted folding is shown in Figure 4.39 . Unfolded proteins bind to the hydrophobic central cavity enclosed by the rings. When folding is complete, the protein is released by hydrolysis of the bound ATP molecules. The actual pathway is more complicated and requires an additional component that serves as a cap sealing one end of the central cavity while the folding process takes place.

Figure 4.39 ►
Chaperonin-assisted protein folding. The unfolded polypeptide enters the central cavity of chaperonin, where it folds. The hydrolysis of several ATP molecules is required for chaperonin function.



Downloaded From Vet-e-Books.com

The cap contains seven GroES subunits forming an additional ring (Figure 4.38c). The conformation of the GroEL ring can be altered during folding to increase the size of the cavity and the role of the cap is to prevent the unfolded protein from being released prematurely.

As mentioned earlier, some proteins tend to aggregate during folding in the absence of chaperones. Aggregation is probably due to temporary formation of hydrophobic surfaces on folding intermediates. The intermediates bind to each other and the result is that they are taken out of solution and are no longer able to explore the conformations represented by the energy funnel shown in Figure 4.36. Chaperonins isolate polypeptide chains in the folding cavity and thus prevent folding intermediates from aggregating. The folding cavity serves as an "Anfinsen cage" that allows the chain to reach the correct low-energy conformation without interference from other folding intermediates.

The central cavity of chaperonin is large enough to accommodate a polypeptide chain of about 630 amino acid residues (M_r = 70,000). Thus, the folding of most small and medium-sized proteins can be assisted by chaperonin. However, only about 5% to 10% of *E. coli* proteins (i.e., about 300 different proteins) appear to interact with chaperonin during protein synthesis. Medium-sized proteins and those of the α/β structural class are more likely to require chaperonin-assisted folding. Smaller proteins are able to fold quickly on their own. Many of the remaining proteins in the cell require other chaperones, such as HSP70/DnaK.

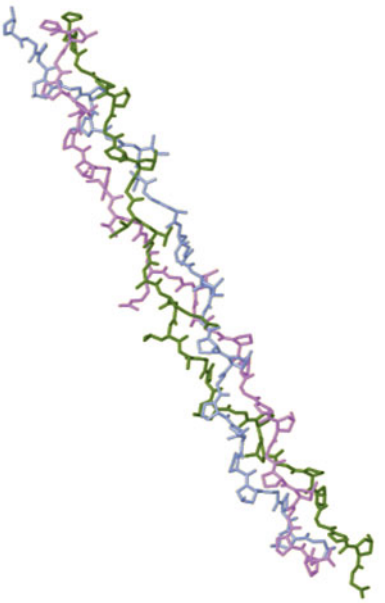
Chaperones appear to inhibit incorrect folding and assembly pathways by forming stable complexes with surfaces on polypeptide chains that are exposed only during synthesis, folding, and assembly. Even in the presence of chaperones, protein folding is spontaneous; for this reason, chaperone-assisted protein folding has been described as assisted self-assembly.

4.12 Collagen, a Fibrous Protein

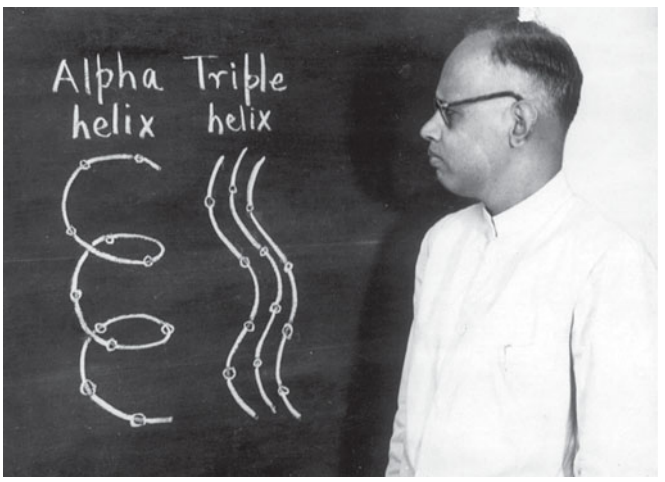
To conclude our examination of the three-dimensional structure of proteins, we examine several proteins to see how their structures are related to their biological functions. The proteins selected for more detailed study are the structural protein collagen, the oxygen-binding proteins myoglobin and hemoglobin (Sections 4.12 to 4.13), and antibodies (Section 4.14).

Collagen is the major protein component of the connective tissue of vertebrates. It makes up about 30% of the total protein in mammals. Collagen molecules have remarkably diverse forms and functions. For example, collagen in tendons forms stiff, ropelike fibers of tremendous tensile strength whereas in skin, collagen takes the form of loosely woven fibers permitting expansion in all directions.

The structure of collagen was worked out by G. N. Ramachandran (famous for his Ramachandran plots, Section 4.3). The molecule consists of three left-handed helical chains coiled around each other to form a right-handed supercoil (Figure 4.40).

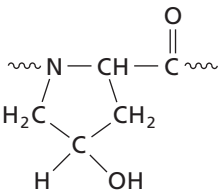


▲ **Figure 4.40**
The human type III collagen triple helix. The extended region of collagen contains three identical subunits (purple, light blue, and green). Three left-handed collagen helices are coiled around one another to form a right-handed supercoil. [PDB 1BKV]

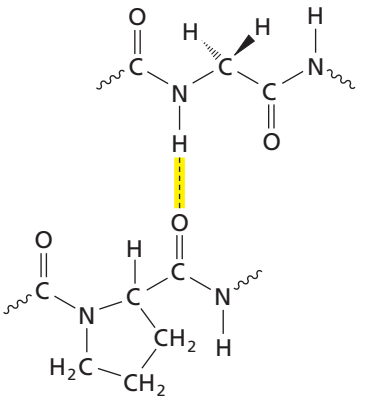


◀ **G.N. Ramachandran (1922–2001).** In this photograph he is illustrating the difference between an α helix and the left-handed triple helix of collagen. Note that he has deliberately drawn the α helix as a left-handed helix and not the standard right-handed form found in most proteins.

Downloaded From Vet-eBooks.com



▲ **Figure 4.41**
4-Hydroxyproline residue. 4-Hydroxyproline residues are formed by enzyme-catalyzed hydroxylation of proline residues.



◄ **Figure 4.42**
Interchain hydrogen bonding in collagen. The amide hydrogen of a glycine residue in one chain is hydrogen-bonded to the carbonyl oxygen of a residue, often proline, in an adjacent chain.

Each left-handed helix in collagen has 3.0 amino acid residues per turn and a pitch of 0.94 nm giving a rise of 0.31 nm per residue. Consequently, a collagen helix is more extended than an α helix and the coiled-coil structure of collagen is not the same as the coiled-coil motif discussed in Section 4.7. (Several proteins unrelated to collagen also form similar three-chain supercoils.)

The collagen triple helix is stabilized by interchain hydrogen bonds. The sequence of the protein in the helical region consists of multiple repeats of the form –Gly–X–Y–, where X is often proline and Y is often a modified proline called 4-hydroxyproline (Figure 4.41). The glycine residues are located along the central axis of the triple helix, where tight packing of the protein strands can accommodate no other residue. For each –Gly–X–Y– triplet, one hydrogen bond forms between the amide hydrogen atom of glycine in one chain and the carbonyl oxygen atom of residue X in an adjacent chain (Figure 4.42). Hydrogen bonds involving the hydroxyl group of hydroxyproline may also stabilize the collagen triple helix. Unlike the more common α helix, the collagen helix has no intrachain hydrogen bonds.

In addition to hydroxyproline, collagen contains an additional modified amino acid residue called 5-hydroxylysine (Figure 4.43). Some hydroxylysine residues are covalently bonded to carbohydrate residues, making collagen a glycoprotein. The role of this glycosylation is not known.

The requirement for vitamin C is explained in Section 7.9.

Hydroxyproline and hydroxylysine residues are formed when specific proline and lysine residues are hydroxylated after incorporation into the polypeptide chains of collagen. The hydroxylation reactions are catalyzed by enzymes and require ascorbic acid (vitamin C). Hydroxylation is impaired in the absence of vitamin C, and the triple helix of collagen is not assembled properly.

The limited conformational flexibility of proline and hydroxyproline residues prevents the formation of α helices in collagen chains and also makes collagen somewhat rigid. (Recall that proline is almost never found in α helices.) The presence of glycine residues at every third position allows collagen chains to form a tightly wound left-handed helix that accommodates the proline residues. (Recall that the flexibility of glycine residues tends to disrupt the right-handed α helix.)

Collagen triple helices aggregate in a staggered fashion to form strong, insoluble fibers. The strength and rigidity of collagen fibers result in part from covalent

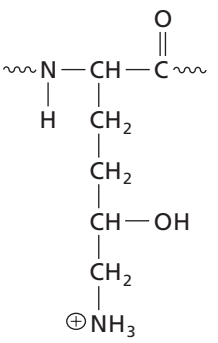
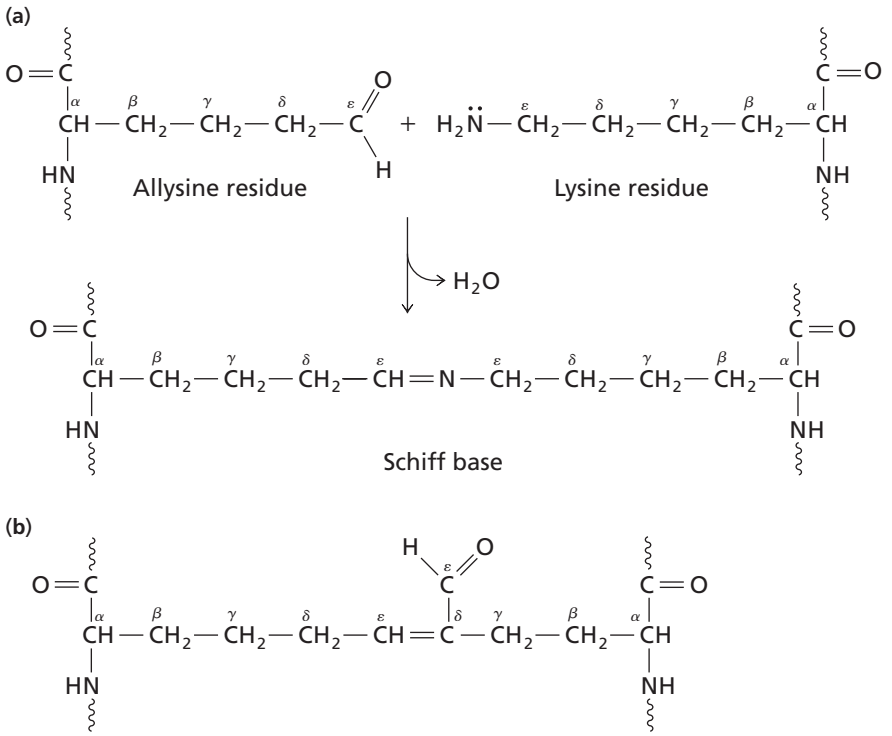


Figure 4.43 ►
5-Hydroxylysine residue. 5-Hydroxylysine residues are formed by enzyme-catalyzed hydroxylation of lysine residues.

Downloaded From Vet-e-Books.com



◄ **Figure 4.44**
Covalent cross-links in collagen. (a) An allysine residue condenses with a lysine residue to form an intermolecular Schiff-base cross-link. **(b)** Two allysine residues condense to form an intramolecular cross-link.

cross-links. The $—CH_2NH_3^+$ groups of the side chains of some lysine and hydroxylysine residues are converted enzymatically to aldehyde groups (—CHO), producing allysine and hydroxyallysine residues. Allysine residues (and their hydroxy derivatives) react with the side chains of lysine and hydroxylysine residues to form **Schiff bases**, complexes formed between carbonyl groups and amines (Figure 4.44a). These Schiff bases usually form between collagen molecules. Allysine residues also react with other allysine residues by aldol condensation to form cross-links, usually between the individual strands of the triple helix (Figure 4.44b). Both types of cross-links are converted to more stable bonds during the maturation of tissues, but the chemistry of these conversions is unknown.

BOX 4.3 STRONGER THAN STEEL

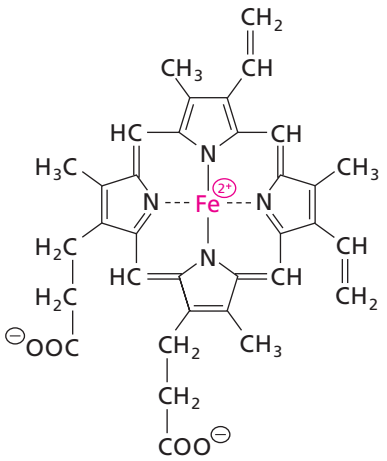
Not all fibrous proteins are composed of α helices. Silk is composed of a number of proteins that are predominantly β strands. The dragline silk of the spider, *Nephila clavipes*, for example, contains two proteins called spidroin 1 and spidroin 2. Both proteins contain multiple stretches of alanine residues separated by residues that are mostly glycine. The structure of this silk is not known in spite of major efforts by many laboratories. However, it is known that the proteins contain extensive regions of β strands.

There are many different kinds of spider silk and spiders have specialized glands for each type. The silk fiber produced by the major ampulate gland is called dragline silk; it is the fiber that spiders use to drop out of danger or anchor their webs. This silk fiber is quite literally stronger than steel cable. Materials manufactured from dragline silk would be very useful in a number of applications, one of which would be personal armor because dragline silk is stronger than Kevlar. So far it has not been possible to make significant amounts of silk in the laboratory without relying on spiders.

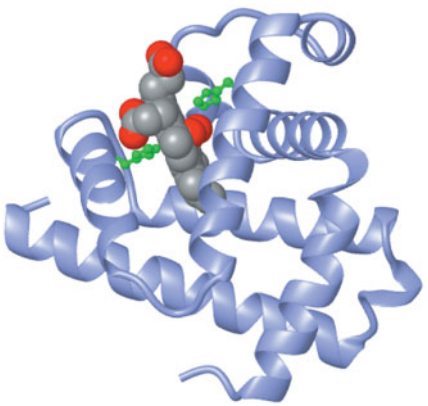


***Nephila clavipes*, the golden silk spider.** ►

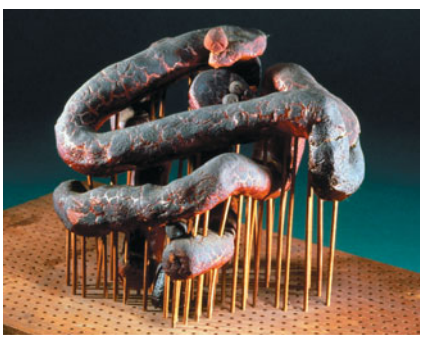
Downloaded From Vet-e-Books.com



▲ **Figure 4.45**
Chemical structure of the Fe(II)-protoporphyrin IX heme group in myoglobin and hemoglobin. The porphyrin ring provides four of the six ligands that surround the iron atom.



▲ **Figure 4.46**
Sperm whale (*Physeter catodon*) oxymyoglobin. Myoglobin consists of eight α helices. The heme prosthetic group binds oxygen (red). His-64 (green) forms a hydrogen bond with oxygen, and His-93 (green) is complexed to the iron atom of the heme. [PDB 1A6M].



▲ John Kendrew's original model of myoglobin determined from his X-ray diffraction data in the 1950s. The model is made of plasticine. It was the first three-dimensional model of a protein.

4.13 Structures of Myoglobin and Hemoglobin

Like most proteins, myoglobin (Mb) and the related protein hemoglobin (Hb) carry out their biological functions by selectively and reversibly binding other molecules—in this case, molecular oxygen (O_2). Myoglobin is a relatively small monomeric protein that facilitates the diffusion of oxygen in vertebrates. It is responsible for supplying oxygen to muscle tissue in reptiles, birds, and mammals. Hemoglobin is a larger tetrameric protein that carries oxygen in blood.

The red color associated with the oxygenated forms of myoglobin and hemoglobin (e.g., the red color of oxygenated blood) is due to a heme prosthetic group (Figure 4.45). (A prosthetic group is a protein-bound organic molecule essential for the activity of the protein.) Heme consists of a tetrapyrrole ring system (protoporphyrin IX) complexed with iron. The four pyrrole rings of this system are linked by methene (—CH═) bridges so that the unsaturated porphyrin is highly conjugated and planar. The bound iron is in the ferrous, or Fe^{2+}, oxidation state; it forms a complex with six ligands, four of which are the nitrogen atoms of protoporphyrin IX. (Other proteins, such as cytochrome *a* and cytochrome *c*, contain different porphyrin/heme groups.)

Myoglobin is a member of a family of proteins called globins. The tertiary structure of sperm whale myoglobin shows that the protein consists of a bundle of eight α helices (Figure 4.46). It is a member of the all-α structural category. The globin fold has several groups of α helices that form a layered structure. Adjacent helices in each layer are tilted at an angle that allows the side chains of the amino acid residues to interdigitate.

The interior of myoglobin is made up almost exclusively of hydrophobic amino acid residues, particularly those that are highly hydrophobic—valine, leucine, isoleucine, phenylalanine, and methionine. The surface of the protein contains both hydrophilic and hydrophobic residues. As is the case with most proteins, the tertiary structure of myoglobin is stabilized by hydrophobic interactions within the core. Folding of the polypeptide chain is driven by the energy minimization that results from formation of this hydrophobic core.

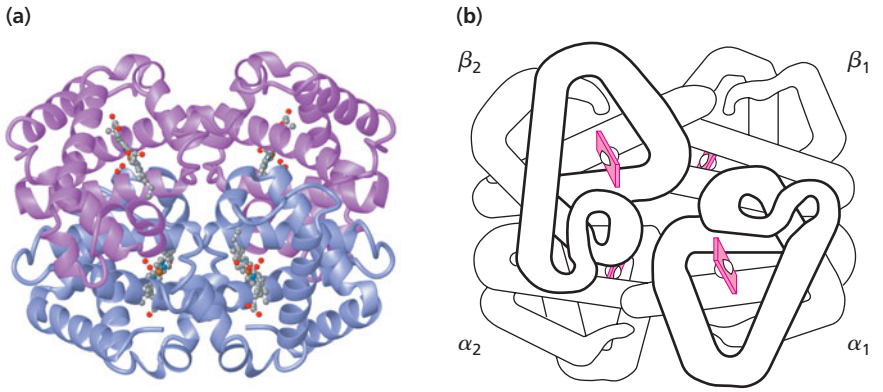
The heme prosthetic group of myoglobin occupies a hydrophobic cleft formed by three α helices and two loops. The binding of the porphyrin moiety to the polypeptide is due to a number of weak interactions including hydrophobic interactions, van der Waals contacts, and hydrogen bonds. There are no covalent bonds between the porphyrin and the amino acid side chains of myoglobin. The iron atom of heme is the site of oxygen binding as shown in Figure 4.46. Two histidine residues interact with the iron atom and the bound oxygen. Accessibility of the heme group to molecular oxygen depends on slight movement of nearby amino acid side chains. We will see later that the hydrophobic crevices of myoglobin and hemoglobin are essential for the reversible binding of oxygen.

In vertebrates, O_2 is bound to molecules of hemoglobin for transport in red blood cells, or erythrocytes. Viewed under a microscope, a mature mammalian erythrocyte is a biconcave disk that lacks a nucleus or other internal membrane-enclosed compartments (Figure 4.47). A typical human erythrocyte is filled with approximately 3×10^8 hemoglobin molecules.

Hemoglobin is more complex than myoglobin because it is a multisubunit protein. In adult mammals, hemoglobin contains two different globin subunits called *α-globin* and *β-globin*. Hemoglobin is an $\alpha_2\beta_2$ tetramer—it contains two α chains and two β chains. Each of these globin subunits is similar in structure and sequence to myoglobin reflecting their evolution from a common ancestral globin gene in primitive chordates.

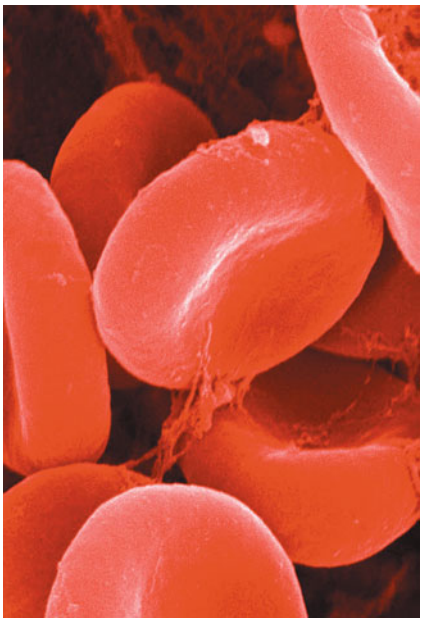
Each of the four globin subunits contains a heme prosthetic group identical to that found in myoglobin. The α and β subunits face each other across a central cavity (Figure 4.48). The tertiary structure of each of the four chains is almost identical to that of myoglobin (Figure 4.49). The α chain has seven α helices, and the β chain has eight. (Two short α helices found in β-globin and myoglobin are fused into one larger one in α-globin) Hemoglobin, however, is not simply a tetramer of myoglobin molecules. Each α chain interacts extensively with a β chain so hemoglobin is actually a dimer of $\alpha\beta$ subunits. We will see in the following section that the presence of multiple subunits is responsible for oxygen-binding properties that are not possible with single-chain myoglobin.

Downloaded From Vet-e-Books.com



▲ **Figure 4.48**
Human (*Homo sapiens*) oxyhemoglobin. (a) Structure of human oxyhemoglobin showing two α and two β subunits. Heme groups are shown as stick models. [PDB 1HND]. **(b)** Schematic diagram of the hemoglobin tetramer. The heme groups are red.

▼ **Figure 4.47**
Scanning electron micrograph of mammalian erythrocytes. Each cell contains approximately 300 million hemoglobin molecules. The cells have been artificially colored.



4.14 Oxygen Binding to Myoglobin and Hemoglobin

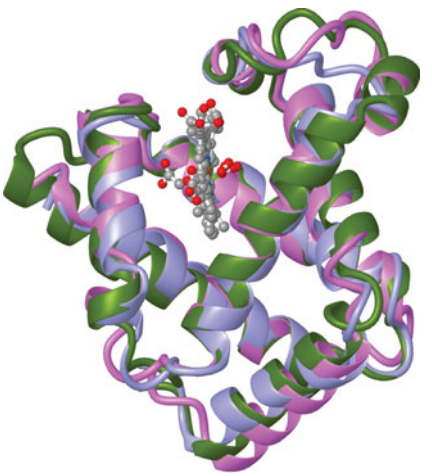
The oxygen-binding activities of myoglobin and hemoglobin provide an excellent example of how protein structure relates to physiological function. These proteins are among the most intensely studied proteins in biochemistry. They were the first complex proteins whose structure was determined by X-ray crystallography (Section 4.2). A number of the principles described here for oxygen-binding proteins also hold true for the enzymes that we will study in Chapters 5 and 6. In this section we examine the chemistry of oxygen binding to heme, the physiology of oxygen binding to myoglobin and hemoglobin, and the regulatory properties of hemoglobin.

A. Oxygen Binds Reversibly to Heme

We will use myoglobin as an example of oxygen binding to the heme prosthetic group but the same principles apply to hemoglobin. The reversible binding of oxygen is called **oxygenation**. Oxygen-free myoglobin is called *deoxymyoglobin* and the oxygen-bearing molecule is called *oxymyoglobin*. (The two forms of hemoglobin are called *deoxyhemoglobin* and *oxyhemoglobin*.)

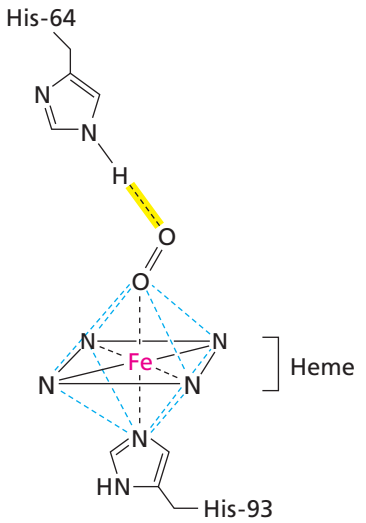
Some substituents of the heme prosthetic group are hydrophobic—this feature allows the prosthetic group to be partially buried in the hydrophobic interior of the myoglobin molecule. Recall from Figure 4.46 that there are two polar residues, His-64 and His-93, situated near the heme group. In oxymyoglobin, six ligands are coordinated to the ferrous iron, with the ligands in octahedral geometry around the metal cation (Figures 4.50 and 4.51). Four of the ligands are the nitrogen atoms of the tetrapyrrole ring system; the fifth ligand is an imidazole nitrogen from His-93 (called the proximal histidine); and the sixth ligand is molecular oxygen bound between the iron and the imidazole side chain of His-64 (called the distal histidine). In deoxymyoglobin, the iron is coordinated to only five ligands because oxygen is not present. The nonpolar side chains of Val-68 and Phe-43, shown in Figure 4.51, contribute to the hydrophobicity of the oxygen-binding pocket and help hold the heme group in place. Several side chains block the entrance to the heme-containing pocket in both oxymyoglobin and deoxymyoglobin. The protein structure in this region must vibrate, or breathe, rapidly to allow oxygen to bind and dissociate.

The hydrophobic crevice of the globin polypeptide holds the key to the ability of myoglobin and hemoglobin to suitably bind and release oxygen. Free heme does not reversibly bind oxygen in aqueous solution; instead, the Fe^{2+} of the heme is almost instantly oxidized to Fe^{3+}. (Oxidation is equivalent to the loss of an electron, as described in Section 6.1C. Reduction is the gain of an electron. Oxidation and reduction refer to the transfer of electrons and not to the presence or absence of oxygen molecules.)

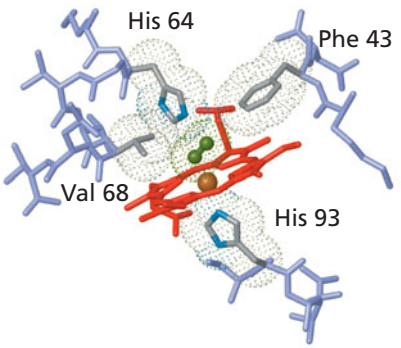


▲ **Figure 4.49**
Tertiary structure of myoglobin, α-globin, and β-globin. The orientations of the individual α-globin and β-globin subunits of hemoglobin have been shifted in order to reveal the similarities in tertiary structure. The three structures have been superimposed. All of the structures are from the oxygenated forms shown in Figures 4.46 and 4.48. Color code: α-globin (blue), β-globin (purple), myoglobin (green).

Downloaded From Vet-e-Books.com



▲ **Figure 4.50**
Oxygen-binding site of sperm whale oxymyoglobin. The heme prosthetic group is represented by a parallelogram with a nitrogen atom at each corner. The blue dashed lines illustrate the octahedral geometry of the coordination complex.



▲ **Figure 4.51**
The oxygen-binding site in sperm whale myoglobin. Fe(II) (orange) lies in the plane of the heme group. Oxygen (green) is bound to the iron atom and the amino acid side chain of His-64. Val-68 and Phe-43 contribute to the hydrophobic environment of the oxygen-binding site. [PDB 1AGM].

The structure of myoglobin and hemoglobin prevents the permanent transfer of an electron or irreversible oxidation thereby ensuring the reversible binding of molecular oxygen for transport. The ferrous iron atom of heme in hemoglobin is partially oxidized when O_2 is bound. An electron is temporarily transferred toward the oxygen atom that is attached to the iron so that the molecule of dioxygen is partially reduced. If the electron were transferred completely to the oxygen, the complex would be $Fe^{3+}—O_2^{\ominus}$ (a superoxide anion attached to ferric iron). The globin crevice prevents complete electron transfer and enforces return of the electron to the iron atom when O_2 dissociates.

B. Oxygen-Binding Curves of Myoglobin and Hemoglobin

Oxygen binds reversibly to myoglobin and hemoglobin. The extent of binding at equilibrium depends on the concentration of the protein and the concentration of oxygen. This relationship is depicted in oxygen-binding curves (Figure 4.52). In these figures, the fractional saturation (Y) of a fixed amount of protein is plotted against the concentration of oxygen (measured as the partial pressure of gaseous oxygen, pO_2). The fractional saturation of myoglobin or hemoglobin is the fraction of the total number of molecules that are oxygenated.

$$Y = \frac{[MbO_2]}{[MbO_2] + [Mb]} \qquad \textbf{(4.3)}$$

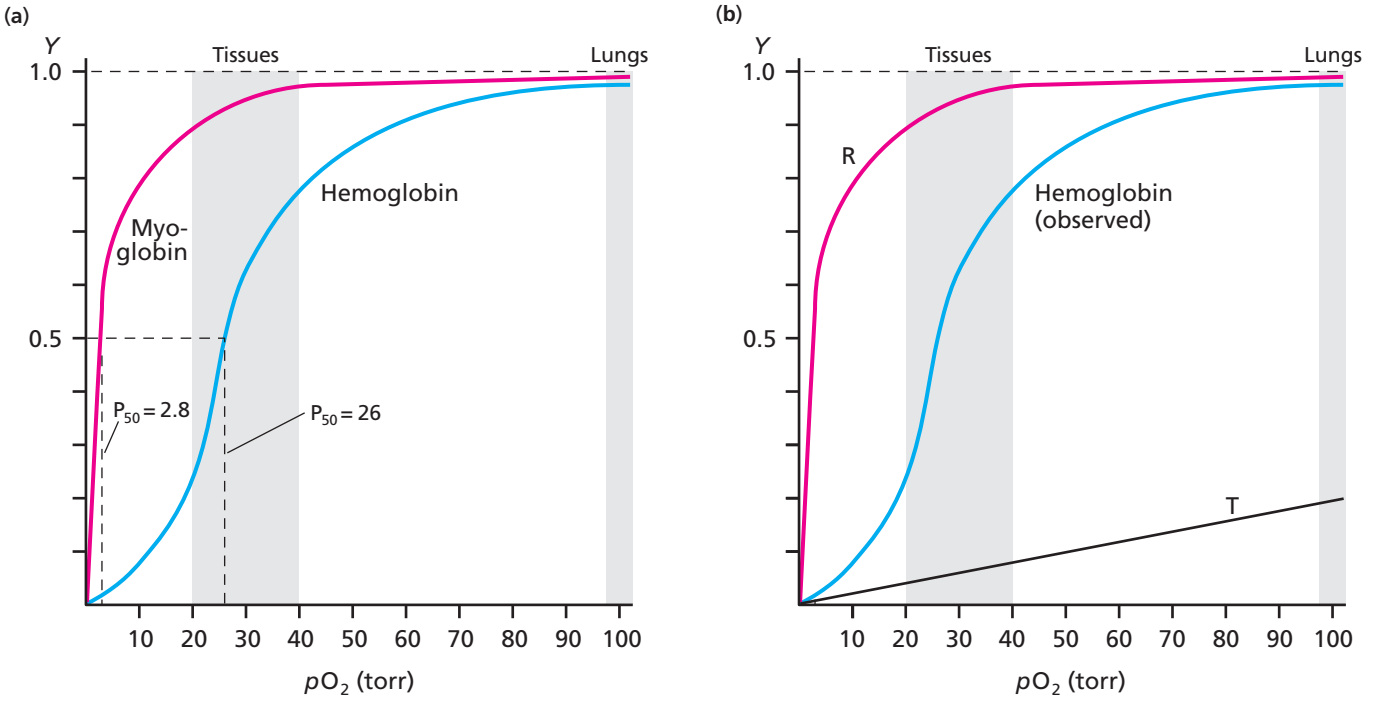
The oxygen-binding curve of myoglobin is hyperbolic (Figure 4.52), indicating that there is a single equilibrium constant for the binding of O_2 to the macromolecule. In contrast, the curve depicting the relationship between oxygen concentrations and binding to hemoglobin is sigmoidal. Sigmoidal (S-shaped) binding curves indicate that more than one molecule of ligand is binding to each protein. In this case, up to four molecules of O_2 bind to hemoglobin, one per heme group of the tetrameric protein. The shape of the curve indicates that the oxygen-binding sites of hemoglobin interact such that the binding of one molecule of oxygen to one heme group facilitates binding of oxygen molecules to the other hemes. The oxygen affinity of hemoglobin increases as each oxygen molecule is bound. This interactive binding phenomenon is termed *positive cooperativity* of binding.

The partial pressure at half-saturation (P_{50}) is a measure of the affinity of the protein for O_2. A low P_{50} indicates a high affinity for oxygen since the protein is half-saturated with oxygen at a low oxygen concentration; similarly, a high P_{50} signifies a low affinity. Myoglobin molecules are half-saturated at a pO_2 of 2.8 torr (1 atmosphere = 760 torr). The P_{50} for hemoglobin is much higher (26 torr) reflecting its lower affinity for oxygen. The heme prosthetic groups of myoglobin and hemoglobin are identical but the affinities of these groups for oxygen differ because the microenvironments provided by the proteins are slightly different. Oxygen affinity is an intrinsic property of the protein. It is similar to the equilibrium binding/dissociation constants that are commonly used to describe the binding of ligands to other proteins and enzymes (Section 4.9).

As Figure 4.52 shows, at the high pO_2 found in the lungs (about 100 torr) both myoglobin and hemoglobin are nearly saturated. However, at pO_2 values below about 50 torr, myoglobin is still almost fully saturated whereas hemoglobin is only partially saturated. Much of the oxygen carried by hemoglobin in erythrocytes is released within the capillaries of tissues where pO_2 is low (20 to 40 torr). Myoglobin in muscle tissue then binds oxygen released from hemoglobin. The differential affinities of myoglobin and hemoglobin for oxygen thus lead to an efficient system for oxygen delivery from the lungs to muscle.

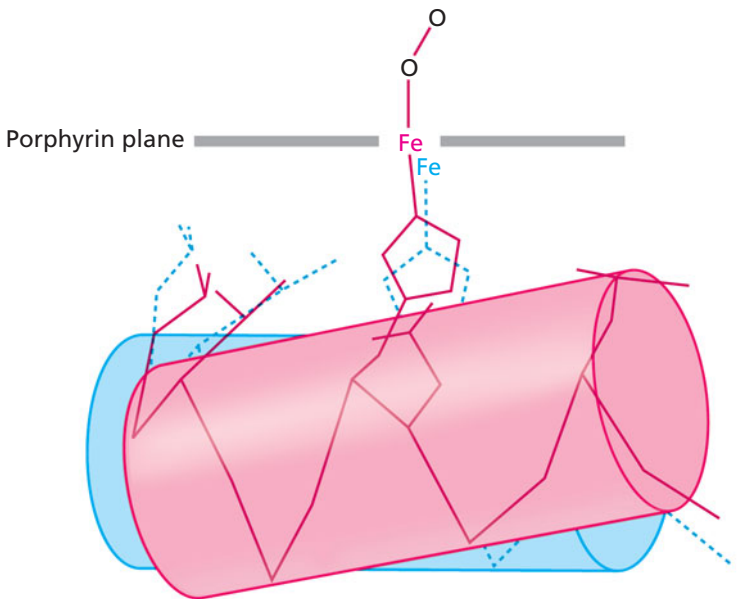
The cooperative binding of oxygen by hemoglobin can be related to changes in the protein conformation that occur on oxygenation. Deoxyhemoglobin is stabilized by several intra- and intersubunit ion pairs. When oxygen binds to one of the subunits, it causes a movement that disrupts these ion pairs and favors a slightly different conformation. The movement is triggered by the reactivity of the heme iron atom (Figure 4.53). In deoxyhemoglobin, the iron atom is bound to only five ligands (as in myoglobin). It is slightly larger than the cavity within the porphyrin ring and lies below the plane of the ring. When O_2—the sixth ligand—binds to the iron atom, the electronic structure of the iron

Downloaded From Vet-eBooks.com



▲ **Figure 4.52**

Oxygen-binding curves of myoglobin and hemoglobin. (a) Comparison of myoglobin and hemoglobin. The fractional saturation (Y) of each protein is plotted against the partial pressure of oxygen (pO_2). The oxygen-binding curve of myoglobin is hyperbolic, with half-saturation ($Y = 0.5$) at an oxygen pressure of 2.8 torr. The oxygen-binding curve of hemoglobin in whole blood is sigmoidal, with half-saturation at an oxygen pressure of 26 torr. Myoglobin has a greater affinity than hemoglobin for oxygen at all oxygen pressures. In the lungs, where the partial pressure of oxygen is high, hemoglobin is nearly saturated with oxygen. In tissues, where the partial pressure of oxygen is low, oxygen is released from oxygenated hemoglobin and transferred to myoglobin. **(b)** O_2 binding by the different states of hemoglobin. The oxy (R, or high-affinity) state of hemoglobin has a hyperbolic binding curve. The deoxy (T, or low-affinity) state of hemoglobin would also have a hyperbolic binding curve but with a much higher concentration for half-saturation. Solutions of hemoglobin containing mixtures of low- and high-affinity forms show sigmoidal binding curves with intermediate oxygen affinities.



▲ **Figure 4.53**

Conformational changes in a hemoglobin chain induced by oxygenation. When the heme iron of a hemoglobin subunit is oxygenated (red), the proximal histidine residue is pulled toward the porphyrin ring. The helix containing the histidine also shifts position, disrupting ion pairs that cross-link the subunits of deoxyhemoglobin (blue).

Downloaded From Vet-eBooks.com

changes, its diameter decreases, and it moves into the plane of the porphyrin ring pulling the helix that contains the proximal histidine. The change in tertiary structure results in a slight change in quaternary structure and this allows the remaining subunits to bind oxygen more readily. The entire tetramer appears to shift from the deoxy to the oxy conformation only after at least one oxygen molecule binds to each $\alpha\beta$ dimer. (For further discussion, see Section 5.9C.)

The conformational change of hemoglobin is responsible for the positive cooperativity of binding seen in the binding curve (Figure 4.52a). The shape of the curve is due to the combined effect of the two conformations (Figure 4.52b). The completely deoxygenated form of hemoglobin has a low affinity for oxygen and thus exhibits a hyperbolic binding curve with a very high concentration of half-saturation. Only a small amount of hemoglobin is saturated at low oxygen concentrations. As the concentration of oxygen increases, some of the hemoglobin molecules bind a molecule of oxygen and this increases their affinity for oxygen so that they are more likely to bind additional oxygen. This causes the sigmoidal curve and also a sharp rise in binding. More molecules of hemoglobin are in the oxy conformation. If all of the hemoglobin molecules were in the oxy conformation, a solution would exhibit a hyperbolic binding curve. Release of the oxygen molecules allows the hemoglobin molecule to re-form the ion pairs and resume the deoxy conformation.

The two conformations of hemoglobin are called the T (tense) and R (relaxed) states, using the standard terminology for such conformational changes. In hemoglobin, the deoxy conformation, which resists oxygen binding, is considered the inactive (T) state, and the oxy conformation, which facilitates oxygen binding, is considered the active (R) state. The R and T states are in dynamic equilibrium.

BOX 4.4 EMBRYONIC AND FETAL HEMOGLOBINS

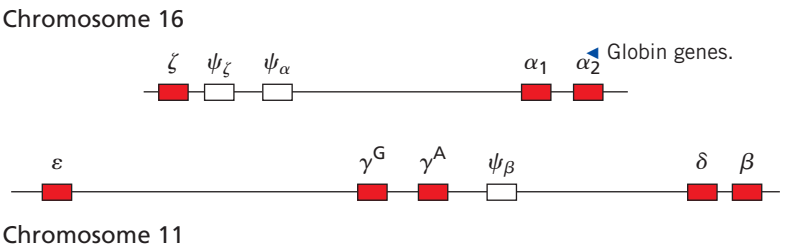
The human α globin genes are located on chromosome 16 in a cluster of related members of the globin gene family. There are two different genes encoding α globin: α_1 and α_2. Upstream of these genes there is another functional gene called ζ (zeta). The locus includes two nonfunctional pseudogenes, one related to ζ ($\psi\zeta$) and the other derived from a duplicated α globin gene (ψ_α).

The β globin gene is on chromosome 11 and it is also located at a locus where there are other members of the globin gene family. The functional genes are δ, two related γ globin genes (γ^A and γ^G), and an ε (epsilon) gene. This locus also contains a pseudogene related to β (ψ_β).

The other globin genes encode hemoglobin subunits that are expressed in the early embryo and in the fetus. The embryonic hemoglobins are called Gower 1 ($\zeta_2\varepsilon_2$), Gower 2 ($\alpha_2\varepsilon_2$), and Portland ($\zeta_2\gamma_2$). The fetal hemoglobin has the subunit composition $\alpha_2\gamma_2$. The adult hemoglobins are $\alpha_2\beta_2$ and $\alpha_2\delta_2$.

During early embryogenesis, the growing embryo gets oxygen from the mother's blood through the placenta. The concentration of oxygen in the embryo is much lower than the concentration of oxygen in adult blood. The embryonic hemoglobins compensate by binding oxygen much more tightly, their P_{50} values range from 4 to 12 torr—much lower than the value of adult hemoglobin (26 torr). The fetal hemoglobins bind oxygen less tightly than the embryonic hemoglobin but tighter than the adult hemoglobins (P_{50} = 20 torr).

Expression of the various globin genes is carefully regulated so that the right genes are transcribed at the right time. Sometimes mutations arise where the fetal γ globin genes are inappropriately expressed in adults. The result is a phenotype known as Hereditary Persistence of Fetal Hemoglobin (HPFH). This is just one of hundreds of hemoglobin variants that have been detected in humans. You can read about them on a database called Online Mendelian Inheritance in Man (OMIM), the most complete and accurate database of human genetic diseases (ncbi.nlm.nih.gov/omim).



► Human fetus.



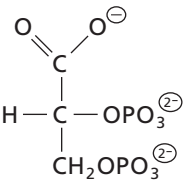
Downloaded From Vet-e-Books.com



▲ Julian Voss-Andreae created a sculpture called "Heart of Steel (Hemoglobin)" in 2005 in the City of Lake Oswego, Oregon. The sculpture is a depiction of a hemoglobin molecule with a bound oxygen atom. The original sculpture was shiny steel (left). After 10 days (middle) it had started to rust as the iron in the steel reacted with oxygen in the atmosphere. After several months (right) the sculpture was completely rust colored.

C. Hemoglobin Is an Allosteric Protein

The binding and release of oxygen by hemoglobin are regulated by **allosteric interactions** (from the Greek *allos*, "other"). In this respect, hemoglobin—a carrier protein, not an enzyme—resembles certain regulatory enzymes (Section 5.9). Allosteric interactions occur when a specific small molecule, called an **allosteric modulator**, or **allosteric effector**, binds to a protein (usually an enzyme) and modulates its activity. The allosteric modulator binds reversibly at a site separate from the functional binding site of the protein. An effector molecule may be an activator or an inhibitor. A protein whose activity is modulated by allosteric effectors is called an **allosteric protein**.



▲ 2,3-Bisphospho-D-glycerate (2,3BPG).

Allosteric modulation is accomplished by small but significant changes in the conformations of allosteric proteins. It involves cooperativity of binding that is regulated by binding of the allosteric effector to a distinct site that doesn't overlap the normal binding site of a substrate, product, or transported molecule such as oxygen. An allosteric protein is in an equilibrium in which its active shape (R state) and its inactive shape (T state) are rapidly interconverting. A substrate, which obviously binds at the active site (to heme in hemoglobin), binds most avidly when the protein is in the R state. An allosteric inhibitor, which binds at an allosteric or regulatory site, binds most avidly to the T state. The binding of an allosteric inhibitor to its own site causes the allosteric protein to change rapidly from the R state to the T state. The binding of a substrate to the active site (or an allosteric activator to the allosteric site) causes the reverse change. The change in conformation of an allosteric protein caused by binding or release of an effector extends from the allosteric site to the functional binding site (the active site). The activity level of an allosteric protein depends on the relative proportions of molecules in the R and T forms and these, in turn, depend on the relative concentrations of the substrates and modulators that bind to each form.

The synthesis of 2,3BPG in red blood cells is described in Box 11.2 (Chapter 11).

The molecule 2,3-*bis*phospho-D-glycerate (2,3BPG) is an allosteric effector of mammalian hemoglobin. The presence of 2,3BPG in erythrocytes raises the P_{50} for binding of oxygen to adult hemoglobin to about 26 torr—much higher than the P_{50} for oxygen binding to purified hemoglobin in aqueous solution (about 12 torr). In other words, 2,3BPG in erythrocytes substantially lowers the affinity of deoxyhemoglobin for oxygen. The concentrations of 2,3BPG and hemoglobin within erythrocytes are nearly equal (about 4.7 mM).

Downloaded From Vet-e-Books.com

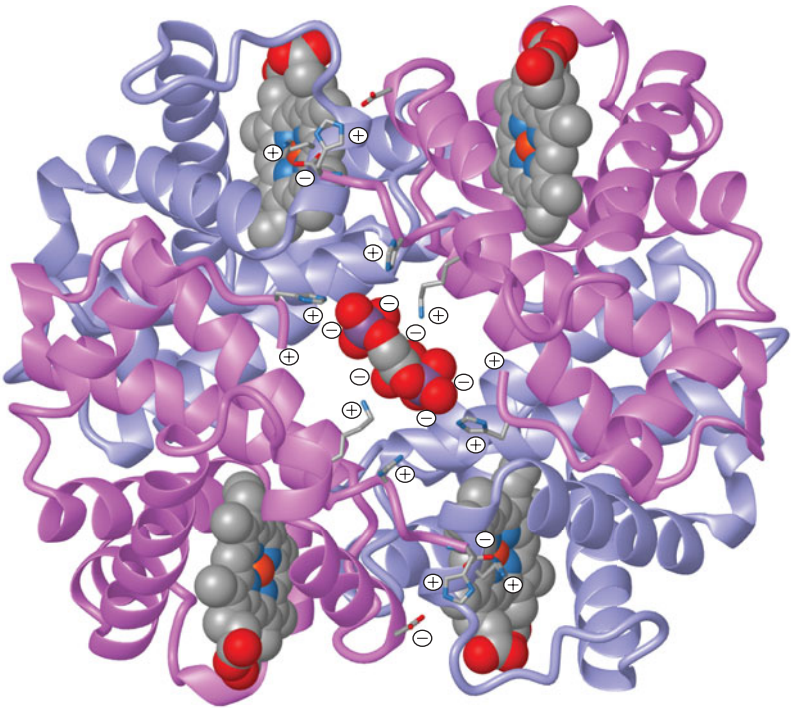


Figure 4.54 ▶
Binding of 2,3BPG to deoxyhemoglobin. The central cavity of deoxyhemoglobin is lined with positively charged groups that are complementary to the carboxylate and phosphate groups of 2,3BPG. Both 2,3BPG and the ion pairs shown help stabilize the deoxy conformation. The α subunits are shown in pink, the β subunits in blue, and the heme prosthetic groups in red.

R and T conformations are explained more thoroughly in Section 5.10, "Theory of Allostery."

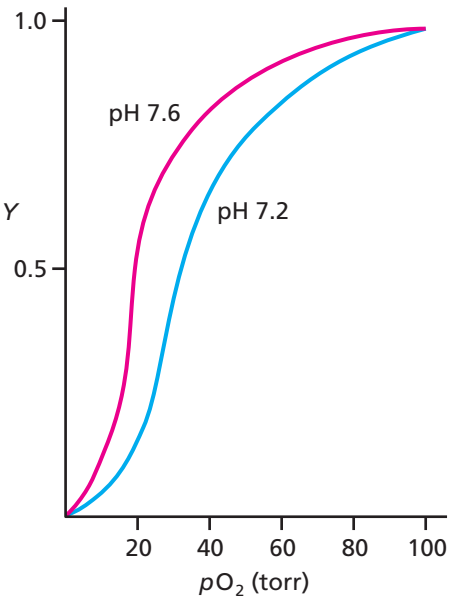
The effector 2,3BPG binds in the central cavity of hemoglobin between the two β subunits. In this binding pocket there are six positively charged side chains and the N-terminal α-amino group of each β chain forming a cationic binding site (Figure 4.54). In deoxyhemoglobin, these positively charged groups can interact electrostatically with the five negative charges of 2,3BPG. When 2,3BPG is bound, the deoxy conformation (the T state, which has a low affinity for O_2) is stabilized and conversion to the oxy conformation (the R or high-affinity state) is inhibited. In oxyhemoglobin, the β chains are closer together and the allosteric binding site is too small to accommodate 2,3BPG. The reversibly bound ligands O_2 and 2,3BPG have opposite effects on the R $\rightleftharpoons$ T equilibrium. Oxygen binding increases the proportion of hemoglobin molecules in the oxy (R) conformation and 2,3BPG binding increases the proportion of hemoglobin molecules in the deoxy (T) conformation. Because oxygen and 2,3BPG have different binding sites, 2,3BPG is a true allosteric effector.

In the absence of 2,3BPG, hemoglobin is nearly saturated at an oxygen pressure of about 20 torr. Thus, at the low partial pressure of oxygen that prevails in the tissues (20 to 40 torr), hemoglobin without 2,3BPG would not unload its oxygen. In the presence of equimolar 2,3BPG, however, hemoglobin is only about one-third saturated at 20 torr. The allosteric effect of 2,3BPG causes hemoglobin to release oxygen at the low partial pressures of oxygen in the tissues. In muscle, myoglobin can bind some of the oxygen that is released.

Additional regulation of the binding of oxygen to hemoglobin involves carbon dioxide and protons, both of which are products of aerobic metabolism. CO_2 decreases the affinity of hemoglobin for O_2 by lowering the pH inside red blood cells. Enzyme-catalyzed hydration of CO_2 in erythrocytes produces carbonic acid, H_2CO_3, which dissociates to form bicarbonate and a proton thereby lowering the pH.

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^{\oplus} + HCO_3^{\ominus} \qquad \textbf{(4.4)}$$

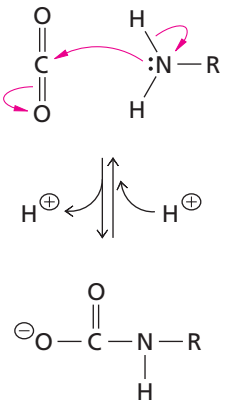
The lower pH leads to protonation of several groups in hemoglobin. These groups then form ion pairs that help stabilize the deoxy conformation. The increase in the concentration of CO_2 and the concomitant decrease in pH raise the P_{50} of hemoglobin (Figure 4.55). This phenomenon, called the Bohr effect, increases the efficiency of the oxygen delivery system. In inhaling lungs, where the CO_2 level is low, O_2 is readily picked up by hemoglobin; in metabolizing tissues, where the CO_2 level is relatively high and the pH is relatively low, O_2 is readily unloaded from oxyhemoglobin.



▲ **Figure 4.55**
Bohr effect. Lowering the pH decreases the affinity of hemoglobin for oxygen.

Downloaded From Vet-e-Books.com

Carbon dioxide is transported from the tissues to the lungs in two ways. Most CO_2 produced by metabolism is transported as dissolved bicarbonate ions but some carbon dioxide is carried by hemoglobin itself the form of carbamate adducts (Figure 4.56). At the pH of red blood cells (7.2) and at high concentrations of CO_2, the unprotonated amino groups of the four N-terminal residues of deoxyhemoglobin (pK_a values between 7 and 8) can react reversibly with CO_2 to form carbamate adducts. The carbamates of oxyhemoglobin are less stable than those of deoxyhemoglobin. When hemoglobin reaches the lungs, where the partial pressure of CO_2 is low and the partial pressure of O_2 is high, hemoglobin is converted to its oxygenated state and the CO_2 that was bound is released.



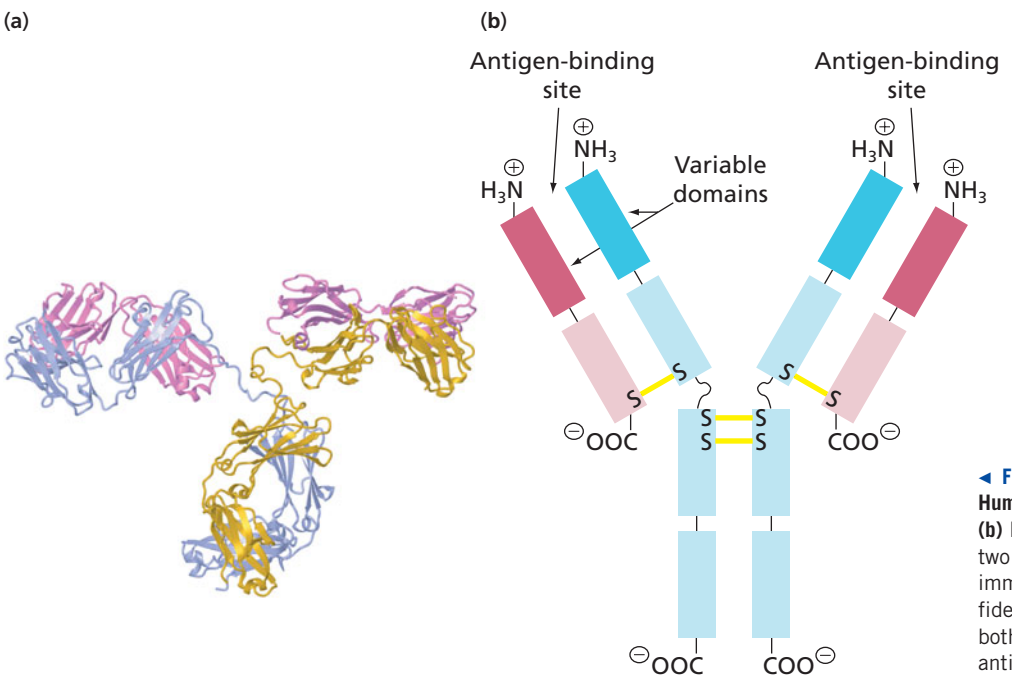
▲ **Figure 4.56**
Carbamate adduct. Carbon dioxide produced by metabolizing tissues can react reversibly with the N-terminal residues of the globin chains of hemoglobin, converting them to carbamate adducts.

4.15 Antibodies Bind Specific Antigens

Vertebrates possess a complex immune system that eliminates foreign substances including infectious bacteria and viruses. As part of this defense system, vertebrates synthesize proteins called **antibodies** (also known as immunoglobulins) that specifically recognize and bind **antigens**. Many different types of foreign compounds can serve as antigens that produce an immune response. Antibodies are synthesized by white blood cells called lymphocytes—each lymphocyte and its descendants synthesize the same antibody. Because animals are exposed to many foreign substances over their lifetimes, they develop a huge array of antibody-producing lymphocytes that persist at low levels for many years and can later respond to the antigen during reinfection. The memory of the immune system is the reason certain infections do not recur in an individual despite repeated exposure. Vaccines (inactivated pathogens or analogs of toxins) administered to children are effective because immunity established in childhood lasts through adulthood.

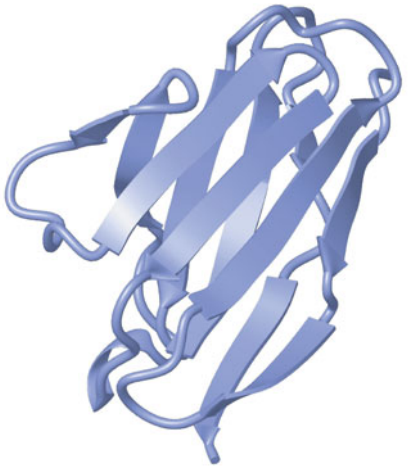
When an antigen—either novel or previously encountered—binds to the surface of lymphocytes, these cells are stimulated to proliferate and produce soluble antibodies for secretion into the bloodstream. The soluble antibodies bind to the foreign organism or substance forming antibody–antigen complexes that precipitate and mark the antigen for destruction by a series of interacting proteases or by lymphocytes that engulf the antigen and digest it intracellularly.

The most abundant antibodies in the bloodstream are of the immunoglobulin G class (IgG). These are Y-shaped oligomers composed of two identical light chains and two identical heavy chains connected by disulfide bonds (Figure 4.57). Immunoglobulins are glycoproteins containing covalently bound carbohydrates attached to the heavy chains. The N-termini of pairs of light and heavy chains are close together. Light chains contain two domains and heavy chains contain four domains. Each of the domains consists of



◀ **Figure 4.57**
Human antibody structure. (a) Structure. **(b)** Diagram. Two heavy chains (blue) and two light chains (red) of antibodies of the immunoglobulin G class are joined by disulfide bonds (yellow). The variable domains of both the light and heavy chains (where antigen binds) are colored more darkly.

Downloaded From Vet-e-Books.com

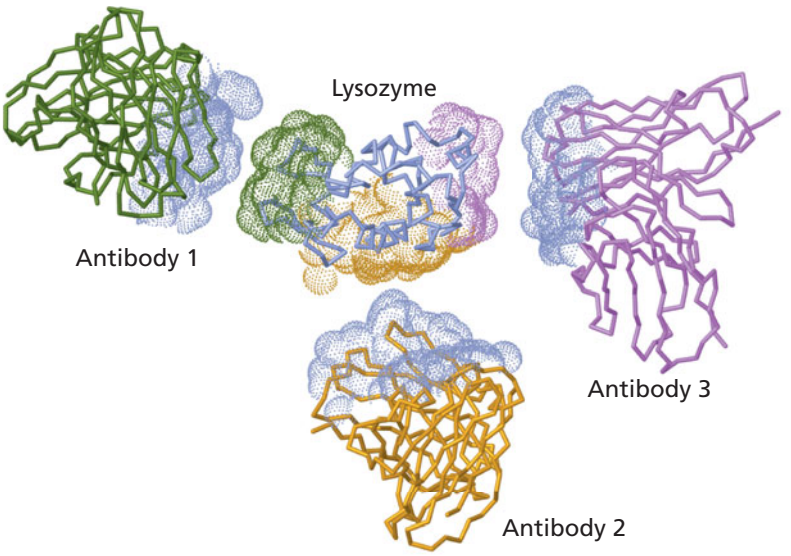


▲ **Figure 4.58**
The immunoglobulin fold. The domain consists of a sandwich of two antiparallel β sheets. [PDB 1REI].

about 110 residues assembled into a common motif called the immunoglobulin fold whose characteristic feature is a sandwich composed of two antiparallel β sheets (Figure 4.58). This domain structure is found in many other proteins of the immune system.

The N-terminal domains of antibodies are called the variable domains because of their sequence diversity. They determine the specificity of antigen binding. X-ray crystallographic studies have shown that the antigen-binding site of a variable domain consists of three loops, called hypervariable regions, that differ widely in size and sequence. The loops from a light chain and a heavy chain combine to form a barrel, the upper surface of which is complementary to the shape and polarity of a specific antigen. The match between the antigen and antibody is so close that there is no space for water molecules between the two. The forces that stabilize the interaction of antigen with antibody are primarily hydrogen bonds and electrostatic interactions. An example of the interaction of antibodies with a protein antigen is shown in Figure 4.59.

Antibodies are used in the laboratory for the detection of small quantities of various substances because of their remarkable antigen-binding specificity. In a common type of immunoassay, fluid containing an unknown amount of antigen is mixed with a solution of labeled antibody and the amount of antibody–antigen complex formed is measured. The sensitivity of these assays can be enhanced in a variety of ways to make them suitable for diagnostic tests.



▲ **Figure 4.59**
Binding of three different antibodies to an antigen (the protein lysozyme). The structures of the three antigen–antibody complexes have been determined by X-ray crystallography. This composite view, in which the antigen and antibodies have been separated, shows the surfaces of the antigen and antibodies that interact. Only parts of the three antibodies are shown.

Summary

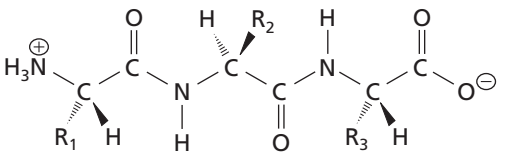
1. Proteins fold into many different shapes, or conformations. Many proteins are water-soluble, roughly spherical, and tightly folded. Others form long filaments that provide mechanical support to cells and tissues. Membrane proteins are integral components of membranes or are associated with membranes.
2. There are four levels of protein structure: primary (sequence of amino acid residues), secondary (regular local conformation, stabilized by hydrogen bonds), tertiary (compacted shape of the entire polypeptide chain), and quaternary (assembly of two or more polypeptide chains into a multisubunit protein).
3. The three-dimensional structures of biopolymers, such as proteins can be determined by X-ray crystallography and NMR spectroscopy.
4. The peptide group is polar and planar. Rotation around the N—Cα and Cα—C bonds is described by φ and ψ.
5. The α helix, a common secondary structure, is a coil containing approximately 3.6 amino acid residues per turn. Hydrogen bonds between amide hydrogens and carbonyl oxygens are roughly parallel to the helix axis.

Downloaded From Vet-e-Books.com

6. The other common type of secondary structure, β structure, often consists of either parallel or antiparallel β strands that are hydrogen-bonded to each other to form β sheets.
7. Most proteins include stretches of nonrepeating conformation, including turns and loops that connect α helices and β strands.
8. Recognizable combinations of secondary structural elements are called motifs.
9. The tertiary structure of proteins consists of one or more domains, which may have recognizable structures and may be associated with particular functions.
10. In proteins that possess quaternary structure, subunits are usually held together by noncovalent interactions.
11. The native conformation of a protein can be disrupted by the addition of denaturing agents. Renaturation may be possible under certain conditions.
12. Folding of a protein into its biologically active state is a sequential, cooperative process driven primarily by the hydrophobic effect. Folding can be assisted by chaperones.
13. Collagen is the major fibrous protein of connective tissues. The three left-handed helical chains of collagen form a right-handed supercoil.
14. The compact, folded structures of proteins allow them to selectively bind other molecules. The heme-containing proteins myoglobin and hemoglobin bind and release oxygen. Oxygen binding to hemoglobin is characterized by positive cooperativity and allosteric regulation.
15. Antibodies are multidomain proteins that bind foreign substances, or antigens, marking them for destruction. The variable domains at the ends of the heavy and light chains interact with the antigen.

Problems

1. Examine the following tripeptide:

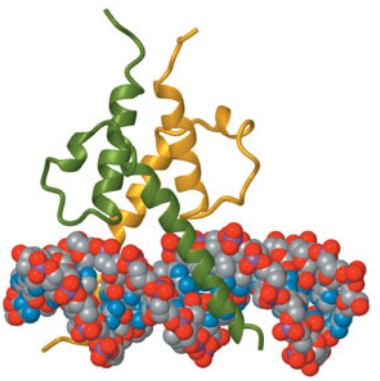


   (a) Label the α-carbon atoms and draw boxes around the atoms of each peptide group.
   (b) What do the R groups represent?
   (c) Why is there limited free rotation around the carbonyl C═O to N amide bonds?
   (d) Assuming that the chemical structure represents the correct conformation of the peptide linkage, are the peptide groups in the *cis* or the *trans* conformation?
   (e) Which bonds allow rotation of peptide groups with respect to each other?

2. (a) Characterize the hydrogen-bonding pattern of (1) an α helix and (2) a collagen triple helix.
   (b) Explain how the amino acid side chains are arranged in each of these helices.
3. Explain why (1) glycine and (2) proline residues are not commonly found in α helices.
4. A synthetic 20 amino acid polypeptide named Betanova was designed as a small soluble molecule that would theoretically form stable β-sheet structures in the absence of disulfide bonds. NMR of Betanova in solution indicates that it does, in fact, form a three-stranded antiparallel β sheet. Given the sequence of Betanova below:
   (a) Draw a ribbon diagram for Betanova indicating likely residues for each hairpin turn between the β strands.
   (b) Show the interactions that are expected to stabilize this β-sheet structure.

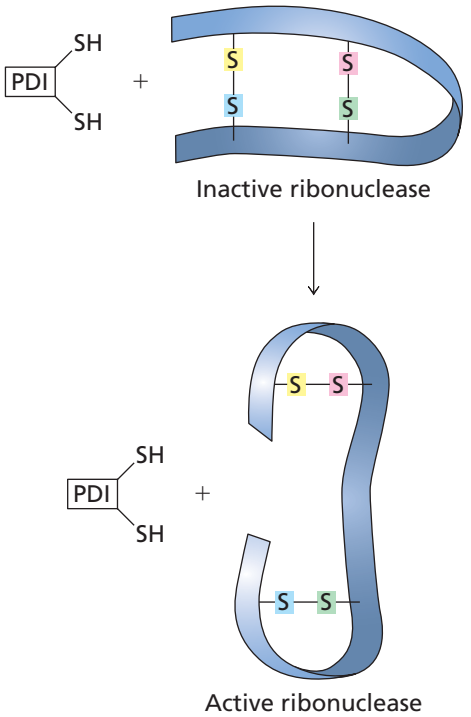
   Betanova RGWSVQNGKYTNNGKTTEGR

5. Each member of an important family of 250 different DNA-binding proteins is composed of a dimer with a common protein motif. This motif permits each DNA-binding protein to recognize and bind to specific DNA sequences. What is the common protein motif in the structure below?



6. Refer to Figure 4.21 to answer the following questions.
   (a) To which of the four major domain categories does the middle domain of pyruvate kinase (PK) belong (all α all β, α/β, $\alpha + \beta$)?
   (b) Describe any characteristic domain "fold" that is prominent in this middle domain of PK.
   (c) Identify two other proteins that have the same fold as the middle domain of pyruvate kinase.
7. Protein disulfide isomerase (PDI) markedly increases the rate of correct refolding of the inactive ribonuclease form with random disulfide bonds (Figure 4.35). Show the mechanism for the PDI-catalyzed rearrangement of a nonnative (inactive) protein with incorrect disulfide bonds to the native (active) protein with correct disulfide bonds.

Downloaded From Vet-e-Books.com



8. Myoglobin contains eight α helices, one of which has the following sequence:

–Gln–Gly–Ala–Met–Asn–Lys–Ala–Leu–Glu–His–Phe–Arg–Lys–Asp–Ile–Ala–Ala–

Which side chains are likely to be on the side of the helix that faces the interior of the protein? Which are likely to be facing the aqueous solvent? Account for the spacing of the residues facing the interior.

9. Homocysteine is an α-amino acid containing one more methylene group in its side chain than cysteine (side chain = $—CH_2CH_2SH$). Homocysteinuria is a genetic disease characterized by elevated levels of homocysteine in plasma and urine, as well as skeletal deformities due to defects in collagen structure. Homocysteine reacts readily with allysine under physiological conditions. Show this reaction and suggest how it might lead to defective cross-linking in collagen.

10. The larval form of the parasite *Schistosoma mansoni* infects humans by penetrating the skin. The larva secretes enzymes that catalyze the cleavage of peptide bonds between residues X and Y in the sequence –Gly–Pro–X–Y– (X and Y can be any of several amino acids). Why is this enzyme activity important for the parasite?

11. (a) How does the reaction of carbon dioxide with water help explain the Bohr effect? Include the equation for the formation of bicarbonate ion from CO_2 and water, and explain the effects of $H^{\oplus}$ and CO_2 on hemoglobin oxygenation.

(b) Explain the physiological basis for the intravenous administration of bicarbonate to shock victims.

12. Fetal hemoglobin (Hb F) contains serine in place of the cationic histidine at position 143 of the β chains of adult hemoglobin (Hb A). Residue 143 faces the central cavity between the β chains.

(a) Why does 2,3BPG bind more tightly to deoxy Hb A than to deoxy Hb F?

(b) How does the decreased affinity of Hb F for 2,3BPG affect the affinity of Hb F for O_2?

(c) The P_{50} for Hb F is 18 torr, and the P_{50} for Hb A is 26 torr. How do these values explain the efficient transfer of oxygen from maternal blood to the fetus?

13. Amino acid substitutions at the $\alpha\beta$ subunit interfaces of hemoglobin may interfere with the $R \rightleftharpoons T$ quaternary structural changes that take place on oxygen binding. In the hemoglobin variant Hb_{Yakima}, the R form is stabilized relative to the T form, and P_{50} = 12 torr. Explain why the mutant hemoglobin is less efficient than normal hemoglobin (P_{50} = 26 torr) in delivering oxygen to working muscle, where O_2 may be as low as 10 to 20 torr.

14. The spider venom from the Chilean Rose Tarantula (*Grammostola spatulata*) contains a toxin that is a 34-amino acid protein. It is thought to be a globular protein that partitions into the lipid membrane to exert its effect. The sequence of the protein is:

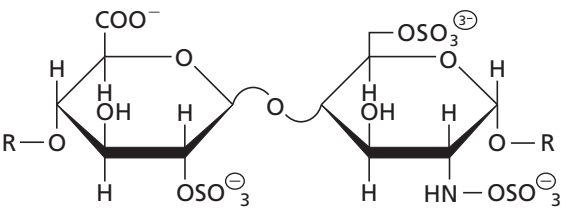
ECGKFMWKCKNSNDCCKDLVCSSRWKWCVLASPF

(a) Identify the hydrophobic and highly hydrophilic amino acids in the protein.

(b) The protein is thought to have a hydrophobic face that interacts with the lipid membrane. How can the hydrophobic amino acids far apart in sequence interact to form a hydrophobic face?

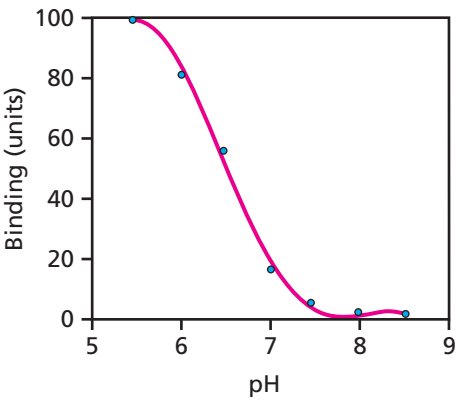
[Adapted from Lee, S. and MacKinnon, R. (2004). *Nature* 430: 232–235.]

15. Selenoprotein P is an unusual extracellular protein that contains 8–10 selenocysteine residues and has a high content of cysteine and histidine residues. Selenoprotein P is found both as a plasma protein and as a protein strongly associated with the surface of cells. The association of selenoprotein P with cells is proposed to occur through the interaction of selenoprotein P with high-molecular-weight carbohydrate compounds classified as glycosaminoglycans. One such compound is heparin (see structure on next page). Binding studies of selenoprotein P to heparin were carried out under different pH conditions. The results are shown in the graph on next page.



(a) How is the binding of selenoprotein P to heparin dependent upon pH?

(b) Give possible structural reasons for the binding dependence.



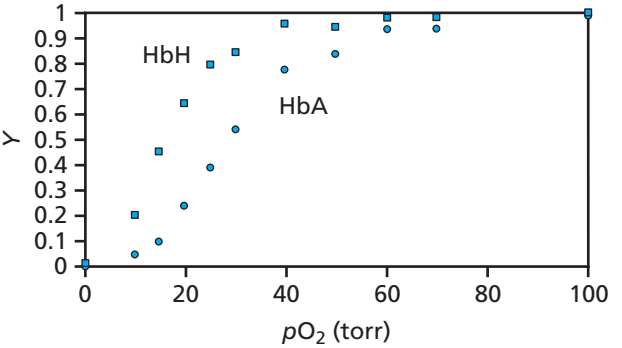
(*Hint:* Use the information about which amino acids are abundant in selenoprotein P in your answer).

[Adapted from Arteel, G. E., Franken, S., Kappler, J., and Sies, H. (2000). *Biol. Chem.* 381:265–268.]

Downloaded From Vet-e-Books.com

16. Gelatin is processed collagen that comes from the joints of animals. When gelatin is mixed with hot water, the triple helix structure unwinds and the chains separate, becoming random coils that dissolve in the water. As the dissolved gelatin mixture cools, the collagen forms a matrix that traps water; as a result, the mixture turns into the jiggling semisolid mass that is recognizable as Jell-O™. The directions on a box of gelatin include the following: "Chill until slightly thickened, then add 1 to 2 cups cooked or raw fruits or vegetables. Fresh or frozen pineapple must be cooked before adding." If the pineapple is not cooked, the gelatin will not set properly. Pineapple belongs to a group of plants called Bromeliads and contains a protease called bromelain. Explain why pineapple must be cooked before adding to gelatin.

17. Hb Helsinki (HbH) is a hemoglobin mutant in which the lysine residue at position 82 has been replaced with methionine. The mutation is in the beta chain, and residue 82 is found in the central cavity of hemoglobin. The oxygen binding curves for normal adult hemoglobin (HbA, ●) and HbH (■) at pH 7.4 in the presence of a physiological concentration of 2,3BPG are shown in the graph.



[Adapted from Ikkala, E., Koskela, J., Pikkarainen, P., Rahiala, E.L., El-Hazmi, M. A., Nagai, K., Lang, A., and Lehmann, H. *Acta Haematol.* (1976). 56:257–275.]

Explain why the curve for HbH is shifted from the curve for HbA. Does this mutation stabilize the R or T state? What result does this mutation have on oxygen affinity?

Selected Readings

General

Clothia, C., and Gough, J. (2009). Genomic and structural aspects of protein evolution. *Biochem. J.* 419:15–28. doi: 10,1042/BJ20090122.

Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties*, 2nd ed. (New York: W. H. Freeman), Chapters 4–7.

Fersht, A. (1998). *Structure and Mechanism in Protein Structure* (New York: W. H. Freeman).

Goodsell, D., and Olson, A. J. (1993). Soluble proteins: size, shape, and function. *Trends Biochem. Sci.* 18:65–68.

Goodsell, D. S., and Olson, A. J. (2000). Structural symmetry and protein function. *Annu. Rev. Biophys, Biomolec. Struct.* 29:105–153.

Kyte, J. (1995). *Structure in Protein Chemistry* (New York: Garland).

Protein Structure

Branden, C., and Tooze, J. (1991). *Introduction to Protein Structure* 2nd ed. (New York: Garland).

Chothia, C., Hubbard, T., Brenner, S., Barns, H., and Murzin, A. (1997). Protein folds in the all-β and all-α classes. *Annu. Rev. Biophys. Biomol. Struct.* 26:597–627.

Edison, A. S. (2001). Linus Pauling and the planar peptide bond. *Nat. Struct. Biol.* 8:201–202.

Harper, E. T., and Rose, G. D. (1993). Helix stop signals in proteins and peptides: the capping box. *Biochemistry* 32:7605–7609.

Phizicky, E., and Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiol. Rev.* 59:94–123.

Rhodes, G. (1993). *Crystallography Made Crystal Clear* (San Diego: Academic Press).

Richardson, J. S., and Richardson, D. C. (1989). Principles and patterns of protein conformation. In *Prediction of Protein Structure and the Principles of Protein Conformation*, G. D. Fasman, ed. (New York: Plenum), pp. 1–98.

Wang, Y., Liu, C., Yang, D., and Yu, H. (2010). *Pin1At* encoding a peptidyl-prolyl *cis/trans* isomerase regulates flowering time in arabidopsis. *Molec. Cell.* 37:112–122.

Uversky, V. N., and Dunker, A. K. (2010). Understanding protein non-folding. *Biochim. Biophys. Acta.* 1804:1231–1264.

Protein Folding and Stability

Daggett, V., and Fersht, A. R. (2003). Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* 28:18–25.

Dill, K. A. Ozkan, S. B., Shell, M. S., and Weik, T. R. (2008). The protein folding problem. *Annu. Rev. Biophys.* 37:289–316.

Feldman, D. E., and Frydman, J. (2000). Protein folding *in vivo*: the importance of molecular chaperones. *Curr. Opin. Struct. Biol.* 10:26–33.

Kryshtafovych, A., Fidelis, K., and Moult, J. (2009). CASP8 results in context of previous experiments. *Proteins.* 77(suppl 9):217–228.

Matthews, B. W. (1993). Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62:139–160.

Saibil, H. R. and Ranson, N. A. (2002). The chaperonin folding machine. *Trends Biochem. Sci.* 27:627–632.

Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998). Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.* 67:581–608.

Smith, C. A. (2000). How do proteins fold? *Biochem. Ed.* 28:76–79.

Specific Proteins

Ackers, G. K., Doyle, M. L., Myers, D., and Daugherty, M. A. (1992). Molecular code for cooperativity in hemoglobin. *Science* 255:54–63.

Brittain, T. (2002). Molecular aspects of embryonic hemogloin function. *Molec. Aspects Med.* 23:293–342.

Davies, D. R., Padlan, E. A., and Sheriff, S. (1990). Antibody-antigen complexes. *Annu. Rev. Biochem.* 59:439–473.

Eaton, W. A., Henry, E. R., Hofrichter, J., and Mozzarelli, A. (1999). Is cooperative binding by hemoglobin really understood? *Nature Struct. Biol.* 6(4):351–357.

Kadler, K. (1994). Extracellular matrix 1: fibril-forming collagens. *Protein Profile* 1:519–549.

Liu, R., and Ochman, H. (2007). Stepwise formation of the bacterial flagellar system. *Proc. Natl. Acad. Sci.* (USA). 104:7116–7121.

Perutz, M. F. (1978). Hemoglobin structure and respiratory transport. *Sci. Am.* 239(6):92–125.

Perutz, M. F., Wilkinson, A. J., Paoli, M., and Dodson, G. G. (1998). The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annu. Rev. Biophys. Biomol. Struct.* 27:1–34.

Downloaded From Vet-e-Books.com





Properties of Enzymes

We have seen how the three-dimensional shapes of proteins allow them to serve structural and transport roles. We now discuss their functions as enzymes. Enzymes are extraordinarily efficient, selective, biological catalysts. Every living cell has hundreds of different enzymes catalyzing the reactions essential for life—even the simplest living organisms contain hundreds of different enzymes. In multicellular organisms, the complement of enzymes differentiates one cell type from another but most of the enzymes we discuss in this book are among the several hundred common to all cells. These enzymes catalyze the reactions of the central metabolic pathways necessary for the maintenance of life.

In the absence of the enzymes, metabolic reactions will not proceed at significant rates under physiological conditions. The primary role of enzymes is to enhance the rates of these reactions to make life possible. Enzyme-catalyzed reactions are 10^3 to 10^{20} times faster than the corresponding uncatalyzed reactions. A catalyst is defined as a substance that speeds up the attainment of equilibrium. It may be temporarily changed during the reaction but it is unchanged in the overall process since it recycles to participate in multiple reactions. Reactants bind to a catalyst and products dissociate from it. Note that a catalyst does not change the *position* of the reaction's equilibrium (i.e., it does not make an unfavorable reaction favorable). Instead, it lowers the amount of energy needed in order for the reaction to proceed. Catalysts speed up both the forward and reverse reactions by converting a one- or two-step process into several smaller steps each needing less energy than the uncatalyzed reaction.

Enzymes are highly specific for the reactants, or **substrates**, they act on, but the degree of substrate specificity varies. Some enzymes act on a group of related substrates, and others on only a single compound. Many enzymes exhibit **stereospecificity** meaning

I was awed by enzymes and fell instantly in love with them. I have since had love affairs with many enzymes (none as enduring as with DNA polymerase), but I have never met a dull or disappointing one.

—Arthur Kornberg (2001)

KEY CONCEPT

Catalysts speed up the rate of forward and reverse reactions but they don't change the equilibrium concentrations.

Top: The enzyme acetylcholinesterase with the reversible inhibitor donepezil hydrochloride (Aricept; shown in red) occupying the active site. Aricept is used to improve mental functioning in patients with Alzheimer's disease. It is thought to act by inhibiting the breakdown of the neurotransmitter acetylcholine in the brain, thus prolonging the neurotransmitter effects. (It does not, however, affect the course of the disease.) [PDB 1EVE]

Downloaded From Vet-eBooks.com



▲ **Enzyme reaction.** This is a large-scale enzyme reaction where milk is being curdled to make Appenzeller cheese. The reaction is catalyzed by rennet (rennin), which was originally derived from cow stomach. Rennet contains the enzyme chymosin, a protease that cleaves the milk protein casein between phenylalanine and methionine residues. The reaction releases a hydrophobic fragment of casein that aggregates and precipitates forming curd.

that they act on only a single stereoisomer of the substrate. Perhaps the most important aspect of enzyme specificity is **reaction specificity**—that is, the lack of formation of wasteful by-products. Reaction specificity is reflected in the exceptional purity of product (essentially 100%)—much higher than the purity of products of typical catalyzed reactions in organic chemistry. The specificity of enzymes not only saves energy for cells but also precludes the buildup of potentially toxic metabolic by-products.

Enzymes can do more than simply increase the rate of a single, highly specific reaction. Some can also combine, or couple, two reactions that would normally occur separately. This property allows the energy gained from one reaction to be used in a second reaction. Coupled reactions are a common feature of many enzymes—the hydrolysis of ATP, for example, is often coupled to less favorable metabolic reactions.

Some enzymatic reactions function as control points in metabolism. As we will see, metabolism is regulated in a variety of ways including alterations in the concentrations of enzymes, substrates, and enzyme inhibitors and modulation of the activity levels of certain enzymes. Enzymes whose activity is regulated generally have a more complex structure than unregulated enzymes. With few exceptions, regulated enzymes are oligomeric molecules that have separate binding sites for substrates and effectors, the compounds that act as regulatory signals. The fact that enzyme activity can be regulated is an important property that distinguishes biological catalysts from those encountered in a chemistry lab.

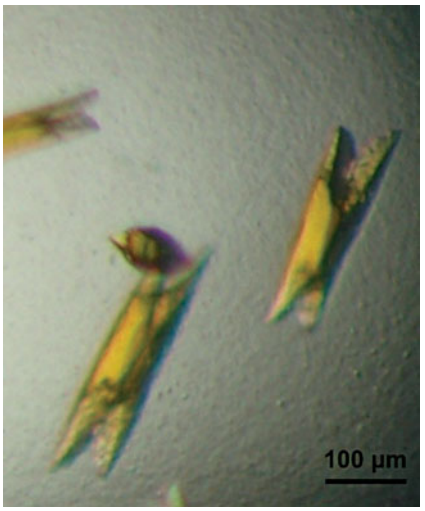
Some of the first biochemistry departments in universities were called Departments of Zymology.

The word *enzyme* is derived from a Greek word meaning "in yeast." It indicates that these catalysts are present inside cells. In the late 1800s, scientists studied the fermentation of sugars by yeast cells. Vitalists (who maintained that organic compounds could be made only by living cells) said that intact cells were needed for fermentation. Mechanists claimed that enzymes in yeast cells catalyze the reactions of fermentation. The latter conclusion was supported by the observation that cell-free extracts of yeast can catalyze fermentation. This finding was soon followed by the identification of individual reactions and the enzymes that catalyze them.

Catalytic RNA molecules are discussed in Chapters 21 and 22.

A generation later, in 1926, James B. Sumner crystallized the first enzyme (urease) and proved that it is a protein. Five more enzymes were purified in the next decade and also found to be proteins: pepsin, trypsin, chymotrypsin, carboxypeptidase, and Old Yellow Enzyme (a flavoprotein NADPH oxidase). Since then, almost all enzymes have been shown to be proteins or proteins plus cofactors. Certain RNA molecules also exhibit catalytic activity but they are not usually referred to as enzymes.

Downloaded From Vet-e-Books.com



▲ **Crystals of a bacterial (*Shewanella oneidensis*) homologue of Old Yellow Enzyme.** (Courtesy of J. Elegheert and S. N. Savvides)

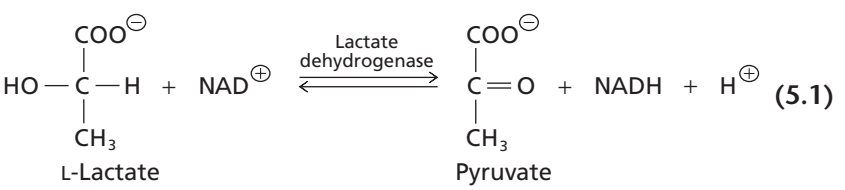
We begin this chapter with a description of enzyme classification and nomenclature. Next, we discuss kinetic analysis (measurements of reaction rates) emphasizing how kinetic experiments can reveal the properties of an enzyme and the nature of the complexes it forms with substrates and inhibitors. Finally, we describe the principles of inhibition and activation of regulatory enzymes. Chapter 6 explains how enzymes work at the chemical level and uses serine proteases to illustrate the relationship between protein structure and enzymatic function. Chapter 7 is devoted to the biochemistry of coenzymes, the organic molecules that assist some enzymes in their catalytic roles by providing reactive groups not found on amino acid side chains. In the remaining chapters we will present many other examples illustrating the four main properties of enzymes: (1) they function as catalysts, (2) they catalyze highly specific reactions, (3) they can couple reactions, and (4) their activity can be regulated.

5.1 The Six Classes of Enzymes

Most of the classical metabolic enzymes are named by adding the suffix *-ase* to the name of their substrates or to a descriptive term for the reactions they catalyze. For example, urease has urea as a substrate. Alcohol dehydrogenase catalyzes the removal of hydrogen from alcohols (i.e., the oxidation of alcohols). A few enzymes, such as trypsin and amylase, are known by their historic names. Many newly discovered enzymes are named after their genes or for some nondescriptive characteristic. For example, RecA is named after the *rec*A gene and HSP70 is a heat shock protein—both enzymes catalyze the hydrolysis of ATP.

A committee of the International Union of Biochemistry and Molecular Biology (IUBMB) maintains a classification scheme that categorizes enzymes according to the general class of organic chemical reaction that is catalyzed. The six categories—oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases—are defined below with an example of each type of enzyme. The IUBMB classification scheme assigns a unique number, called the enzyme classification number, or EC number, to each enzyme. IUBMB also assigns a unique systematic name to each enzyme; it may be different from the common name of an enzyme. This book usually refers to enzymes by their common names.

1. **Oxidoreductases** catalyze oxidation–reduction reactions. Most of these enzymes are commonly referred to as **dehydrogenases**. Other enzymes in this class are called oxidases, peroxidases, oxygenases, or reductases. There is a trend in biochemistry to refer to more and more of these enzymes by their systematic name, oxidoreductases, rather than the more common names in the older biochemical literature. One example of an oxidoreductase is lactate dehydrogenase (EC 1.1.1.27) also called lactate:NAD oxidoreductase. This enzyme catalyzes the reversible conversion of L-lactate to pyruvate. The oxidation of L-lactate is coupled to the reduction of the coenzyme nicotinamide adenine dinucleotide ($NAD^{\oplus}$).



(5.1)

2. **Transferases** catalyze group transfer reactions and many require the presence of coenzymes. In group transfer reactions a portion of the substrate molecule usually binds covalently to the enzyme or its coenzyme. This group includes kinases, enzymes that catalyze the transfer of a phosphoryl group from ATP. Alanine transaminase, whose systematic name is L-alanine:2-oxyglutarate aminotransferase

Downloaded From Vet-eBooks.com

BOX 5.1 ENZYME CLASSIFICATION NUMBERS

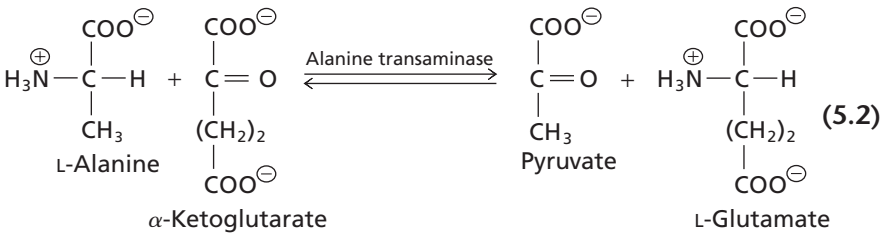
The enzyme classification number for malate dehydrogenase is EC 1.1.1.37. This enzyme has an activity similar to that of lactate dehydrogenase described under oxidoreductases (see Figure 4.23, Box 13.3).

The first number identifies this enzyme as a member of the first class of enzymes (oxidoreductases). The second number identifies the substrate group that malate dehydrogenase recognizes. Subclass 1.1 means that the substrate is a HC—OH group. The third number specifies the electron acceptor for this class of enzymes. Subclass 1.1.1 is for enzymes that use NAD^+ or $NADP^+$ as an acceptor. The final number means that malate dehydrogenase is the 37th enzyme in this category.

Compare the EC number of malate dehydrogenase with that of lactate dehydrogenase to see how similar enzymes have similar classification numbers.

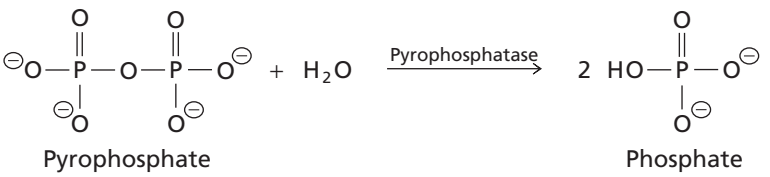
Accurate enzyme identification and classification is an important and essential part of modern biological databases. The entire classification database can be seen at **www.chem.qmul.ac.uk/iubmb/enzyme/.**

(EC 2.6.1.2), is a typical transferase. It transfers an amino group from L-alanine to α-ketoglutarate (2-oxoglutarate).



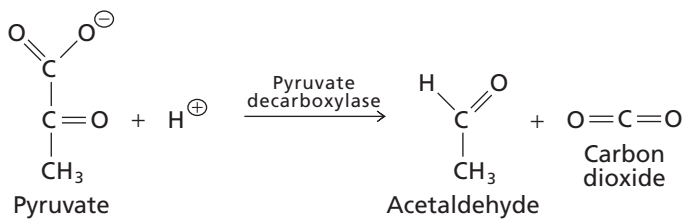
(5.2)

3. **Hydrolases** catalyze hydrolysis. They are a special class of transferases with water serving as the acceptor of the group transferred. Pyrophosphatase is a simple example of a hydrolase. The systematic name of this enzyme is diphosphate phosphohydrolase (EC 3.6.1.1).



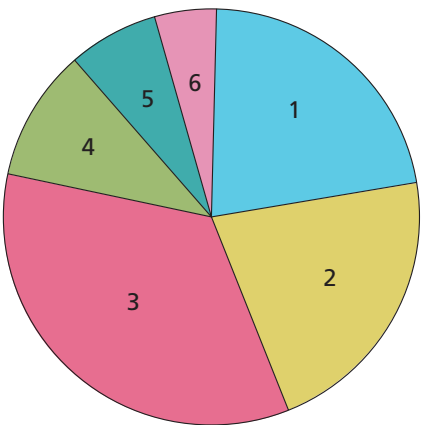
(5.3)

4. **Lyases** catalyze lysis of a substrate generating a double bond in nonhydrolytic, nonoxidative, elimination reactions. In the reverse direction, lyases catalyze the addition of one substrate to the double bond of a second substrate. Pyruvate decarboxylase belongs to this class of enzymes since it splits pyruvate into acetaldehyde and carbon dioxide. The systematic name for pyruvate decarboxylase, 2-oxo-acid carboxy-lyase (EC 4.1.1.1), is rarely used.



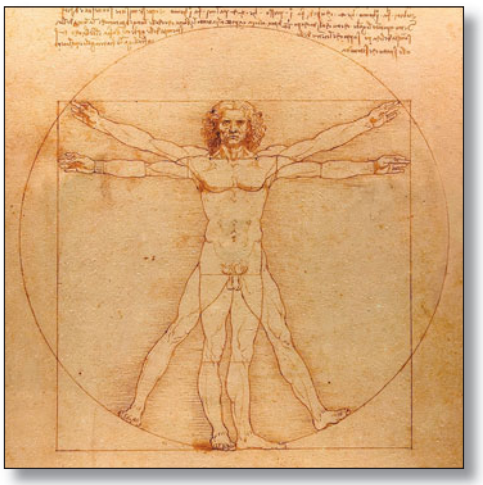
(5.4)

5. **Isomerases** catalyze structural change within a single molecule (isomerization reactions). Because these reactions have only one substrate and one product, they are among the simplest enzymatic reactions. Alanine racemase (EC 5.1.1.1) is an



▲ **Distribution of all known enzymes by EC classification number.** 1. oxidoreductases; 2. transferases; 3. hydrolases; 4. lyases; 5. isomerases; 6. ligases.

Downloaded From Vet-e-Books.com



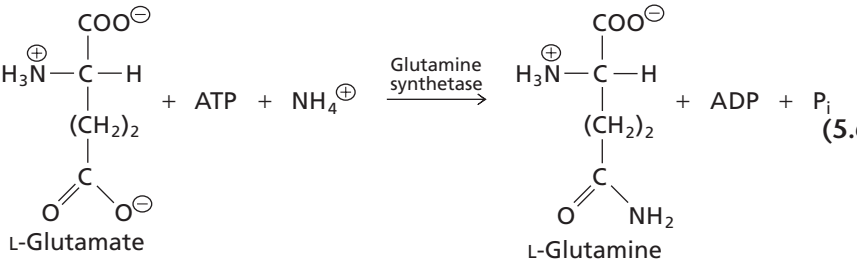
The human genome contains genes for about 1000 different enzymes catalyzing reactions in several hundred metabolic pathways (humancyc.org/). Since many enzymes have multiple subunits there are about 3000 different genes devoted to making enzymes. We have about 20,000 genes so most of the genes in our genome do not encode enzymes or enzyme subunits.

isomerase that catalyzes the interconversion of L-alanine and D-alanine. The common name is the same as the systematic name.



(5.5)

6. **Ligases** catalyze ligation, or joining, of two substrates. These reactions require the input of chemical potential energy in the form of a nucleoside triphosphate such as ATP. Ligases are usually referred to as **synthetases**. Glutamine synthetase, or L-glutamate:ammonia ligase (ADP-forming) (EC 6.3.1.2), uses the energy of ATP hydrolysis to join glutamate and ammonia to produce glutamine.



(5.6)

From the examples given above we see that most enzymes have more than one substrate although the second substrate may be only a molecule of water or a proton. Although enzymes catalyze both forward and reverse reactions, one-way arrows are often used when the equilibrium favors a great excess of product over substrate. Remember that when a reaction reaches equilibrium the enzyme must be catalyzing both the forward and reverse reactions at the same rate.

5.2 Kinetic Experiments Reveal Enzyme Properties

We begin our study of enzyme properties by examining the *rates* of enzyme-catalyzed reactions. Such studies fall under the category of enzyme kinetics (from the Greek *kinetikos*, "moving"). This is an appropriate place to begin since the most important property of enzymes is that they act as catalysts, speeding up the rates of reactions. Enzyme kinetics provides indirect information about the specificities and catalytic mechanisms of enzymes. Kinetic experiments also reveal whether an enzyme is regulated.

Recall that concentrations are indicated by square brackets: [P] signifies the concentration of product, [E] the concentration of enzyme, and [S] the concentration of the substrate.

Most enzyme research in the first half of the 20th century was limited to kinetic experiments. This research revealed how the rates of reactions are affected by variations in experimental conditions or changes in the concentration of enzyme or substrate. Before discussing enzyme kinetics in depth, let's review the principles of kinetics for nonenzymatic chemical systems. These principles are then applied to enzymatic reactions.

A. Chemical Kinetics

Kinetic experiments examine the relationship between the amount of product (P) formed in a unit of time ($\Delta[P]/\Delta t$) and the experimental conditions under which the reaction takes place. The basis of most kinetic measurements is the observation that the rate, or **velocity** (v), of a reaction varies directly with the concentration of each reactant (Section 1.4). This observation is expressed in a **rate equation**. For example, the rate equation for the nonenzymatic conversion of substrate (S) to product in an isomerization reaction is written as

$$\frac{\Delta[P]}{\Delta t} = v = k[S] \qquad (5.7)$$

Downloaded From Vet-eBooks.com

The rate equation reflects the fact that the velocity of a reaction depends on the concentration of the substrate ([S]). The symbol k is the rate constant and indicates the speed or efficiency of a reaction. Each reaction has a different rate constant. The units of the rate constant for a simple reaction are s^{-1}.

As a reaction proceeds, the amount of product ([P]) increases and the amount of substrate ([S]) decreases. An example of the progress of several reactions is shown in Figure 5.1a. The velocity is the slope of the progress curve over a particular interval of time. The shape of the curves indicates that the velocity is decreasing over time as expected since the substrate is being depleted.

In this hypothetical example, the velocity of the reaction might eventually become zero when the substrate is used up. This would explain why the curve flattens out at extended time points. (See below for another explanation.) We are interested in the relationship between substrate concentration and the velocity of a reaction since if we know these two values we can use Equation 5.7 to calculate the rate constant. The only accurate substrate concentration is the one we prepare at the beginning of the experiment because the concentration changes during the experiment. The velocity of the reaction at the very beginning is the value that we want to know. This value represents the rate of the reaction at a known substrate concentration before it changes.

The initial velocity (v_0) can be determined from the slope of the progress curves (Figure 5.1a) or from the derivatives of the curves. A graph of initial velocity versus substrate concentration at the beginning of the experiment gives a straight line as shown in Figure 5.1b. The slope of the curve in Figure 5.1b is the rate constant.

The experiment shown in Figure 5.1 will only determine the forward rate constant since the data were collected under conditions where there was no reverse reaction. This is another important reason for calculating initial velocity (v_0) rather than the rate at later time points. In a reversible reaction, the flattening of the progress curves does not represent zero velocity. Instead, it simply indicates that there is no net increase in product over time because the reaction has reached equilibrium.

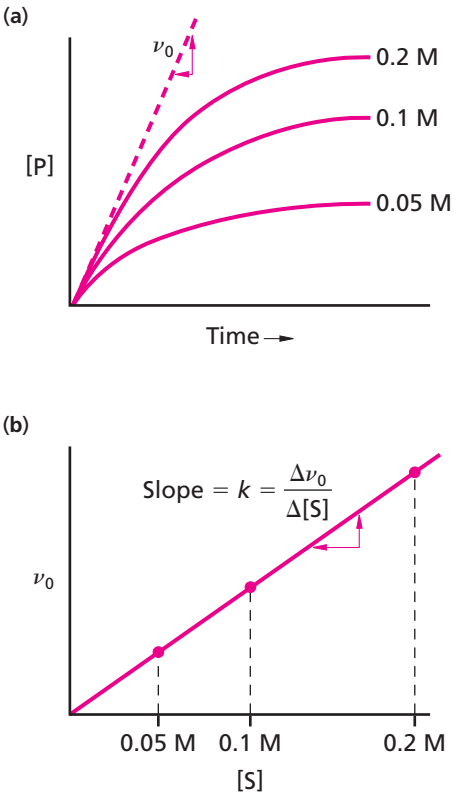
A better description of our simple reaction would be

$$S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P \qquad (5.8)$$

For a more complicated single-step reaction, such as the reaction $S_1 + S_2 \rightarrow P_1 + P_2$, the rate is determined by the concentrations of both substrates. If both substrates are present at similar concentrations, the rate equation is

$$v = k[S_1][S_2] \qquad (5.9)$$

The rate constant for reactions involving two substrates has the units $M^{-1}\ s^{-1}$. These rate constants can be easily determined by setting up conditions where the concentration of one substrate is very high and the other is varied. The rate of the reaction will depend on the concentration of the rate-limiting substrate.



▲ **Figure 5.1**
Rate of a simple chemical reaction. (a) The amount of product produced over time is plotted for several different initial substrate concentrations. The initial velocity v_0 is the slope of the progress curve at the beginning of the reaction. **(b)** The initial velocity as a function of initial substrate concentration. The slope of the curve is the rate constant.

KEY CONCEPT

The rate or velocity of a reaction depends on the concentration of substrate.

B. Enzyme Kinetics

One of the first great advances in biochemistry was the discovery that enzymes bind substrates transiently. In 1894, Emil Fischer proposed that an enzyme is a rigid template, or lock, and that the substrate is a matching key. Only specific substrates can fit into a given enzyme. Early studies of enzyme kinetics confirmed that an enzyme (E) binds a substrate to form an **enzyme–substrate complex (ES)**. ES complexes are formed when ligands bind noncovalently in their proper places in the active site. The substrate interacts transiently with the protein catalyst (and with other substrates in a multisubstrate reaction) on its way to forming the product of the reaction.

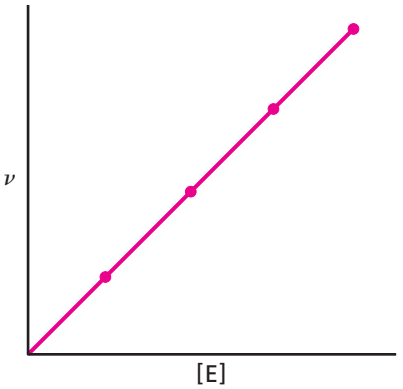
Let's consider a simple enzymatic reaction; namely, the conversion of a single substrate to a product. Although most enzymatic reactions have two or more substrates, the general principles of enzyme kinetics can be described by assuming the simple case of one substrate and one product.

$$E + S \longrightarrow ES \longrightarrow E + P \qquad (5.10)$$

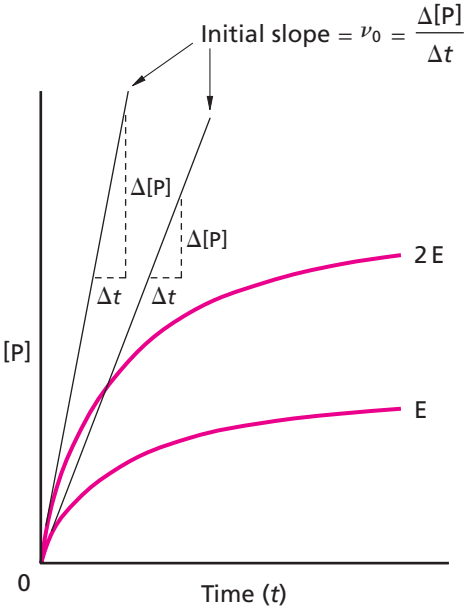
Downloaded From Vet-eBooks.com

KEY CONCEPT

The enzyme–substrate complex (ES) is a transient intermediate in an enzyme catalyzed reaction.



▲ **Figure 5.2**
Effect of enzyme concentration ([E]), on the initial velocity (v) of an enzyme-catalyzed reaction at a fixed, saturating [S]. The reaction rate is affected by the concentration of enzyme but not by the concentration of the other reactant, S.



▲ **Figure 5.3 Progress curve for an enzyme-catalyzed reaction.** [P], the concentration of product, increases as the reaction proceeds. The initial velocity of the reaction, v_0, is the slope of the initial linear portion of the curve. Note that the rate of the reaction doubles when twice as much enzyme (2E, upper curve) is added to an otherwise identical reaction mixture.

This reaction takes place in two distinct steps—the formation of the enzyme–substrate complex and the actual chemical reaction accompanied by the dissociation of the enzyme and product. Each step has a characteristic rate. The overall rate of an enzymatic reaction depends on the concentrations of both the substrate and the catalyst (enzyme). When the amount of enzyme is much less than the amount of substrate the reaction will depend on the amount of enzyme.

The straight line in Figure 5.2 illustrates the effect of enzyme concentration on the reaction velocity in a pseudo first-order reaction. The more enzyme present, the faster the reaction. These conditions are used in enzyme assays to determine the concentrations of enzymes. The concentration of enzyme in a test sample can be easily determined by comparing its activity to a reference curve similar to the model curve in Figure 5.2. Under these experimental conditions, there are sufficient numbers of substrate molecules so that every enzyme molecule binds a molecule of substrate to form an ES complex, a condition called saturation of E with S. Enzyme assays measure the amount of product formed in a given time period. In some assay methods, a recording spectrophotometer can be used to record data continuously; in other methods, samples are removed and analyzed at intervals. The assay is performed at a constant pH and temperature, generally chosen for optimal enzyme activity or for approximation to physiological conditions.

If we begin an enzyme-catalyzed reaction by mixing substrate and enzyme then there is no product present during the initial stages of the reaction. Under these conditions we can ignore the reverse reaction where P binds to E and is converted to S. The reaction can be described by

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \qquad \textbf{(5.11)}$$

The rate constants k_1 and k_{-1} in Reaction 5.11 govern the rates of association of S with E and dissociation of S from ES, respectively. This first step is an equilibrium binding interaction similar to the binding of oxygen to hemoglobin. The rate constant for the second step is k_2, the rate of formation of product from ES. Note that conversion of the ES complex to free enzyme and product is shown by a one-way arrow because the rate of the reverse reaction (E + P → EP) is negligible at the start of a reaction. The velocity measured during this short period is the **initial velocity** (v_0) described in the previous section. The formation and dissociation of ES complexes are usually very rapid reactions because only noncovalent bonds are being formed and broken. In contrast, the conversion of substrate to product is usually rate limiting. It is during this step that the substrate is chemically altered.

Enzyme kinetics differs from simple chemical kinetics because the rates of enzyme-catalyzed reactions depend on the concentration of enzyme and the enzyme is neither a product nor a substrate of the reaction. The rates also differ because substrate has to bind to enzyme before it can be converted to product. In an enzyme-catalyzed reaction, the initial velocities are obtained from progress curves, just as they are in chemical reactions. Figure 5.3 shows the progress curves at two different enzyme concentrations in the presence of a high initial concentration of substrate ([S] >> [E]). In this case, the rate of product formation depends on enzyme concentration and not on the substrate concentration. Data from experiments such as those shown in Figure 5.3 can be used to plot the curve shown in Figure 5.2.

5.3 The Michaelis–Menten Equation

Enzyme-catalyzed reactions, like any chemical reaction, can be described mathematically by rate equations. Several constants in the equations indicate the efficiency and specificity of an enzyme and are therefore useful for comparing the activities of several enzymes or for assessing the physiological importance of a given enzyme. The first rate equations were derived in the early 1900s by examining the effects of variations in substrate concentration. Figure 5.4 a shows a typical result where the initial velocity (v_0) of a reaction is plotted against the substrate concentration ([S]).

Downloaded From Vet-e-Books.com

The data can be explained by the reaction shown in Reaction 5.11. The first step is a bimolecular interaction between the enzyme and substrate to form an ES complex. At high substrate concentrations (right-hand side of the curve in Figure 5.4) the initial velocity doesn't change very much as more S is added. This indicates that the amount of enzyme has become rate-limiting in the reaction. The concentration of enzyme is an important component of the overall reaction as expected for formation of an ES complex. At low substrate concentrations (left-hand side of the curve in Figure 5.4), the initial velocity is very sensitive to changes in the substrate concentration. Under these conditions most enzyme molecules have not yet bound substrate and the formation of the ES complex depends on the substrate concentration.

The shape of the v_0 vs. [S] curve is that of a rectangular hyperbola. Hyperbolic curves indicate processes involving simple dissociation as we saw for the dissociation of oxygen from oxymyoglobin (Section 4.13B). This is further evidence that the simple reaction under study is bimolecular involving the association of E and S to form an ES complex. The equation for a rectangular hyperbola is

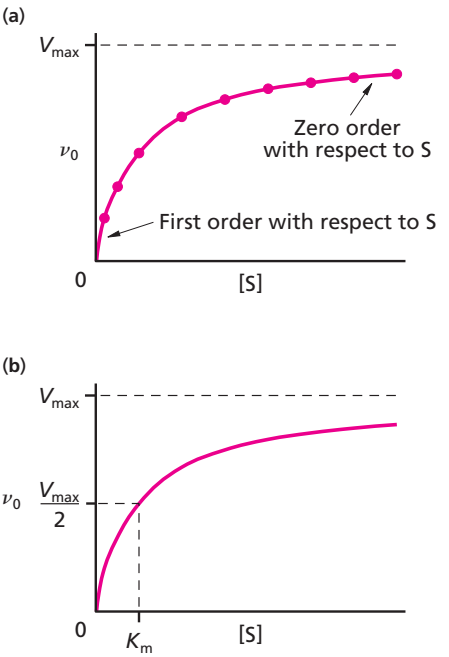
$$y = \frac{ax}{b + x} \tag{5.12}$$

where a is the asymptote of the curve (the value of y at an infinite value of x) and b is the point on the x axis corresponding to a value of $a/2$. In enzyme kinetic experiments, $y = v_0$ and $x = [S]$. The asymptote value (a) is called V_{max}. It's the maximum velocity of the reaction at infinitely large substrate concentrations. We often show the V_{max} value on v_0 vs. [S] plots but if you look at the figure it's not obvious why this particular asymptote was chosen. One of the characteristics of hyperbolic curves is that the curve seems to flatten out at moderate substrate concentrations at a level that seems far less than the V_{max} value. The true V_{max} is not determined by trying to estimate the position of the asymptote from the shape of the curve; instead, it is precisely and correctly determined by fitting the data to the general equation for a rectangular hyperbola.

The b term in the general equation for a rectangular hyperbola is called the **Michaelis constant** (K_m) defined as the concentration of substrate when v_0 is equal to one-half V_{max} (Figure 5.4b). The complete rate equation is

$$v_0 = \frac{V_{max}[S]}{K_m + [S]} \tag{5.13}$$

This is called the **Michaelis–Menten equation**, named after Leonor Michaelis and Maud Menten. Note how the general form of the equation compares to Equation 5.12. The Michaelis–Menten equation describes the relationship between the initial velocity of a reaction and the substrate concentration. In the following section we derive the Michaelis–Menten equation by a kinetic approach and then consider the meaning of the various constants.



▲ **Figure 5.4**
Plots of initial velocity (v_0) versus substrate concentration ([S]) for an enzyme-catalyzed reaction. (a) Each experimental point is obtained from a separate progress curve using the same concentration of enzyme. The shape of the curve is hyperbolic. At low substrate concentrations, the curve approximates a straight line that rises steeply. In this region of the curve, the reaction is highly dependent on the concentration of substrate. At high concentrations of substrate, the enzyme is almost saturated, and the initial rate of the reaction does not change much when substrate concentration is further increased. **(b)** The concentration of substrate that corresponds to half-maximum velocity is called the Michaelis constant (K_m). The enzyme is half-saturated when S = K_m.

A. Derivation of the Michaelis–Menten Equation

One common derivation of the Michaelis–Menten equation is termed the *steady state derivation*. It was proposed by George E. Briggs and J. B. S. Haldane. This derivation postulates a period of time (called the steady state) during which the ES complex is formed at the same rate that it decomposes so that the concentration of ES is constant. The initial velocity is used in the steady state derivation because we assume that the concentration of product ([P]) is negligible. The steady state is a common condition for metabolic reactions in cells.

If we assume a constant steady state concentration of ES then the rate of formation of product depends on the rate of the chemical reaction and the rate of dissociation of P from the enzyme. The rate limiting step is the right-hand side of Reaction 5.11 and the velocity depends on the rate constant k_2 and the concentration of ES.

$$ES \xrightarrow{k_2} E + P \qquad v_0 = k_2[ES] \tag{5.14}$$

Downloaded From Vet-e-Books.com



▲ Leonor Michaelis (1875–1949).

The steady-state derivation solves Equation 5.14 for [ES] using terms that can be measured such as the rate constant, the total enzyme concentration ($[E]_{total}$), and the substrate concentration ([S]). [S] is assumed to be greater than $[E]_{total}$ but not necessarily saturating. For example, soon after a small amount of enzyme is mixed with substrate [ES] becomes constant because the overall rate of decomposition of ES (the sum of the rates of conversion of ES to E + S and to E + P) is equal to the rate of formation of the ES complex from E + S. The rate of formation of ES from E + S depends on the concentration of free enzyme (enzyme molecules not in the form of ES) which is $[E]_{total} - [ES]$. The concentration of the ES complex remains constant until consumption of S causes [S] to approach $[E]_{total}$. We can express these statements as a mathematical equation.

$$\underset{\text{Rate of ES formation}}{k_1([E]_{total} - [ES])[S]} = \underset{\text{Rate of ES decomposition}}{(k_{-1} + k_2)[ES]} \quad \textbf{(5.15)}$$

Equation 5.15 is rearranged to collect the rate constants.

$$\frac{k_{-1} + k_2}{k_1} = K_m = \frac{1[E]_{total} - [ES]2[S]}{[ES]} \quad \textbf{(5.16)}$$

The ratio of rate constants on the left-hand side of Equation 5.16 is the Michaelis constant, K_m. Next, this equation is solved for [ES] in several steps.

$$[ES]K_m = ([E]_{total} - [ES])[S] \quad \textbf{(5.17)}$$

Expanding,

$$[ES]K_m = ([E]_{total}[S]) - ([ES][S]) \quad \textbf{(5.18)}$$

Collecting [ES] terms,

$$[ES](K_m + [S]) = [E]_{total}[S] \quad \textbf{(5.19)}$$

and

$$[ES] = \frac{[E]_{total}[S]}{K_m + [S]} \quad \textbf{(5.20)}$$



▼ Maud Menten (1879–1960).



Downloaded From Vet-eBooks.com

Equation 5.20 describes the steady-state ES concentration using terms that can be measured in an experiment. Substituting the value of [ES] into the velocity equation (Equation 5.14) gives

$$v_0 = k_2[\text{ES}] = \frac{k_2[\text{E}]_{\text{total}}[\text{S}]}{K_m + [\text{S}]} \tag{5.21}$$

As indicated by Figure 5.4a, when the concentration of S is very high the enzyme is saturated and essentially all the molecules of E are present as ES. Adding more S has almost no effect on the reaction velocity. The only way to increase the velocity is to add more enzyme. Under these conditions the velocity is at its maximum rate (V_{max}) and this velocity is determined by the total enzyme concentration and the rate constant k_2. Thus, by definition,

$$V_{max} = k_2[\text{E}]_{\text{total}} \tag{5.22}$$

Substituting this in Equation 5.21 gives the most familiar form of the Michaelis–Menten equation.

$$v_0 = \frac{V_{max}[\text{S}]}{K_m + [\text{S}]} \tag{5.23}$$

We've already seen that this form of the Michaelis–Menten equation adequately describes the data from kinetic experiments. In this section we've shown that the same equation can be derived from a theoretical consideration of the implications of Reaction 5.11, the equation for an enzyme-catalyzed reaction. The agreement between theory and data gives us confidence that the theoretical basis of enzyme kinetics is sound.

KEY CONCEPT

The constant k_{cat} is the number of moles of substrate converted to product per second per mole of enzyme.

B. The Catalytic Constant k_{cat}

At high substrate concentration, the overall velocity of the reaction is V_{max} and the rate is determined by the enzyme concentration. The rate constant observed under these conditions is called the **catalytic constant, k_{cat}**, defined as

$$V_{max} = k_{cat}[\text{E}]_{\text{total}} \qquad k_{cat} = \frac{V_{max}}{[\text{E}]_{\text{total}}} \tag{5.24}$$

where k_{cat} represents the number of moles of substrate converted to product per second per mole of enzyme (or per mole of active site for a multisubunit enzyme) under saturating conditions. In other words, k_{cat} indicates the maximum number of substrate molecules converted to product each second by each active site. This is often called the **turnover number**. The catalytic constant measures how quickly a given enzyme can catalyze a specific reaction—it's a very useful way of describing the effectiveness of an enzyme. The unit for k_{cat} is s^{-1} and the reciprocal of k_{cat} is the time required for one catalytic event. Note that the enzyme concentration must be known in order to calculate k_{cat}.

For a simple reaction, such as Reaction 5.11, the rate-limiting step is the conversion of substrate to product and the dissociation of product from the enzyme (ES → E + P). Under these conditions k_{cat} is equal to k_2 (Equation 5.14). Many enzyme reactions are more complex. If one step is clearly rate-limiting then its rate constant is the k_{cat} for that reaction. If the mechanism is more complex then k_{cat} may be a combination of several different rate constants. This is why we need a different rate constant (k_{cat}) to describe the overall rate of the enzyme-catalyzed reaction. In most cases you can assume that k_{cat} is a good approximation of k_2.

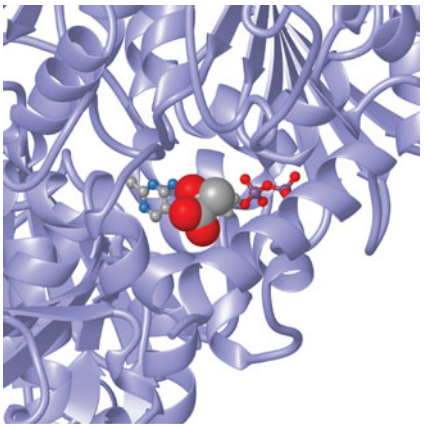
Representative values of k_{cat} are listed in Table 5.1. Most enzymes are potent catalysts with k_{cat} values of 10^2 to 10^3 s^{-1}. This means that at high substrate concentrations a single

Table 5.1 Examples of catalytic constants

Enzyme	$k_{cat}(s^{-1})$*
Papain	10
Ribonuclease	10^2
Carboxypeptidase	10^2
Trypsin	10^2 (to 10^3)
Acetylcholinesterase	10^3
Kinases	10^3
Dehydrogenases	10^3
Transaminases	10^3
Carbonic anhydrase	10^6
Superoxide dismutase	10^6
Catalase	10^7

*The catalytic constants are given only as orders of magnitude.

Downloaded From Vet-e-Books.com



▲ **Substrate binding.** Pyruvate carboxylase binds pyruvate, HCO_3^-, and ATP. The structure of the active site of the yeast (*Saccharomyces cerevisiae*) enzyme is shown here with a bound molecule of pyruvate (space-filling representation) and the cofactor biotin (ball-and-stick). The K_m value for pyruvate binding is 4×10^{-4} M. The K_m values for HCO_3^-, and ATP binding are 1×10^{-3} M and 6×10^{-5} M. [PDB 2VK1]

enzyme molecule will convert 100–1000 molecules of substrate to product every second. This rate is limited by a number of factors that will be discussed in the next chapter (Chapter 6: *Mechanisms of Enzymes*).

Some enzymes are extremely rapid catalysts with k_{cat} values of 10^6 s^{-1} or greater. Mammalian carbonic anhydrase, for example, must act very rapidly in order to maintain equilibrium between aqueous CO_2 and bicarbonate (Section 2.10). As we will see in Section 6.4B, superoxide dismutase and catalase are responsible for rapid decomposition of the toxic oxygen metabolites superoxide anion and hydrogen peroxide, respectively. Enzymes that catalyze a million reactions per second often act on small substrate molecules that diffuse rapidly inside the cell.

C. The Meanings of K_m

The Michaelis constant has a number of meanings. Equation 5.16 defined K_m as the ratio of the combined rate constants for the breakdown of ES divided by the constant for its formation. If the rate constant for product formation (k_2) is much smaller than either k_1 or k_{-1}, as is often the case, k_2 can be neglected and K_m is equivalent to k_{-1}/k_1. In this case K_m is the same as the equilibrium constant for dissociation of the ES complex to E +S. Thus, K_m becomes a measure of the affinity of E for S. The lower the value of K_m, the more tightly the substrate is bound. K_m is also one of the parameters that determines the shape of the v_0 vs. [S] curve shown in Figure 5.4b. It is the substrate concentration when the initial velocity is one-half the V_{max} value. This meaning follows directly from the general equation for a rectangular hyperbola.

K_m values are sometimes used to distinguish between different enzymes that catalyze the same reaction. For example, mammals have several different forms of lactate dehydrogenase, each with a distinct K_m value. Although it is useful to think of K_m as representing the equilibrium dissociation constant for ES, this is not always valid. For many enzymes K_m is a more complex function of the rate constants. This is especially true when the reaction occurs in more than two steps.

Typical K_m values for enzymes range from 10^{-2} to 10^{-5} M. Since these values often represent apparent dissociation constants their reciprocal is an apparent association (binding) constant. You can see by comparison with protein–protein interactions (Section 4.9) that the binding of enzymes to substrates is much weaker.

5.4 Kinetic Constants Indicate Enzyme Activity and Catalytic Proficiency

KEY CONCEPT

K_m is the substrate concentration when the rate of the reaction is one-half the V_{max} value. It is often an approximation of the equilibrium dissociation constant of the reaction ES $\rightleftharpoons$ E + S.

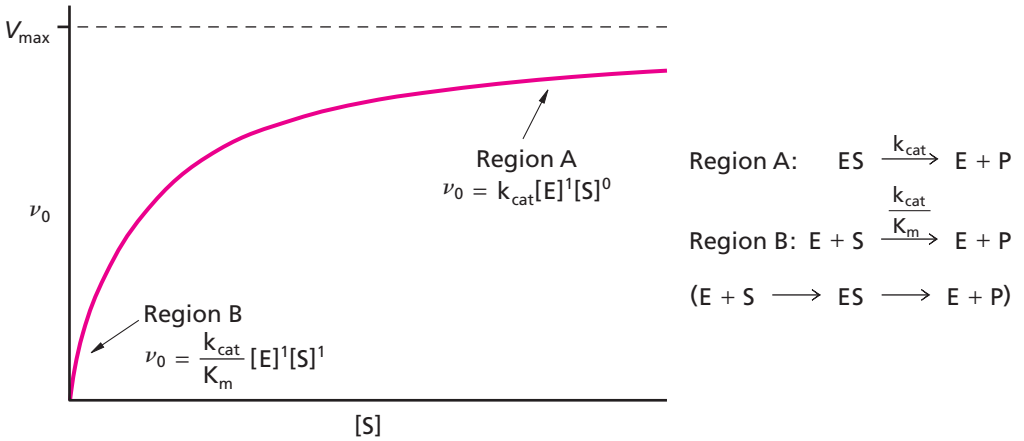
We've seen that the kinetic constants K_m and k_{cat} can be used to gauge the relative activities of enzymes and substrates. In most cases, K_m is a measure of the stability of the ES complex and k_{cat} is similar to the rate constant for the conversion of ES to E + P when the substrate is not limiting (region A in Figure 5.5). Recall that k_{cat} is a measure of the catalytic activity of an enzyme indicating how many reactions a molecule of enzyme can catalyze per second.

Examine region B of the hyperbolic curve in Figure 5.5. The concentration of S is very low and the curve approximates a straight line. Under these conditions, the reaction rate depends on the concentrations of both substrate and enzyme. In chemical terms, this is a second-order reaction and the velocity depends on a second-order rate constant defined by

$$v_0 = k[E][S] \qquad \textbf{(5.25)}$$

We are interested in knowing how to determine this second-order rate constant since it tells us the rate of the enzyme-catalyzed reaction under physiological conditions. When Michaelis and Menten first wrote the full rate equation they used the form that included $k_{cat}[E]_{total}$ rather than V_{max} (Equation 5.24). Now that we understand the meaning of k_{cat}

Downloaded From Vet-e-Books.com



▲ **Figure 5.5 Meanings of k_{cat} and k_{cat}/K_m.** The catalytic constant (k_{cat}) is the rate constant for conversion of the ES complex to E + P. It is measured most easily when the enzyme is saturated with substrate (region A on the Michaelis–Menten curve shown). The ratio k_{cat}/K_m is the rate constant for the conversion of E + S to E + P at very low concentrations of substrate (region B). The reactions measured by these rate constants are summarized below the graph.

we can substitute $k_{cat}[E]_{total}$ in the Michaelis–Menten equation (Equation 5.23) in place of V_{max}. If we consider only the region of the Michaelis–Menten curve at a very low [S] then this equation can be simplified by neglecting the [S] in the denominator since [S] is much less than K_m.

$$v_0 = \frac{k_{cat}[E][S]}{K_m + [S]} = \frac{k_{cat}}{K_m}[E][S] \qquad \textbf{(5.26)}$$

Comparing Equations 5.25 and 5.26 reveals that the second-order rate constant is closely approximated by k_{cat}/K_m. Thus, the ratio k_{cat}/K_m is an apparent second-order rate constant for the formation of E + P from E + S when the overall reaction is limited by the encounter of S with E. This ratio approaches 10^8 to 10^9 M^{-1} s^{-1}, the fastest rate at which two uncharged solutes can approach each other by diffusion at physiological temperature. Enzymes that can catalyze reactions at this extremely rapid rate are discussed in Section 6.4.

The k_{cat}/K_m ratio is useful for comparing the activities of different enzymes. It is also possible to assess the efficiency of an enzyme by measuring its catalytic proficiency. This value is equal to the rate constants for a reaction in the presence of the enzyme (k_{cat}/K_m) divided by the rate constant for the same reaction in the absence of the enzyme (k_n). Surprisingly few catalytic proficiency values are known because most chemical reactions occur extremely slowly in the absence of enzymes—so slowly that their nonenzymatic rates are very difficult to measure. The reaction rates are often measured in special steel-enclosed glass vessels at temperatures in excess of 300°C.

Table 5.2 lists several examples of known catalytic proficiencies. Typical values range from 10^{14} to 10^{20} but some are quite a bit higher (up to 10^{24}). The current record holder is uroporphyrinogen decarboxylase, an enzyme required for a step in the porphyrin synthesis pathway. The difficulty in obtaining rate constants for nonenzymatic reactions is illustrated by the half-life for the uncatalyzed reaction—about 2 billion years! The catalytic proficiency values in Table 5.2 emphasize one of the main properties of enzymes, namely, their ability to increase the rates of reactions that would normally occur too slowly to be useful.

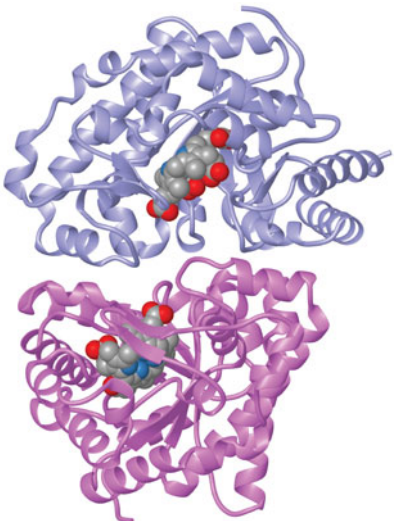
5.5 Measurement of K_m and V_{max}

The kinetic parameters of an enzymatic reaction can provide valuable information about the specificity and mechanism of the reaction. The key parameters are K_m and V_{max} because k_{cat} can be calculated if V_{max} is known.

Downloaded From Vet-e-Books.com

Table 5.2 Catalytic proficiencies of some enzymes

	Nonenzymatic rate constant (k_n in s^{-1})	Enzymatic rate constant (k_{cat}/K_m in $M^{-1}s^{-1}$)	Catalytic proficiency
Carbonic anhydrase	10^{-1}	7×10^6	7×10^7
Chymotrypsin	4×10^{-9}	9×10^7	2×10^{16}
Chorismate mutase	10^{-5}	2×10^6	2×10^{11}
Triose phosphate isomerase	4×10^{-6}	4×10^8	10^{14}
Cytidine deaminase	10^{-10}	3×10^6	3×10^{16}
Adenosine deaminase	2×10^{-10}	10^7	5×10^{16}
Mandelate racemase	3×10^{-13}	10^6	3×10^{18}
β-Amylase	7×10^{-14}	10^7	10^{20}
Fumarase	10^{-13}	10^9	10^{21}
Arginine decarboxylase	9×10^{-16}	10^6	10^{21}
Alkaline phosphatase	10^{-15}	3×10^7	3×10^{22}
Orotidine 5′-phosphate decarboxylase	3×10^{-16}	6×10^7	2×10^{23}
Uroporphyrinogen decarboxylase	10^{-17}	2×10^7	2×10^{24}



▲ **Maximum catalytic proficiency.** Uroporphyrinogen decarboxylase is the current record holder for maximum catalytic proficiency. It catalyzes a step in the heme synthesis pathway. The enzyme shown here is a human *(Homo sapiens)* variant with a bound porphoryrin molecule at the active site of each monomer. [PDB 2Q71]

K_m and V_{max} for an enzyme-catalyzed reaction can be determined in several ways. Both values can be obtained by the analysis of initial velocities at a series of substrate concentrations and a fixed concentration of enzyme. In order to obtain reliable values for the kinetic constants the [S] points must be spread out both below and above K_m to produce a hyperbola. It is difficult to determine either K_m or V_{max} directly from a graph

BOX 5.2 HYPERBOLAS VERSUS STRAIGHT LINES

We have seen that a plot of substrate concentration ([S]) versus the initial velocity of a reaction (v_0) produces a hyperbolic curve as shown in Figures 5.4 and 5.5. The general equation for a rectangular hyperbola (Equation 5.12) and the Michaelis–Menten equation have the same form (Equation 5.13).

It's very difficult to determine V_{max} from a plot of enzyme kinetic data since the hyperbolic curve that shows the relationship between substrate concentration and initial velocity is asymptotic to V_{max} and it is experimentally difficult to achieve the concentration of substrate required to estimate V_{max}. For these reasons, it is often easier to convert the hyperbolic curve to a linear form that matches the general formula $y = mx + b$, where m is the slope of the line and b is the y-axis intercept. The first step in transforming the original Michaelis–Menten equation to this general form of a linear equation is to invert the terms so that the K_m + [S] term is on top of the right-hand side. This is done by taking the reciprocal of each side—a transformation that will be familiar to many who are familiar with hyperbolic curves.

The next two steps involve separating terms and canceling [S] in the second term on the right-hand side of the equation. This form of the Michaelis–Menten equation is called the Lineweaver–Burk equation and it resembles the general form of a linear equation, $y = mx + b$, where y is the reciprocal of v_0 and x values are the reciprocal of [S]. A plot of data in this form is referred to as a double-reciprocal plot. The slope of the line will be K_m/V_{max} and the y-axis intercept will be $1/V_{max}$.

The original reason for this sort of transformation was to calculate K_m and V_{max} from experimental data. It was easier to plot the reciprocal values of v_0 and [S] and draw a straight line through the points in order to calculate the kinetic constants. Nowadays, there are computer programs that can accurately fit the data to a hyperbolic curve and calculate the constants so the Lineweaver–Burk plot is no longer necessary for this type of analysis. In this book we will still use the Lineweaver–Burk plots to illustrate some general features of enzyme kinetics but they are rarely used for their original purpose of data analysis.

$$\frac{1}{v_0} = \frac{K_m + [S]}{V_{max}[S]} \quad \frac{1}{v_0} = \frac{K_m}{V_{max}[S]} + \frac{[S]}{V_{max}[S]} \quad \frac{1}{v_0} = \left(\frac{K_m}{V_{max}}\right)\frac{1}{[S]} + \frac{1}{V_{max}}$$

Downloaded From Vet-eBooks.com

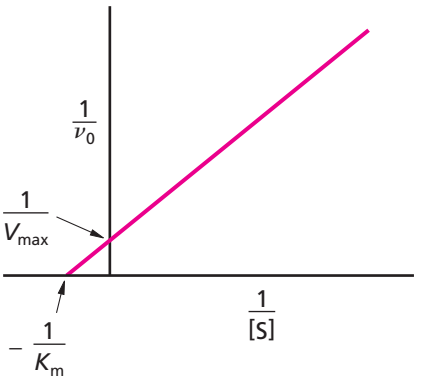
of initial velocity versus concentration because the curve approaches V_{max} asymptotically. However, accurate values can be determined by using a suitable computer program to fit the experimental results to the equation for the hyperbola.

The Michaelis–Menten equation can be rewritten in order to obtain values for V_{max} and K_m from straight lines on graphs. The most commonly used transformation is the double-reciprocal, or Lineweaver–Burk, plot in which the values of $1/v_0$ are plotted against 1/[S] (Figure 5.6). The absolute value of $1/K_m$ is obtained from the intercept of the line at the x axis, and the value of $1/V_{max}$ is obtained from the y intercept. Although double-reciprocal plots are not the most accurate methods for determining kinetic constants, they are easily understood and provide recognizable patterns for the study of enzyme inhibition, an extremely important aspect of enzymology that we will examine shortly.

Values of k_{cat} can be obtained from measurements of V_{max} only when the absolute concentration of the enzyme is known. Values of K_m can be determined even when enzymes have not been purified provided that only one enzyme in the impure preparation can catalyze the observed reaction.

Lineweaver–Burk equation:

$$\frac{1}{v_0} = \left(\frac{K_m}{V_{max}}\right)\frac{1}{[S]} + \frac{1}{V_{max}}$$



▲ **Figure 5.6**
Double-reciprocal (Lineweaver–Burk) plot. This plot is derived from a linear transformation of the Michaelis–Menten equation. Values of $1/v_0$ are plotted as a function of 1/[S] values.

5.6 Kinetics of Multisubstrate Reactions

Until now, we have only been considering reactions where a single substrate is converted to a single product. Let's consider a reaction in which two substrates, A and B, are converted to products P and Q.

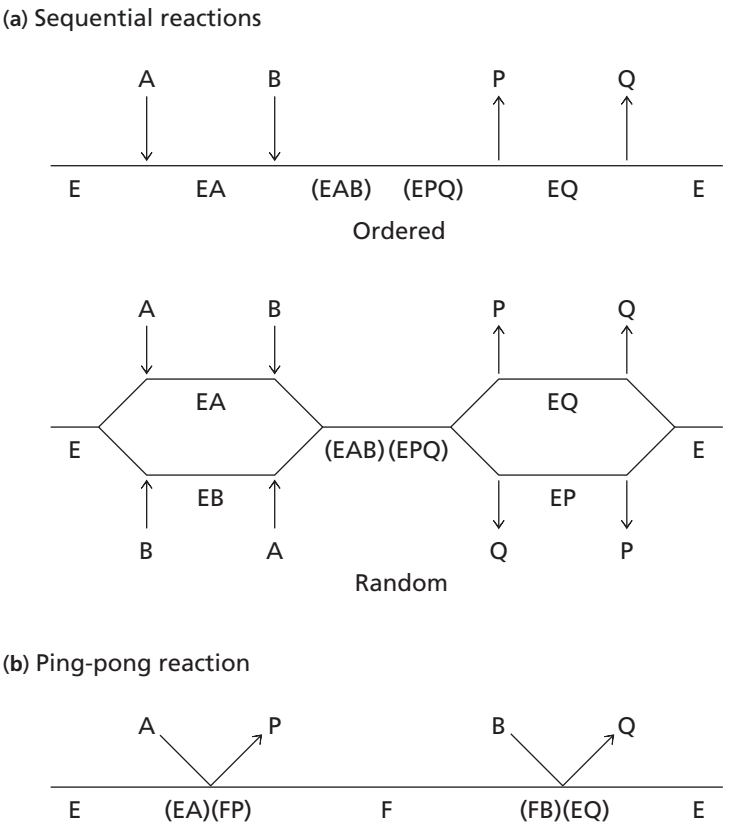
$$E + A + B \rightleftharpoons (EAB) \rightarrow E + P + Q \qquad \textbf{(5.27)}$$

Kinetic measurements for such multisubstrate reactions are a little more complicated than simple one-substrate enzyme kinetics. For many purposes, such as designing an enzyme assay, it's sufficient simply to determine the K_m for each substrate in the presence of saturating amounts of each of the other substrates as we described for chemical reactions (Section 5.2A). The simple enzyme kinetics discussed in this chapter can be extended to distinguish among several mechanistic possibilities for multisubstrate reactions, such as group transfer reactions. This is done by measuring the effect of variations in the concentration of one substrate on the kinetic results obtained for the other.

Multisubstrate reactions can occur by several different kinetic schemes. These schemes are called **kinetic mechanisms** because they are derived entirely from kinetic experiments. Kinetic mechanisms are commonly represented using the notation introduced by W. W. Cleland. The sequence of steps proceeds from left to right (Figure 5.7). The addition of substrate molecules (A, B, C, . . .) to the enzyme and the release of products (P, Q, R, . . .) from the enzyme are indicated by arrows pointing toward (substrate binding) or from (product release) the line. The various forms of the enzyme (free E, ES complexes, or EP complexes) are written under a horizontal line. The ES complexes that undergo chemical transformation when the active site is filled are shown in parentheses.

Sequential reactions (Figure 5.7a) require all the substrates to be present before any product is released. Sequential reactions can be either **ordered**, with an obligatory order for the addition of substrates and release of products, or **random**. In **ping-pong reactions** (Figure 5.7b), a product is released before all the substrates are bound. In a bisubstrate ping-pong reaction, the first substrate is bound, the enzyme is altered by substitution, and the first product is released. Then the second substrate is bound, the altered enzyme is restored to its original form, and the second product is released. A ping-pong mechanism is sometimes called a substituted-enzyme mechanism because of the covalent binding of a portion of a substrate to the enzyme. The binding and release of ligands in a ping-pong mechanism are usually indicated by slanted lines. The two forms of the enzyme are represented by E (unsubstituted) and F (substituted).

Downloaded From Vet-e-Books.com



▲ **Figure 5.7**
Notation for bisubstrate reactions. (a) In sequential reactions, all substrates are bound before a product is released. The binding of substrates may be either ordered or random. **(b)** In ping-pong reactions, one substrate is bound and a product is released, leaving a substituted enzyme. A second substrate is then bound and a second product released, restoring the enzyme to its original form.

5.7 Reversible Enzyme Inhibition

An enzyme **inhibitor** (I) is a compound that binds to an enzyme and interferes with its activity. Inhibitors can act by preventing the formation of the ES complex or by blocking the chemical reaction that leads to the formation of product. As a general rule, inhibitors are small molecules that bind reversibly to the enzyme they inhibit. Cells contain many natural enzyme inhibitors that play important roles in regulating metabolism. Artificial inhibitors are used experimentally to investigate enzyme mechanisms and decipher metabolic pathways. Some drugs, and many poisons, are enzyme inhibitors.

Some inhibitors bind covalently to enzymes causing irreversible inhibition but most biologically relevant inhibition is reversible. Reversible inhibitors are bound to enzymes by the same weak, noncovalent forces that bind substrates and products. The equilibrium between free enzyme (E) plus inhibitor (I) and the EI complex is characterized by a dissociation constant. In this case, the constant is called the **inhibition constant**, K_i.

Irreversible inhibitors are described in Section 5.8.

$$E + I \rightleftharpoons EI \qquad K_d = K_i = \frac{[E][I]}{[EI]} \tag{5.28}$$

KEY CONCEPT

Reversible inhibitors bind to enzymes and either prevent substrate binding or block the reaction leading to formation of product.

The basic types of reversible inhibition are competitive, uncompetitive, noncompetitive and mixed. These can be distinguished experimentally by their effects on the kinetic behavior of enzymes (Table 5.3). Figure 5.8 shows diagrams representing modes of reversible enzyme inhibition.

Downloaded From Vet-e-Books.com

Table 5.3 Effects of reversible inhibitors on kinetic constants

Type of inhibitor	Effect
Competitive (I binds to E only)	Raises K_m V_{max} remains unchanged
Uncompetitive (I binds to ES only)	Lowers V_{max} and K_m Ratio of V_{max}/K_m remains unchanged
Noncompetitive (I binds to E or ES)	Lowers V_{max} K_m remains unchanged

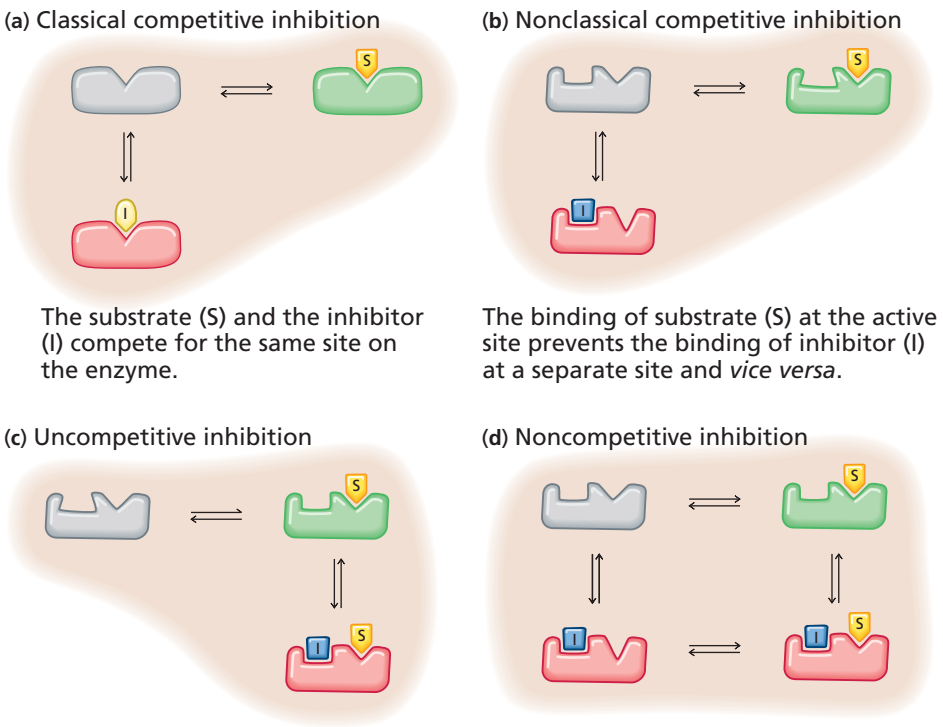
A. Competitive Inhibition

Competitive inhibitors are the most commonly encountered inhibitors in biochemistry. In competitive inhibition, the inhibitor can bind only to free enzyme molecules that have not bound any substrate. **Competitive inhibition** is illustrated in Figure 5.8 and by the kinetic scheme in Figure 5.9a. In this scheme only ES can lead to the formation of product. The formation of an EI complex removes enzyme from the normal pathway.

Once a competitive inhibitor is bound to an enzyme molecule, a substrate molecule cannot bind to that enzyme molecule. Conversely, the binding of substrate to an enzyme molecule prevents the binding of an inhibitor. In other words, S and I compete for binding to the enzyme molecule. Most commonly, S and I bind at the same site on the enzyme, the active site. This type of inhibition is termed classical competitive inhibition (Figure 5.8). This is not the only kind of competitive inhibition (see Figure 5.8). In some cases, such as allosteric enzymes (Section 5.10), the inhibitor binds at a different site and this alters the substrate binding site preventing substrate binding. This type of inhibition is called nonclassical competitive inhibition. When both I and S are

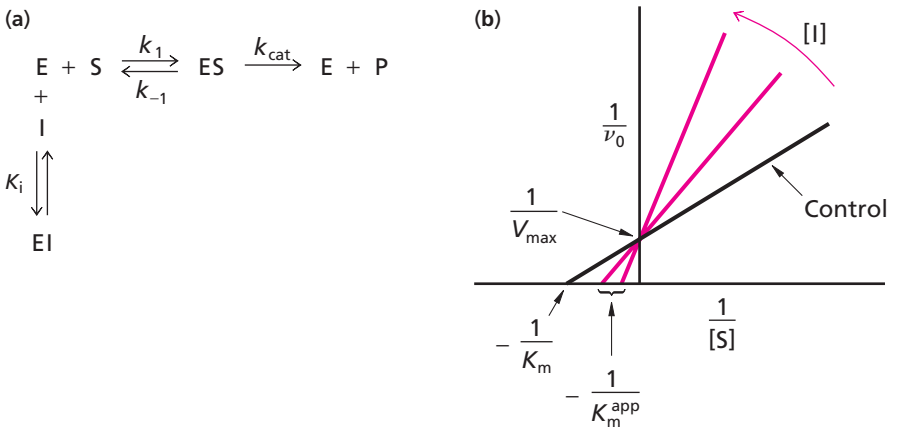


▲ **Competitive inhibition.** The active ingredient in the weed killer Roundup© is glyphosate, a competitive inhibitor of the plant enzyme 5-enolpyruvylshikimate-3-phosphate synthase. (See Box 17.2 in Chapter 17.)

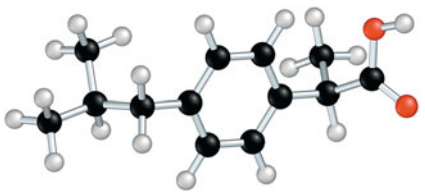


◄ **Figure 5.8**
Diagrams of reversible enzyme inhibition. In this scheme, catalytically competent enzymes are green and inactive enzymes are red.

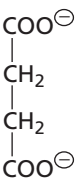
Downloaded From Vet-e-Books.com



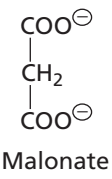
▲ **Figure 5.9**
Competitive inhibition. (a) Kinetic scheme illustrating the binding of I to E. Note that this is an expansion of Equation 5.11 that includes formation of the EI complex. **(b)** Double-reciprocal plot. In competitive inhibition, V_{max} remains unchanged and K_m increases. The black line labeled "Control" is the result in the absence of inhibitor. The red lines are the results in the presence of inhibitor, with the arrow showing the direction of increasing [I].



▲ Ibuprofen, the active ingredient in many over-the-counter painkillers, is a competitive inhibitor of the enzyme cyclooxygenase (COX). (See Box 16.1 Chapter 16.)



Succinate



Malonate

present in a solution, the proportion of the enzyme that is able to form ES complexes depends on the concentrations of substrate and inhibitor and their relative affinities for the enzyme.

The amount of EI can be reduced by increasing the concentration of S. At sufficiently high concentrations the enzyme can still be saturated with substrate. Therefore, the maximum velocity is the same in the presence or in the absence of an inhibitor. The more competitive inhibitor present, the more substrate needed for half-saturation. We have shown that the concentration of substrate at half-saturation is K_m. In the presence of increasing concentrations of a competitive inhibitor, K_m increases. The new value is usually referred to as the apparent K_m (K_m^{app}). On a double-reciprocal plot, adding a competitive inhibitor shows as a decrease in the absolute value of the intercept at the x axis $1/K_m$, whereas the y intercept $1/V_{max}$ remains the same (Figure 5.9b).

Many classical competitive inhibitors are substrate analogs—compounds that are structurally similar to substrates. The analogs bind to the enzyme but do not react. For example, the enzyme succinate dehydrogenase converts succinate to fumarate (Section 13.3#6). Malonate resembles succinate and acts as a competitive inhibitor of the enzyme.

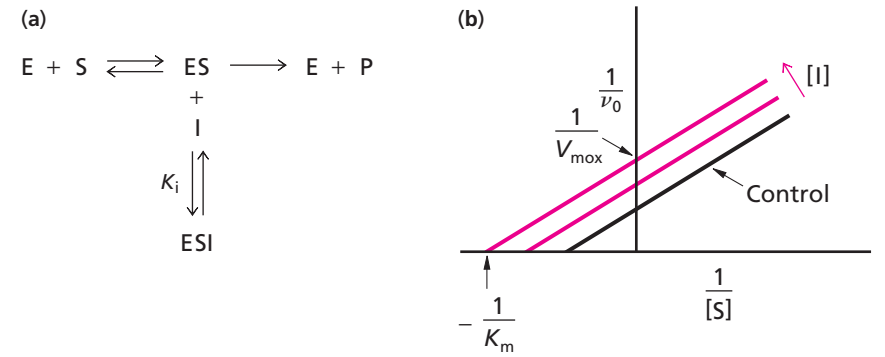
B. Uncompetitive Inhibition

Uncompetitive inhibitors bind only to ES and not to free enzyme (Figure 5.10a). In **uncompetitive inhibition**, V_{max} is decreased ($1/V_{max}$ is increased) by the conversion of some molecules of E to the inactive form ESI. Since it is the ES complex that binds I, the decrease in V_{max} is not reversed by the addition of more substrate. Uncompetitive inhibitors also decrease the K_m (seen as an increase in the absolute value of $1/K_m$ on a double-reciprocal plot) because the equilibria for the formation of both ES and ESI are shifted toward the complexes by the binding of I. Experimentally, the lines on a double-reciprocal plot representing varying concentrations of an uncompetitive inhibitor all have the same slope indicating proportionally decreased values for K_m and V_{max} (Figure 5.10b). This type of inhibition usually occurs only with multisubstrate reactions.

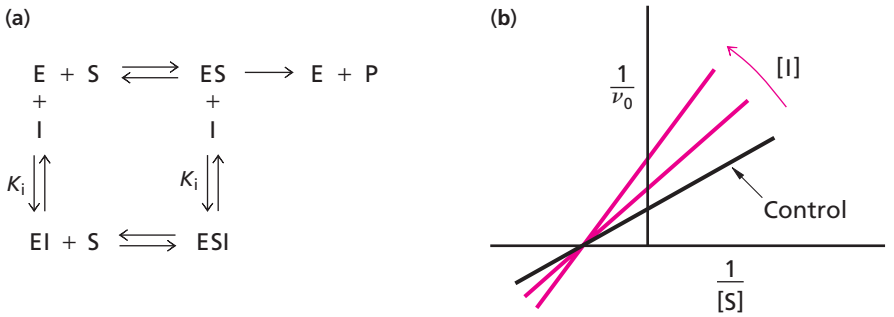
C. Noncompetitive Inhibition

Noncompetitive inhibitors can bind to E or ES forming inactive EI or ESI complexes, respectively (Figure 5.11a). These inhibitors are not substrate analogs and do not bind at the same site as S. The classic case of noncompetitive inhibition is characterized by an

Downloaded From Vet-eBooks.com



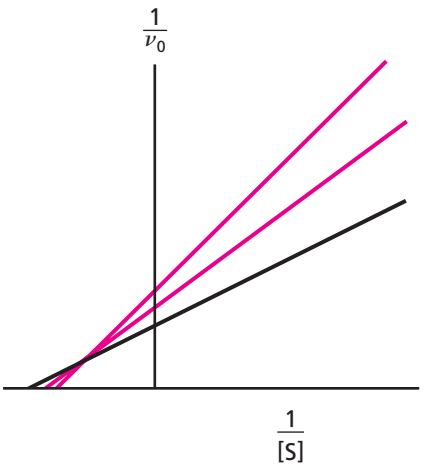
◀ **Figure 5.10**
Uncompetitive inhibition. (a) Kinetic scheme illustrating the binding of I to ES. **(b)** Double-reciprocal plot. In uncompetitive inhibition, both V_{max} and K_m decrease (i.e., the absolute values of both $1/V_{max}$ and $1/K_m$ obtained from the y and x intercepts, respectively, increase). The ratio K_m/V_{max}, the slope of the lines, remains unchanged.



◀ **Figure 5.11**
Classic noncompetitive inhibition. (a) Kinetic scheme illustrating the binding of I to E or ES. **(b)** Double-reciprocal plot. V_{max} decreases, but K_m remains the same.

apparent decrease in V_{max} ($1/V_{max}$ appears to increase) with no change in K_m. On a double-reciprocal plot, the lines for classic noncompetitive inhibition intersect at the point on the x axis corresponding to $1/K_m$ (Figure 5.11b). The common x-axis intercept indicates that K_m isn't affected. The effect of noncompetitive inhibition is to reversibly titrate E and ES with I removing active enzyme molecules from solution. This inhibition cannot be overcome by the addition of S. Classic noncompetitive inhibition is rare but examples are known among allosteric enzymes. In these cases, the noncompetitive inhibitor probably alters the conformation of the enzyme to a shape that can still bind S but cannot catalyze any reaction.

Most enzymes do not conform to the classic form of noncompetitive inhibition where K_m is unchanged. In most cases, both K_m and V_{max} are affected because the affinity of the inhibitor for E is different than its affinity for ES. These cases are often referred to as **mixed inhibition** (Figure 5.12).



▲ **Figure 5.12**
Double-reciprocal plot showing mixed Inhibition. Both V_{max} and K_m are affected when the inhibitor binds with different affinities to E and ES.

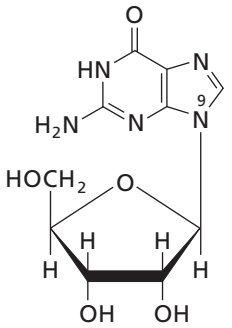
D. Uses of Enzyme Inhibition

Reversible enzyme inhibition provides a powerful tool for probing enzyme activity. Information about the shape and chemical reactivity of the active site of an enzyme can be obtained from experiments involving a series of competitive inhibitors with systematically altered structures.

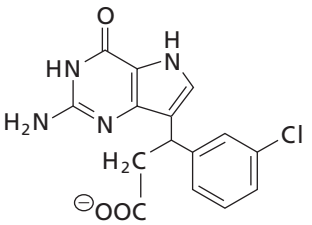
The pharmaceutical industry uses enzyme inhibition studies to design clinically useful drugs. In many cases, a naturally occurring enzyme inhibitor is used as the starting point for drug design. Instead of using random synthesis and testing of potential inhibitors, some investigators are turning to a more efficient approach known as *rational drug design*. Theoretically, with the greatly expanded bank of knowledge about enzyme structure, inhibitors can now be rationally designed to fit the active site of a target enzyme. The effects of a synthetic compound are tested first on isolated enzymes and then in biological systems. However, even if a compound has suitable inhibitory activity, other problems may be encountered. For example, the drug may not enter the target cells, may be rapidly metabolized to an inactive compound, may be toxic to the host organism, or the target cell may develop resistance to the drug.

Downloaded From Vet-e-Books.com

(a)



(b)



▲ **Figure 5.13**
Comparison of a substrate and a designed inhibitor of purine nucleoside phosphorylase. The two substrates of this enzyme are guanosine and inorganic phosphate. **(a)** Guanosine. **(b)** A potent inhibitor of the enzyme. N-9 of guanosine has been replaced by a carbon atom. The chlorinated benzene ring binds to the sugar-binding site of the enzyme, and the acetate side chain binds to the phosphate-binding site.

The advances made in drug synthesis are exemplified by the design of a series of inhibitors of the enzyme purine nucleoside phosphorylase. This enzyme catalyzes a degradative reaction between phosphate and the nucleoside guanosine whose structure is shown in Figure 5.13a. With computer modeling, the structures of potential inhibitors were designed and fit into the active site of the enzyme. One such compound (Figure 5.13b) was synthesized and found to be 100 times more inhibitory than any compound made by the traditional trial-and-error approach. Researchers hope that the rational design approach will produce a drug suitable for treating autoimmune disorders such as rheumatoid arthritis and multiple sclerosis.

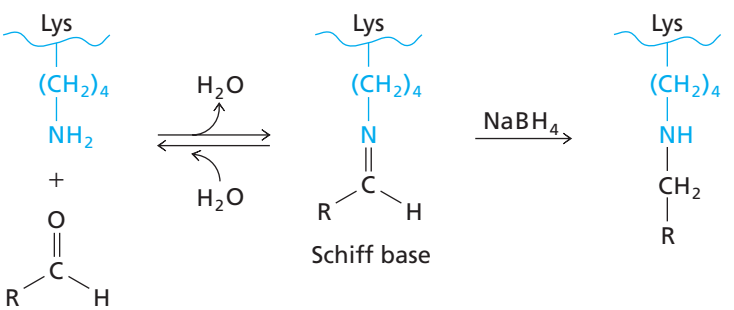
5.8 Irreversible Enzyme Inhibition

In contrast to a reversible enzyme inhibitor, an **irreversible enzyme inhibitor** forms a stable covalent bond with an enzyme molecule thus removing active molecules from the enzyme population. Irreversible inhibition typically occurs by alkylation or acylation of the side chain of an active-site amino acid residue. There are many naturally occurring irreversible inhibitors as well as the synthetic examples described here.

An important use of irreversible inhibitors is the identification of amino acid residues at the active site by specific substitution of their reactive side chains. In this process, an irreversible inhibitor that reacts with only one type of amino acid is incubated with a solution of enzyme that is then tested for loss of activity. Ionizable side chains are modified by acylation or alkylation reactions. For example, free amino groups such as the ϵ-amino group of lysine react with an aldehyde to form a Schiff base that can be stabilized by reduction with sodium borohydride ($NaBH_4$) (Figure 5.14).

The nerve gas diisopropyl fluorophosphate (DFP) is one of a group of organic phosphorus compounds that inactivate hydrolases with a reactive serine as part of the active site. These enzymes are called serine proteases or serine esterases, depending on their reaction specificity. The serine protease chymotrypsin, an important digestive enzyme, is inhibited irreversibly by DFP (Figure 5.15). DFP reacts with the serine residue at chymotrypsin's active site (Ser-195) to produce diisopropylphosphoryl-chymotrypsin.

Some organophosphorus inhibitors are used in agriculture as insecticides; others, such as DFP, are useful reagents for enzyme research. The original organophosphorus nerve gases are extremely toxic poisons developed for military use. The major biological action of these poisons is irreversible inhibition of the serine esterase acetylcholinesterase that catalyzes hydrolysis of the neurotransmitter acetylcholine. When acetylcholine released from an activated nerve cell binds to its receptor on a second nerve cell, it triggers a nerve impulse. The action of acetylcholinesterase restores the cell to its resting state. Inhibition of this enzyme can cause paralysis.



▲ **Figure 5.14**
Reaction of the ϵ-amino group of a lysine residue with an aldehyde. Reduction of the Schiff base with sodium borohydride ($NaBH_4$) forms a stable substituted enzyme.

Downloaded From Vet-e-Books.com

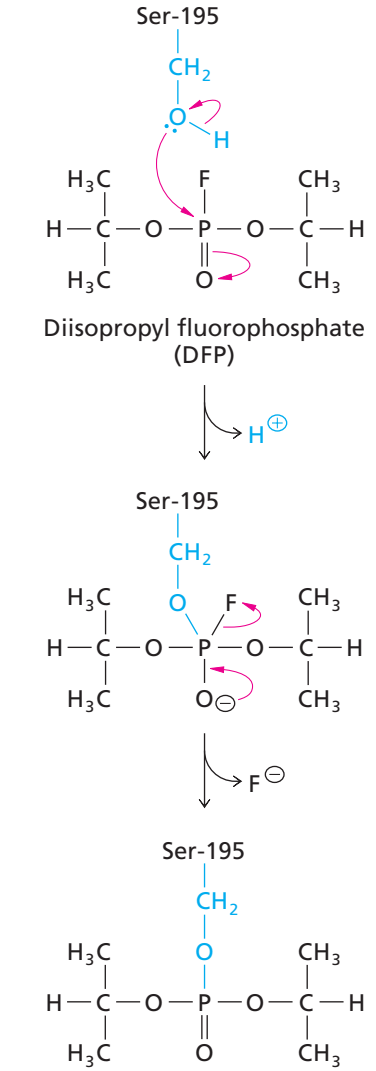


Figure 5.15 ►
Irreversible Inhibition by DFP. Diisopropyl fluorophosphate (DFP) reacts with a single, highly nucleophilic serine residue (Ser-195) at the active site of chymotrypsin, producing inactive diisopropylphosphoryl-chymotrypsin. DFP inactivates serine proteases and serine esterases.

5.9 Regulation of Enzyme Activity

At the beginning of this chapter, we listed several advantages to using enzymes as catalysts in biochemical reactions. Clearly, the most important advantage is to speed up reactions that would otherwise take place too slowly to sustain life. One of the other advantages of enzymes is that their catalytic activity can be regulated in various ways. The *amount* of an enzyme can be controlled by regulating the rate of its synthesis or degradation. This mode of control occurs in all species but it often takes many minutes or hours to synthesize new enzymes or to degrade existing enzymes.

In all organisms, rapid control—on the scale of seconds or less—can be accomplished through reversible modulation of the activity of **regulated enzymes**. In this context, we define regulated enzymes as those enzymes whose activity can be modified in a manner that affects the rate of an enzyme-catalyzed reaction. In many cases, these regulated enzymes control a key step in a metabolic pathway. The activity of a regulated enzyme changes in response to environmental signals, allowing the cell to respond to changing conditions by adjusting the rates of its metabolic processes.

In general, regulated enzymes become more active catalysts when the concentrations of their substrates increase or when the concentrations of the products of their metabolic pathways decrease. They become less active when the concentrations of their substrates decrease or when the products of their metabolic pathways accumulate. Inhibition of the first enzyme unique to a pathway conserves both material and energy by preventing the accumulation of intermediates and the ultimate end product. The activity of regulated enzymes can be controlled by noncovalent allosteric modulation or covalent modification.

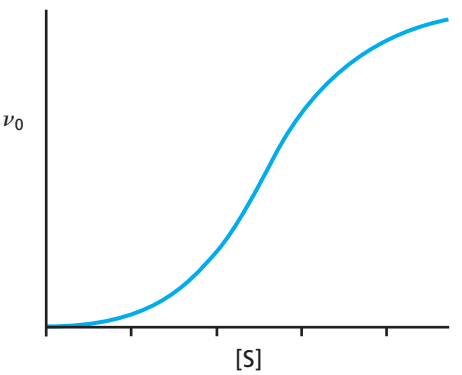
Allosteric enzymes are enzymes whose properties are affected by changes in structure. The structural changes are mediated by interaction with small molecules. We saw an example of allostery in the previous chapter when we examined the binding of oxygen to hemoglobin. Allosteric enzymes often do not exhibit typical Michaelis–Menten kinetics due to cooperative binding of substrate, as is the case with hemoglobin.

Figure 5.16 shows a v_0 versus [S] curve for an allosteric enzyme with cooperative binding of substrate. Sigmoidal curves result from the transition between two states of the enzyme. In the absence of substrate, the enzyme is in the T state. The conformation of each subunit is in a shape that binds substrate inefficiently and the rate of the reaction is slow. As substrate concentration is increased, enzyme molecules begin to bind substrate even though the affinity of the enzyme in the T state is low. When a subunit binds substrate, the enzyme undergoes a conformational change that converts the enzyme to the R state and the reaction takes place. The kinetic properties of the enzyme subunit in the T state and the R state are quite different—each conformation by itself could exhibit standard Michaelis–Menten kinetics.

The conformational change in the subunit that initially binds a substrate molecule affects the other subunits in the multisubunit enzyme. The conformations of these other subunits are shifted toward the R state where their affinity for substrate is much higher. They can now bind substrate at a much lower concentration than when they were in the T state.

Allosteric phenomena are responsible for the reversible control of many regulated enzymes. In Section 4.13C, we saw how the conformation of hemoglobin and its affinity for oxygen change when 2,3-*bis*phosphoglycerate is bound. Many regulated enzymes also undergo allosteric transitions between active (R) states and inactive (T) states. These enzymes have a second ligand-binding site away from their catalytic centers called the **regulatory site** or **allosteric site**. An allosteric inhibitor or activator, also called an allosteric modulator or **allosteric effector**, binds to the regulatory site and causes a conformational change in the regulated enzyme. This conformational change is transmitted

Aspartate transcarbamoylase (ATCase), another well-characterized allosteric enzyme, is described in Chapter 18.

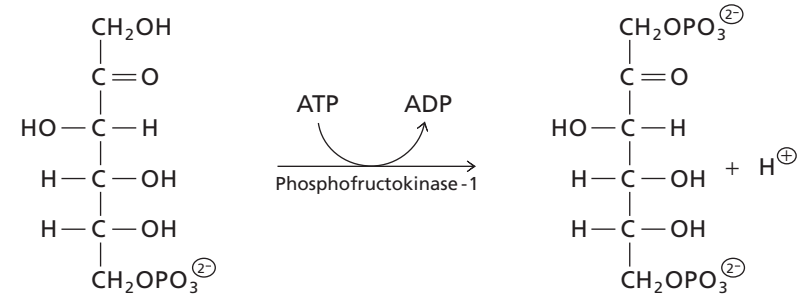


▲ **Figure 5.16**
Cooperativity. Plot of initial velocity as a function of substrate concentration for an allosteric enzyme exhibiting cooperative binding of substrate.

Downloaded From Vet-eBooks.com

KEY CONCEPT

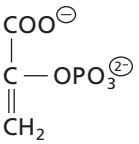
Allosteric enzymes often have multiple subunits and substrate binding is cooperative. This produces a sigmoidal curve when velocity is plotted against substrate concentration.



▲ **Figure 5.17**
Reaction catalyzed by phosphofructokinase-1.

to the active site of the enzyme, which changes shape sufficiently to alter its activity. The regulatory and catalytic sites are physically distinct regions of the protein—usually located on separate domains and sometimes on separate subunits. Allosterically regulated enzymes are often larger than other enzymes.

First, we examine an enzyme that undergoes allosteric (noncovalent) regulation and then we list some general properties of such enzymes. Next, we describe two models that explain allosteric regulation in terms of changes in the conformation of regulated enzymes. Finally, we discuss a closely related group of regulatory enzymes—those subject to covalent modification.



▲ **Figure 5.18**
Phosphoenolpyruvate. This intermediate of glycolysis is an allosteric inhibitor of phosphofructokinase-1 from *Escherichia coli.*

A. Phosphofructokinase Is an Allosteric Enzyme

Bacterial phosphofructokinase-1 (*Escherichia coli*) provides a good example of allosteric inhibition and activation. Phosphofructokinase-1 catalyzes the ATP-dependent phosphorylation of fructose 6-phosphate to produce fructose 1,6-*bis*phosphate and ADP (Figure 5.17). This reaction is one of the first steps of glycolysis, an ATP-generating pathway for glucose degradation described in detail in Chapter 11. Phosphoenolpyruvate (Figure 5.18), an intermediate near the end of the glycolytic pathway, is an allosteric *inhibitor* of *E. coli* phosphofructokinase-1. When the concentration of phosphoenolpyruvate rises, it indicates that the pathway is blocked beyond that point. Further production of phosphoenolpyruvate is prevented by inhibiting phosphofructokinase-1 (see feedback inhibition, Section 10.2C).

ADP is an allosteric activator of phosphofructokinase-1. This may seem strange from looking at Figure 5.17 but keep in mind that the overall pathway of glycolysis results in net synthesis of ATP from ADP. Rising ADP levels indicate a deficiency of ATP and glycolysis needs to be stimulated. Thus, ADP activates phosphofructokinase-1 in spite of the fact that ADP is a product in this particular reaction.

Phosphoenolpyruvate and ADP affect the binding of the substrate fructose 6-phosphate to phosphofructokinase-1. Kinetic experiments have shown that there are four binding sites on phosphofructokinase-1 for fructose 6-phosphate and structural experiments have confirmed that *E. coli* phosphofructokinase-1 (M_r 140,000) is a tetramer consisting of four identical subunits. Figure 5.19 shows the structure of the enzyme complexed with its products, fructose 1,6-*bis*phosphate and ADP, and a second molecule of ADP, an allosteric activator. Two of the subunits shown in Figure 5.19a associate to form a dimer. The two products are bound in the active site located between two domains of each chain—ADP is bound to the large domain and fructose 1,6-*bis*phosphate is bound mostly to the small domain. Two of these dimers interact to form the complete tetrameric enzyme.

A notable feature of the structure of phosphofructokinase-1 (and a general feature of regulated enzymes) is the physical separation of the active site and the regulatory

Downloaded From Vet-e-Books.com

site on each subunit. (In some regulated enzymes the active sites and regulatory sites are on different subunits.) The activator ADP binds at a distance from the active site in a deep hole between the subunits. When ADP is bound to the regulatory site, phosphofructokinase-1 assumes the R conformation, which has a high affinity for fructose 6-phosphate. When the smaller compound phosphoenolpyruvate is bound to the same regulatory site the enzyme assumes a different conformation, the T conformation, which has a lower affinity for fructose 6-phosphate. The transition between conformations is accomplished by a slight rotation of one rigid dimer relative to the other. The cooperativity of substrate binding is tied to the concerted movement of an arginine residue in each of the four fructose 6-phosphate binding sites located near the interface between the dimers. Movement of the side chain of this arginine from the active site lowers the affinity for fructose 6-phosphate. In many organisms, phosphofructokinase-1 is larger and is subject to more complex allosteric regulation than in *E. coli* as you will see in Chapter 11.

Activators can affect either V_{max} or K_m or both. It's important to recognize that the binding of an activator alters the structure of an enzyme and this alteration converts it to a different form that may have quite different kinetic properties. In most cases, the differences between the kinetic properties of the R and T forms are more complex than the differences we saw with enzyme inhibitors in Section 5.7.

B. General Properties of Allosteric Enzymes

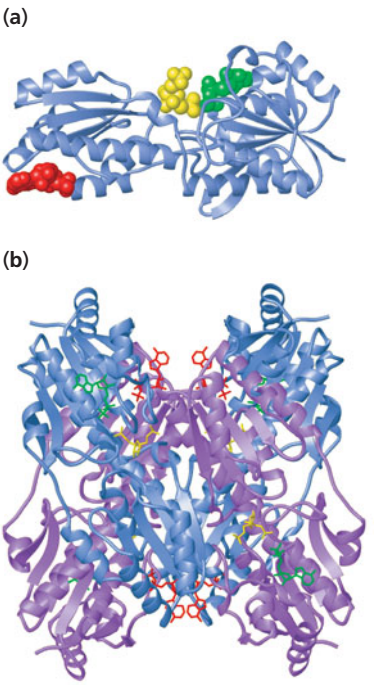
Examination of the kinetic and physical properties of allosteric enzymes has shown that they have the following general features:

1. The activities of allosteric enzymes are changed by metabolic inhibitors and activators. Often these allosteric effectors do not resemble the substrates or products of the enzyme. For example, phosphoenolpyruvate (Figure 5.18) resembles neither the substrate nor the product (Figure 5.17) of phosphofructokinase. Consideration of the structural differences between substrates and metabolic inhibitors originally led to the conclusion that allosteric effectors are bound to regulatory sites separate from catalytic sites.
2. Allosteric effectors bind noncovalently to the enzymes they regulate. (There is a special group of regulated enzymes whose activities are controlled by covalent modification, described in Section 5.10D.) Many effectors alter the K_m of the enzyme for a substrate; but some alter the V_{max}. Allosteric effectors themselves are not altered chemically by the enzyme.
3. With few exceptions, regulated enzymes are multisubunit proteins. (But not all multisubunit enzymes are regulated.) The individual polypeptide chains of a regulated enzyme may be identical or different. For those with identical subunits (such as phosphofructokinase-1 from *E. coli*), each polypeptide chain can contain both the catalytic and regulatory sites and the oligomer is a symmetric complex, most often possessing two or four protein chains. Regulated enzymes composed of nonidentical subunits have more complex, but usually symmetric, arrangements.
4. An allosterically regulated enzyme usually has at least one substrate for which the v_0 versus [S] curve is sigmoidal rather than hyperbolic (Section 5.9). Phosphofructokinase-1 exhibits Michaelis–Menten (hyperbolic) kinetics with respect to one substrate, ATP, but sigmoidal kinetics with respect to its other substrate, fructose 6-phosphate. A sigmoidal curve is caused by positive cooperativity of substrate binding and this is made possible by the presence of multiple substrate binding sites in the enzyme—four binding sites in the case of tetrameric phosphofructokinase-1.

The allosteric R $\rightleftharpoons$ T transition between the active and the inactive conformations of a regulatory enzyme is rapid. The ratio of R to T is controlled by the concentrations of the various ligands and the relative affinities of each conformation for these ligands. In the simplest cases, substrate and activator molecules bind only to enzyme in the R state (E_R) and inhibitor molecules bind only to enzyme in the T state (E_T).

KEY CONCEPT

Allosteric effectors shift the concentrations of the R and T forms of an allosteric enzyme.



▲ **Figure 5.19**
The R conformation of phosphofructokinase-1 from *E. coli*. The enzyme is a tetramer of identical chains. **(a)** Single subunit, shown as a ribbon. The products, fructose 1,6-*bis*phosphate (yellow) and ADP (green), are bound in the active site. The allosteric activator ADP (red) is bound in the regulatory site. **(b)** Tetramer. Two are blue, and two are purple. The products, fructose 1,6-*bis*phosphate (yellow) and ADP (green), are bound in the four active sites. The allosteric activator ADP (red) is bound in the four regulatory sites, at the interface of the subunits. [PDB 1PFK].

The relationship between the regulation of an individual enzyme and a pathway is discussed in Section 10.2B, where we encounter terms such as feedback inhibition and feedforward activation.

Downloaded From Vet-e-Books.com

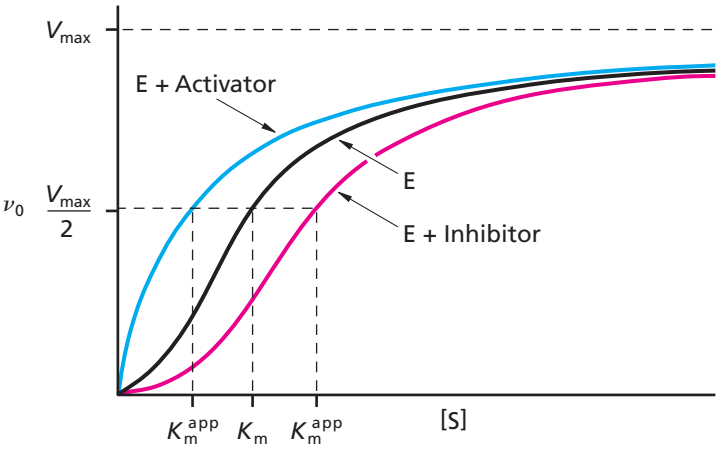
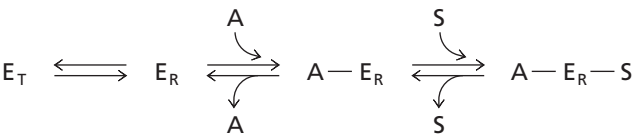


Figure 5.20 ►
Role of cooperativity of binding in regulation. The activity of an allosteric enzyme with a sigmoidal binding curve can be altered markedly when either an activator or an inhibitor is bound to the enzyme. Addition of an activator can lower the apparent K_m raising the activity at a given [S]. Conversely, addition of an inhibitor can raise the apparent K_m producing less activity at a given [S].



(5.29)



(5.30)

These simplified examples illustrate the main property of allosteric effectors—they shift the steady-state concentrations of free E_T and E_R.

Figure 5.20 illustrates the regulatory role that cooperative binding can play. Addition of an activator can shift the sigmoidal curve toward a hyperbolic shape, lowering the apparent K_m (the concentration of substrate required for half-saturation) and raising the activity at a given [S]. The addition of an inhibitor can raise the apparent K_m of the enzyme and lower its activity at any particular concentration of substrate.

The addition of S leads to an increase in the concentration of enzyme in the R conformation. Conversely, the addition of inhibitor increases the proportion of the T species. Activator molecules bind preferentially to the R conformation leading to an increase in the R/T ratio. Note that this simplified scheme does not show that there are multiple interacting binding sites for both S and I.

Some allosteric inhibitors are nonclassical competitive inhibitors (Figure 5.8). For example, Figure 5.20 describes an enzyme that has a higher apparent K_m for its substrate in the presence of the allosteric inhibitor but an unaltered V_{max}. Therefore, the allosteric modulator is a competitive inhibitor.

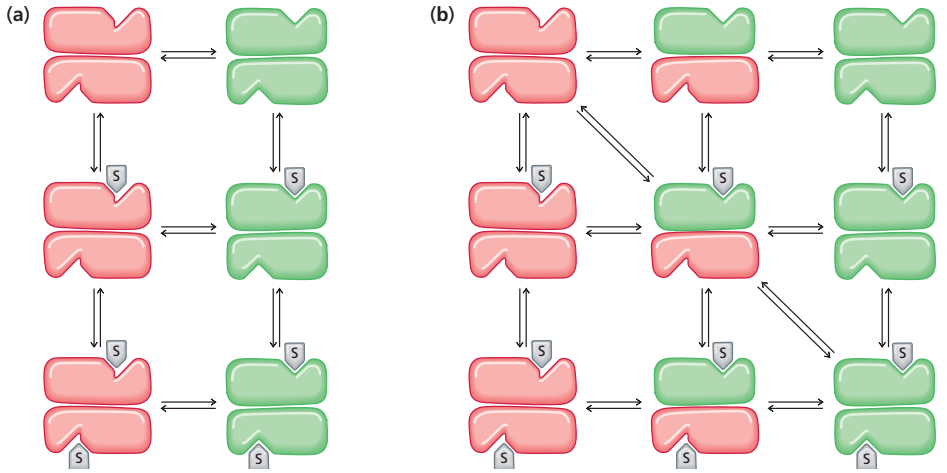
Some regulatory enzymes exhibit noncompetitive inhibition patterns where binding of a modulator at the regulatory site does not *prevent* substrate from binding but appears to distort the conformation of the active site sufficiently to *decrease* the activity of the enzyme.

C. Two Theories of Allosteric Regulation

Recall that most proteins are made up of two or more polypeptide chains (Section 4.8). Enzymes are typical proteins—most of them have multiple subunits. This complicates our understanding of regulation. There are two general models that explain the cooperative binding of ligands to multimeric proteins. Both models describe the cooperative transitions in simple quantitative terms.

The **concerted model**, or symmetry model, was devised to explain the cooperative binding of identical ligands, such as substrates. It was first proposed in 1965 by

Downloaded From Vet-e-Books.com



▲ **Figure 5.21**

Two models for cooperativity of binding of substrate (S) to a tetrameric protein. A two-subunit protein is shown for simplicity. In all cases, the enzymatically active subunit (R) is colored green and the inactive conformation (T) is colored red. **(a)** In the simplified concerted model, both subunits are either in the R conformation or the T conformation. Substrate (S) can bind to subunits in either conformation but binding to T is assumed to be weaker than binding to R. Cooperativity is explained by postulating that when substrate binds to a subunit in the T conformation (red), it shifts the protein into a conformation where both subunits are in the R conformation. **(b)** In the sequential model, one subunit may be in the R conformation while another is in the T conformation. As in the concerted model, both conformations can bind substrate. Cooperativity is achieved by postulating that substrate binding causes the subunit to shift to the R conformation and that when one subunit has adopted the R conformation, the other one is more likely to bind substrate and undergo a conformation change (diagonal lines).

Jacques Monod, Jeffries Wyman, and Jean-Pierre Changeux and it's sometimes known as the MWC model. The concerted model assumes there is one substrate binding site on each subunit. According to the concerted model, the conformation of each subunit is constrained by its association with other subunits and when the protein changes conformation it retains its molecular symmetry (Figure 5.21a). Thus, there are two conformations in equilibrium, R and T. When a subunit is in the R conformation it has a high affinity for the substrate. Subunits in the T conformation have a low affinity for the substrate. The binding of substrate to one subunit shifts the equilibrium since it "locks" the other subunits in the R conformation making it more likely that the other subunits will bind substrate. This explains the cooperativity of substrate binding.

When the conformation of the protein changes, the affinity of its substrate binding sites also changes. The concerted model was extended to include the binding of allosteric effectors and it can be simplified by assuming that the substrate binds only to the R conformation and the allosteric effectors bind preferentially to one of the conformations—inhibitors bind only to subunits in the T conformation and activators bind only to subunits in the R conformation. The concerted model is based on the observed structural symmetry of regulatory enzymes. It suggests that all subunits of a given protein molecule have the same conformation, either all R or all T.

When the enzyme shifts from one conformation to the other, all subunits change conformation in a concerted manner. Experimental data obtained with a number of enzymes can be explained by this simple theory. For example, many of the properties of phosphofructokinase-1 from *E. coli* fit the concerted theory. In most cases, however, the concerted theory does not adequately account for all of the observations concerning a particular enzyme. Their behavior is more complex than that suggested by this simple all-or-nothing model.

The **sequential model** was first proposed by Daniel Koshland, George Némethy, and David Filmer (KNF model). It is a more general model because it allows for both subunits to exist in two different conformations within the same multimeric protein. The specific induced-fit version or the model is based on the idea that a ligand may induce a change in the tertiary structure of each subunit to which it binds. This subunit–ligand

Downloaded From Vet-e-Books.com

complex may change the conformations of neighboring subunits to varying extents. Like the concerted model, the sequential model assumes that only one shape has a high affinity for the ligand but it differs from the concerted model in allowing for the existence of both high- and low-affinity subunits in a multisubunit protein (Figure 5.21b).

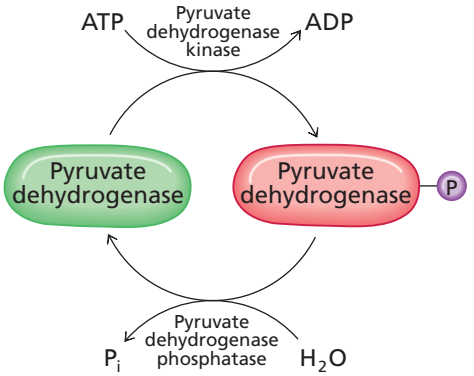
Hundreds of allosteric proteins have been studied and the majority show cooperative binding of substrates and/or effector molecules. It has proven to be very difficult to distinguish between the concerted and sequential models. Many proteins exhibit binding behavior that can best be explained as a mixture of the all-or-nothing shift of the concerted model and the stepwise shift of the sequential model.

D. Regulation by Covalent Modification

The activity of an enzyme can be modified by the covalent attachment and removal of groups on the polypeptide chain. Regulation by covalent modification is usually slower than the allosteric regulation described above. It's important to note that the covalent modification of regulated enzymes must be reversible, otherwise it wouldn't be a form of regulation. The modifications usually require additional modifying enzymes for activation and inactivation. The activities of these modifying enzymes may themselves be allosterically regulated or regulated by covalent modification. Enzymes controlled by covalent modification are believed to generally undergo R $\rightleftharpoons$ T transitions but they may be frozen in one conformation or the other by a covalent substitution.

The most common type of covalent modification is phosphorylation of one or more specific serine residues, although in some cases threonine, tyrosine, or histidine residues are phosphorylated. An enzyme called a protein kinase catalyzes the transfer of the terminal phosphoryl group from ATP to the appropriate serine residue of the regulated enzyme. The phosphoserine of the regulated enzyme is hydrolyzed by the activity of a protein phosphatase, releasing phosphate and returning the enzyme to its dephosphorylated state. Individual enzymes differ as to whether it is their phosphorylated or dephosphorylated forms that are active.

The reactions involved in the regulation of mammalian pyruvate dehydrogenase by covalent modification are shown in Figure 5.22. Pyruvate dehydrogenase catalyzes a reaction that connects the pathway of glycolysis to the citric acid cycle. Phosphorylation of pyruvate dehydrogenase, catalyzed by the allosteric enzyme pyruvate dehydrogenase kinase, inactivates the dehydrogenase. The kinase can be activated by any of several metabolites. Phosphorylated pyruvate dehydrogenase is reactivated under different metabolic conditions by hydrolysis of its phosphoserine residue, catalyzed by pyruvate dehydrogenase phosphatase.



▲ **Figure 5.22**
Regulation of mammalian pyruvate dehydrogenase. Pyruvate dehydrogenase, an interconvertible enzyme, is inactivated by phosphorylation catalyzed by pyruvate dehydrogenase kinase. It is reactivated by hydrolysis of its phosphoserine residue, catalyzed by an allosteric hydrolase called pyruvate dehydrogenase phosphatase.

5.10 Multienzyme Complexes and Multifunctional Enzymes

In some cases, different enzymes that catalyze sequential reactions in the same pathway are bound together in a multienzyme complex. In other cases, different activities may be found on a single multifunctional polypeptide chain. The presence of multiple activities on a single polypeptide chain is usually the result of a gene fusion event.

Some multienzyme complexes are quite stable. We will encounter several of these complexes in other chapters. In other multienzyme complexes the proteins may be associated more weakly (Section 4.9). Because these complexes dissociate easily it has been difficult to demonstrate their existence and importance. Attachment to membranes or cytoskeletal components is another way that enzymes may be associated.

The metabolic advantages of multienzyme complexes and multifunctional enzymes include the possibility of **metabolite channeling**. Channeling of reactants between active sites can occur when the product of one reaction is transferred directly to the next active site without entering the bulk solvent. This can vastly increase the rate of a reaction by decreasing transit times for intermediates between enzymes and by producing local high concentrations of intermediates. Channeling can also protect chemically labile intermediates from degradation by the solvent. Metabolic channeling is one way in which enzymes can effectively couple separate reactions.

Downloaded From Vet-e-Books.com

One of the best-characterized examples of channeling involves the enzyme tryptophan synthase that catalyzes the last two steps in the biosynthesis of tryptophan (Section 17.3F). Tryptophan synthase has a tunnel that conducts a reactant between its two active sites. The structure of the enzyme not only prevents the loss of the reactant to the bulk solvent but also provides allosteric control to keep the reactions occurring at the two active sites in phase.

Several other enzymes have two or three active sites connected by a molecular tunnel. Another mechanism for metabolite channeling involves guiding the reactant along a path of basic amino acid side chains on the surface of coupled enzymes. The metabolites (most of which are negatively charged) are directed between active sites by the electrostatically positive surface path. The fatty acid synthase complex catalyzes a sequence of seven reactions required for the synthesis of fatty acids. The structure of this complex is described in Chapter 16 (Section 16.1).

The search for enzyme complexes and the evaluation of their catalytic and regulatory roles is an extremely active area of research.

The regulation of pyruvate dehydrogenase activity is explained in Section 13.5. An example of a signal transduction pathway involving covalent modification is described in Section 12.6.

Summary

1. Enzymes, the catalysts of living organisms, are remarkable for their catalytic efficiency and their substrate and reaction specificity. With few exceptions, enzymes are proteins or proteins plus cofactors. Enzymes are grouped into six classes (oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases) according to the nature of the reactions they catalyze.
2. The kinetics of a chemical reaction can be described by a rate equation.
3. Enzymes and substrates form noncovalent enzyme–substrate complexes. Consequently, enzymatic reactions are characteristically first order with respect to enzyme concentration and typically show hyperbolic dependence on substrate concentration. The hyperbola is described by the Michaelis–Menten equation.
4. Maximum velocity (V_{max}) is reached when the substrate concentration is saturating. The Michaelis constant (K_m) is equal to the substrate concentration at half-maximal reaction velocity—that is, at half-saturation of E with S.
5. The catalytic constant (k_{cat}), or turnover number, for an enzyme is the maximum number of molecules of substrate that can be transformed into product per molecule of enzyme (or per active site) per second. The ratio k_{cat}/K_m is an apparent second-order rate constant that governs the reaction of an enzyme when the substrate is dilute and nonsaturating. k_{cat}/K_m provides a measure of the catalytic efficiency of an enzyme.
6. K_m and V_{max} can be obtained from plots of initial velocity at a series of substrate concentrations and at a fixed enzyme concentration.
7. Multisubstrate reactions may follow a sequential mechanism with binding and release events being ordered or random, or a ping-pong mechanism.
8. Inhibitors decrease the rates of enzyme-catalyzed reactions. Reversible inhibitors may be competitive (increasing the apparent value of K_m without changing V_{max}), uncompetitive (appearing to decrease K_m and V_{max} proportionally), noncompetitive (appearing to decrease V_{max} without changing K_m), or mixed. Irreversible enzyme inhibitors form covalent bonds with the enzyme.
9. Allosteric modulators bind to enzymes at a site other than the active site and alter enzyme activity. Two models, the concerted model and the sequential model, describe the cooperativity of allosteric enzymes. Covalent modification, usually phosphorylation, of certain regulatory enzymes can also regulate enzyme activity.
10. Multienzyme complexes and multifunctional enzymes are very common. They can channel metabolites between active sites.

Problems

1. Initial velocities have been measured for the reaction of α-chymotrypsin with tyrosine benzyl ester [S] at six different substrate concentrations. Use the data below to make a reasonable estimate of the V_{max} and K_m value for this substrate.

mM [S]	0.00125	0.01	0.04	0.10	2.0	10
(mM/min)	14	35	56	66	69	70

2. Why is the k_{cat}/K_m value used to measure the catalytic proficiency of an enzyme?
   (a) What are the upper limits for k_{cat}/K_m values for enzymes?
   (b) Enzymes with k_{cat}/K_m values approaching these upper limits are said to have reached "catalytic perfection." Explain.
3. Carbonic anhydrase (CA) has a 25,000-fold higher activity ($k_{cat} = 10^6\ s^{-1}$) than orotidine monophosphate decarboxylase (OMPD) ($k_{cat} = 40\ s^{-1}$). However, OMPD provides more than a 10^{10} higher "rate acceleration" than CA (Table 5.2). Explain how this is possible.
4. An enzyme that follows Michaelis–Menten kinetics has a K_m of 1 μM. The initial velocity is 0.1 μM min^{-1} at a substrate concentration of 100 μM. What is the initial velocity when [S] is equal to (a) 1 mM, (b) 1 μM, or (c) 2 μM?
5. Human immunodeficiency virus 1 (HIV-1) encodes a protease (M_r 21,500) that is essential for the assembly and maturation of the virus. The protease catalyzes the hydrolysis of a heptapeptide substrate with a k_{cat} of 1000 s^{-1} and a K_m of 0.075 M.

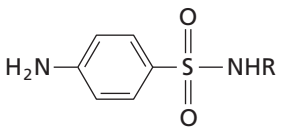
Downloaded From Vet-e-Books.com

(a) Calculate V_{max} for substrate hydrolysis when HIV-1 protease is present at 0.2 mg ml^{-1}.

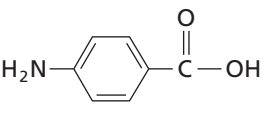
(b) When —C(O)NH— of the heptapeptide is replaced by $—CH_2NH—$, the resulting derivative cannot be cleaved by HIV-1 protease and acts as an inhibitor. Under the same experimental conditions as in part (a), but in the presence of 2.5 μM inhibitor, V_{max} is 9.3×10^{-3} M s^{-1}. What kind of inhibition is occurring? Is this type of inhibition expected for a molecule of this structure?

6. Draw a graph of v_0 versus [S] for a typical enzyme reaction (a) in the absence of an inhibitor, (b) in the presence of a competitive inhibitor, and (c) in the presence of a noncompetitive inhibitor.

7. Sulfonamides (sulfa drugs) such as sulfanilamide are antibacterial drugs that inhibit the enzyme dihydropteroate synthase (DS) that is required for the synthesis of folic acid in bacteria. There is no corresponding enzyme inhibition in animals because folic acid is a required vitamin and cannot be synthesized. If *p* aminobenzoic acid (PABA) is a substrate for DS, what type of inhibition can be predicted for the bacterial synthase enzyme in the presence of sulfonamides? Draw a double reciprocal plot for this type of inhibition with correctly labeled axes and identify the uninhibited and inhibited lines.



Sulfonamides
(R = H, sulfanilamide)



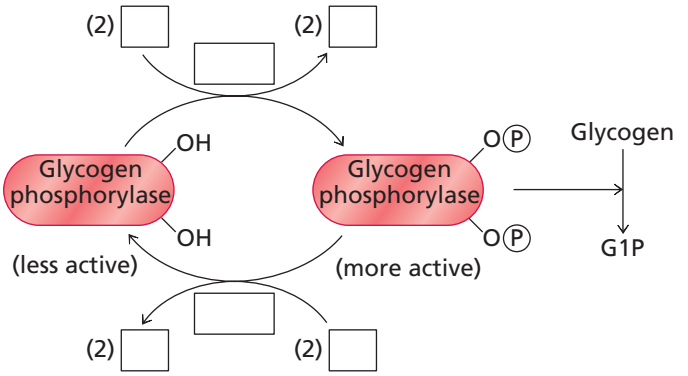
p-Aminobenzoic acid

8. (a) Fumarase is an enzyme in the citric acid cycle that catalyzes the conversion of fumarate to L-malate. Given the fumarate (substrate) concentrations and initial velocities below, construct a Lineweaver–Burk plot and determine the V_{max} and K_m values for the fumarase-catalyzed reaction.

Fumarate (mM)	Rate (mmol l^{-1} min^{-1})
02.0	2.5
03.3	3.1
05.0	3.6
10.0	4.2

(b) Fumarase has a molecular weight of 194,000 and is composed of four identical subunits, each with an active site. If the enzyme concentration is 1×10^{-2} M for the experiment in part (a), calculate the k_{cat} value for the reaction of fumarase with fumarate. *Note*: The units for k_{cat} are reciprocal seconds (s^{-1}).

9. Covalent enzyme regulation plays an important role in the metabolism of muscle glycogen, an energy storage molecule. The active phosphorylated form of glycogen phosphorylase (GP) catalyzes the degradation of glycogen to glucose 1-phosphate. Using pyruvate dehydrogenase as a model (Figure 5.23), fill in the boxes below for the activation and inactivation of muscle glycogen phosphorylase.



10. Regulatory enzymes in metabolic pathways are often found at the first step that is unique to that pathway. How does regulation at this point improve metabolic efficiency?

11. ATCase is a regulatory enzyme at the beginning of the pathway for the biosynthesis of pyrimidine nucleotides. ATCase exhibits positive cooperativity and is activated in vitro by ATP and inhibited by the pyrimidine nucleotide cytidine triphosphate (CTP). Both ATP and CTP affect the K_m for the substrate aspartate but not V_{max}. In the absence of ATP or CTP, the concentration of aspartate required for half-maximal velocity is about 5 mM at saturating concentrations of the second substrate, carbamoyl phosphate. Draw a v_0 versus [aspartate] plot for ATCase, and indicate how CTP and ATP affect v_0 when [aspartate] = 5 mM.

12. The cytochrome P450 family of monooxygenase enzymes are involved in the clearance of foreign compounds (including drugs) from our body. P450s are found in many tissues, including the liver, intestine, nasal tissues, and lung. For every drug that is approved for human use the pharmaceutical company must investigate the metabolism of the drug by cytochrome P450. Many of the adverse drug–drug interactions known to occur are a result of interactions with the cytochrome P450 enzymes. A significant portion of drugs are metabolized by one of the P450 enzymes, P450 3A4. Human intestinal P450 3A4 is known to metabolize midazolam, a sedative, to a hydroxylated product, 1′-hydroxymidazolam. The kinetic data given below are for the reaction catalyzed by P450 3A4.

(a) Focusing on the first two columns, determine the K_m and V_{max} for the enzyme using a Lineweaver–Burk plot.

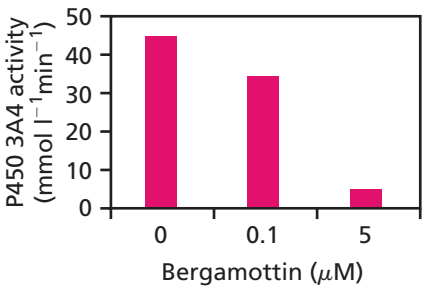
(b) Ketoconazole, an antifungal, is known to cause adverse drug-drug interactions when administered with midazolam. Using the data in the table, determine the type of inhibition that ketoconazole exerts on the P450-catalyzed hydroxylation of midazolam.

Midazolam(μM)	Rate of product formation (pmol l^{-1} min^{-1})	Rate of product formation in the presence of 0.1 μM ketoconazole (pmol l^{-1} min^{-1})
1	100	11
2	156	18
4	222	27
8	323	40

[Adapted from Gibbs, M. A., Thummel, K. E., Shen, D. D., and Kunze, K. L. *Drug Metab. Dispos.* (1999). 27:180–187]

Downloaded From Vet-eBooks.com

13. Patients who are taking certain medications are warned by their physicians to avoid taking these medications with grapefruit juice, which contains many compounds including bergamottin. Cytochrome P450 3A4 is a monooxygenase that is known to metabolize drugs to their inactive forms. The following results were obtained when P450 3A4 activity was measured in the absence or presence of bergamottin.



(a) What is the effect of adding bergamottin to the P450-catalyzed reaction?

(b) Why could it be dangerous for a patient to take certain medications with grapefruit juice?

[Adapted from Wen, Y. H., Sahi, J., Urda, E., Kalkarni, S., Rose, K., Zheng, X., Sinclair, J. F., Cai, H., Strom, S. C., and Kostrubsky, V. E. *Drug Metab. Dispos.* (2002). 30:977–984.]

14. Use the Michaelis-Menten equation (Equation 5.14) to demonstrate the following:

(a) v_0 becomes independent of [S] when [S]>>K_m.

(b) The reaction is first order with respect to S when [S]<<K_m.

(c) [S]>>K_m when v_0 is one-half V_{max}.

Selected Readings

Enzyme Catalysis

Fersht, A. (1985). *Enzyme Structure and Mechanism*, 2nd ed. (New York: W. H. Freeman).

Lewis, C. A., and Wolfenden, R. (2008). Uroporphyrinogen decarboxylation as a benchmark for the catalytic proficiency of enzymes. *Proc. Natl. Acad. Sci. (USA).* 105:17328–17333.

Miller, B. G., and Wolfenden, R. (2002). Catalytic proficiency: the unusual case of OMP decarboxylase. *Annu. Rev. Biochem.* 71, 847–885.

Sigman, D. S., and Boyer, P. D., eds. (1990–1992). *The Enzymes*, Vols. 19 and 20, 3rd ed. (San Diego: Academic Press).

Webb, E. C., ed. (1992). *Enzyme Nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes* (San Diego; Academic Press).

Enzyme Kinetics and Inhibition

Bugg, C. E., Carson, W. M., and Montgomery, J. A. (1993). Drugs by design. *Sci. Am.* 269(6):92–98.

Chandrasekhar, S. (2002). Thermodynamic analysis of enzyme catalysed reactions: new insights into the Michaelis-Menten equation. *Res. Cehm. Intermed.* 28:265–275.

Cleland, W. W. (1970). *Steady State Kinetics. The Enzymes*, Vol. 2, 3rd ed., P. D. Boyer, ed. (New York: Academic Press), pp. 1–65.

Cornish-Bowden, A. (1999). Enzyme kinetics from a metabolic perspective. *Biochem. Soc. Trans.* 27:281–284.

Northrop, D. B. (1998). On the meaning of K_m and V/K in enzyme Kinetics. *J. Chem. Ed.* 75:1153–1157.

Radzicka, A., and Wolfenden, R. (1995). A proficient enzyme. *Science* 267:90–93.

Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady State Enzyme Systems* (New York: Wiley-Interscience).

Regulated Enzymes

Ackers, G. K., Doyle, M. L., Myers, D., and Daugherty, M. A. (1992). Molecular code for cooperativity in hemoglobin. *Science* 255:54–63.

Barford, D. (1991). Molecular mechanisms for the control of enzymic activity by protein phosphorylation. *Biochim. Biophys. Acta* 1133:55–62.

Hilser, V. J. (2010). An ensemble view of allostery. *Science* 327:653–654.

Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., Jr., and Stroud, R. M. (1990). Regulation of an enzyme by phosphorylation at the active site. *Science* 249:1012–1016.

Schirmer, T., and Evans, P. R. (1990). Structural basis of the allosteric behavior of phosphofructokinase. *Nature* 343:140–145.

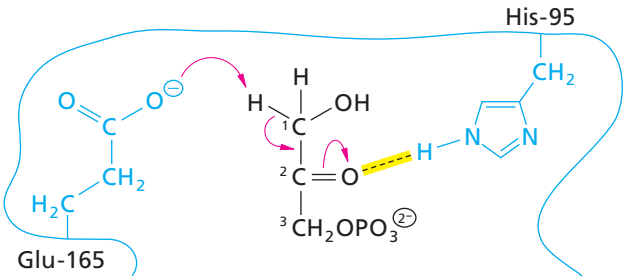
Metabolite Channeling

Pan, P., Woehl, E., and Dunn, M. F. (1997). Protein architecture, dynamics and allostery in tryptophan synthase channeling. *Trends Biochem. Sci.* 22:22–27.

Vélot, C., Mixon, M. B., Teige, M., and Srere, P. A. (1997). Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* 36:14271–14276.

Downloaded From Vet-e-Books.com





Mechanisms of Enzymes

The previous chapter described some general properties of enzymes with an emphasis on enzyme kinetics. In this chapter, we see how enzymes catalyze reactions by studying the molecular details of catalyzed reactions. Individual enzyme mechanisms have been deduced by a variety of methods including kinetic experiments, protein structural studies, and studies of nonenzymatic model reactions. The results of such studies show that the extraordinary catalytic ability of enzymes results from simple physical and chemical properties, especially the binding and proper positioning of reactants in the active sites of enzymes. Chemistry, physics, and biochemistry have combined to take much of the mystery out of enzymes and recombinant DNA technology now allows us to test the theories proposed by enzyme chemists. Observations for which there were no explanations just a half-century ago are now thoroughly understood.

The mechanisms of many enzymes are well established and they give us a general picture of how enzymes function as catalysts. We begin this chapter with a review of simple chemical mechanisms, followed by a brief discussion of catalysis. We then examine the major modes of enzymatic catalysis: acid–base and covalent catalysis (classified as chemical effects) and substrate binding and transition state stabilization (classified as binding effects). We end the chapter with some specific examples of enzyme mechanisms.

I think that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyze.

—Linus Pauling (1948)

6.1 The Terminology of Mechanistic Chemistry

The mechanism of a reaction is a detailed description of the molecular, atomic, and even subatomic events that occur during the reaction. Reactants, products, and any intermediates must be identified. A number of laboratory techniques are used to determine the mechanism of a reaction. For example, the use of isotopically labeled reactants can trace the path of individual atoms and kinetic techniques can measure the changes in chemical bonds of a reactant or solvent during the reaction. Study of the stereochemical changes that occur during the reaction can give a three-dimensional view of the process. For any proposed enzyme mechanism, the mechanistic information about the reactants and intermediates must be coordinated with the three-dimensional structure of the enzyme. This is an important part of understanding structure–function relationships—one of the main themes in biochemistry.

Top: A step from the mechanism of the triose phosphate isomerase reaction.

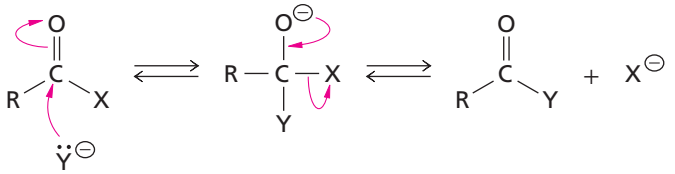
Downloaded From Vet-eBooks.com

Enzymatic mechanisms are described using the same symbolism developed in organic chemistry to represent the breaking and forming of chemical bonds. The movement of electrons is the key to understanding chemical (and enzymatic) reactions. We will review chemical mechanisms in this section and in the following sections we will discuss catalysis and present several specific enzyme mechanisms. This discussion should provide sufficient background for you to understand all the enzyme-catalyzed reactions presented in this book.

A. Nucleophilic Substitutions

Many chemical reactions have ionic substrate, intermediates, or products. There are two types of ionic molecules: one species is electron rich, or **nucleophilic**, and the other species is electron poor, or **electrophilic** (Section 2.6). A nucleophile has a negative charge or an unshared electron pair. We usually think of the nucleophile as attacking the electrophile and call the mechanism a nucleophilic attack or a nucleophilic substitution. In mechanistic chemistry, the movement of a *pair* of electrons is represented by a curved arrow pointing from the available electrons of the nucleophile to the electrophilic center. These "electron pushing" diagrams depict the breaking of an existing covalent bond or the formation of a new covalent bond. The reaction mechanism usually involves an intermediate.

Many biochemical reactions are **group transfer reactions** where a group is moved from one molecule to another. Many of these reactions involve a charged intermediate. The transfer of an acyl group, for example, can be written as the general mechanism

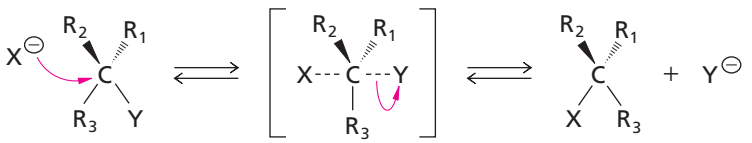


(6.1)

The nucleophile $Y^{\ominus}$ attacks the carbonyl carbon (i.e., adds to the carbonyl carbon atom) to form a tetrahedral addition intermediate from which $X^{\ominus}$ is eliminated. $X^{\ominus}$ is called the leaving group—the group displaced by the attacking nucleophile. This is an example of a **nucleophilic substitution** reaction.

Another type of nucleophilic substitution involves direct displacement. In this mechanism, the attacking group, or molecule, adds to the face of the central atom opposite the leaving group to form a transition state having five groups associated with the central atom. This **transition state** is unstable. It has a structure between that of the reactant and that of the product. (Transition states are shown in square brackets to identify them as unstable, transient entities.)

Transition states are discussed further in Section 6.2.



(6.2)

Transition state

Note that both types of nucleophilic substitution mechanisms involve a transitory state. In the first type (Reaction 6.1), the reaction proceeds in a stepwise manner forming an intermediate molecule that may be stable enough to be detected. In the second type of mechanism (Reaction 6.2), the addition of the attacking nucleophile and the displacement of the leaving group occur simultaneously. The transition state is not a stable intermediate.

B. Cleavage Reactions

We will also encounter cleavage reactions. Covalent bonds can be cleaved in two ways: either both electrons can stay with one atom or one electron can remain with each atom.

Downloaded From Vet-e-Books.com

The two electrons will stay with one atom in most reactions so that an ionic intermediate and a leaving group are formed. For example, cleavage of a C—H bond almost always produces two ions. If the carbon atom retains both electrons then the carbon-containing compound becomes a **carbanion** and the other product is a proton.

$$R_3—C—H \longrightarrow \underset{\text{Carbanion}}{R_3—C:^{\ominus}} + \underset{\text{Proton}}{H^{\oplus}} \qquad \textbf{(6.3)}$$

If the carbon atom loses both electrons, the carbon-containing compound becomes a cationic ion called a **carbocation** and the hydride ion carries a pair of electrons.

$$R_3—C—H \longrightarrow \underset{\text{Carbocation}}{R_3—C^{\oplus}} + \underset{\text{Hydride ion}}{H^{\ominus}} \qquad \textbf{(6.4)}$$

In the second, less common, type of bond cleavage, one electron remains with each product to form two free radicals that are usually very unstable. (A **free radical**, or radical, is a molecule or atom with an unpaired electron.)

$$R_1O—OR_2 \longrightarrow R_1O\cdot + \cdot OR_2 \qquad \textbf{(6.5)}$$

C. Oxidation–Reduction Reactions

Loss of Electrons = Oxidation (LEO)
Gain of Electrons = Reduction (GER)

Remember the phrase: LEO (the lion) says GER

Oxidation Is Loss (OIL)
Reduction Is Gain (RIG)

Remember the phrase: OIL RIG

Oxidation–reduction reactions are central to the supply of biological energy. In an oxidation–reduction (redox) reaction, electrons from one molecule are transferred to another. The terminology here can be a bit confusing so it's important to master the meaning of the words *oxidation* and *reduction*—they will come up repeatedly in the rest of the book. **Oxidation** is the loss of electrons: a substance that is oxidized will have fewer electrons when the reaction is complete. **Reduction** is the gain of electrons: a substance that gains electrons in a reaction is reduced. Oxidation and reduction reactions always occur together. One substrate is oxidized and the other is reduced. An oxidizing agent is a substance that causes an oxidation—it takes electrons from the substrate that is oxidized. Thus, oxidizing agents gain electrons (i.e., they are reduced). A reducing agent is a substance that donates electrons (and is oxidized in the process).

Oxidations can take several forms, such as removal of hydrogen (dehydrogenation), addition of oxygen, or removal of electrons. Dehydrogenation is the most common form of biological oxidation. Recall that oxidoreductases (enzymes that catalyze oxidation–reduction reactions) represent a large class of enzymes and dehydrogenases (enzymes that catalyze removal of hydrogen) are a major subclass of oxidoreductases (Section 5.1).

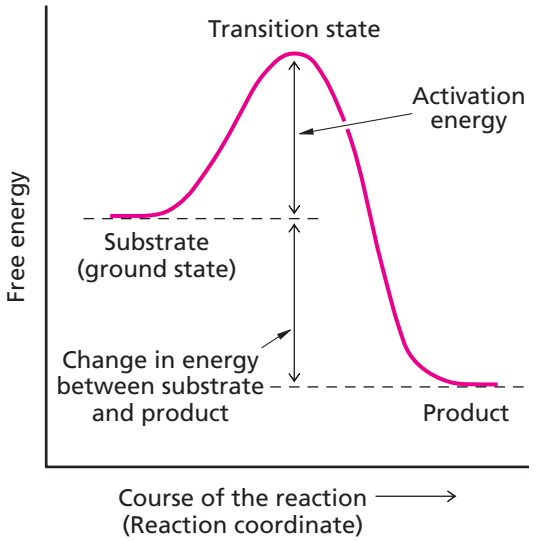
Most dehydrogenations occur by C—H bond cleavage producing a hydride ion ($H^{\ominus}$). The substrate is oxidized because it loses the electrons associated with the hydride ion. Such reactions will be accompanied by a corresponding reduction where another substrate gains electrons by reacting with the hydride ion. The dehydrogenation of lactate (Equation 5.1) is an example of the removal of hydrogen. In this case, the oxidation of lactate is coupled to the reduction of the coenzyme $NAD^{\oplus}$. The role of cofactors in oxidation–reduction reactions will be discussed in the next chapter (Section 7.3) and the free energy of these reactions is described in Section 10.9.

6.2 Catalysts Stabilize Transition States

In order to understand catalysis it's necessary to appreciate the importance of transition states and intermediates in chemical reactions. The rate of a *chemical* reaction depends on how often reacting molecules collide in such a way that a reaction is favored. The colliding substances must be in the correct orientation and must possess sufficient energy to approach the physical configuration of the atoms and bonds of the final product.

As mentioned above, the transition state is an unstable arrangement of atoms in which chemical bonds are in the process of being formed or broken. Transition states

Downloaded From Vet-eBooks.com



◀ **Figure 6.1**
Energy diagram for a single-step reaction. The upper arrow shows the activation energy for the forward reaction. Molecules of substrate that have more free energy than the activation energy pass over the activation barrier and become molecules of product. For reactions with a high activation barrier, energy in the form of heat must be provided in order for the reaction to proceed.

have extremely short lifetimes of about 10^{-14} to 10^{-13} second, the time of one bond vibration. Although they are very difficult to detect, their structures can be predicted. The energy required to reach the transition state from the ground state of the reactants is called the *activation energy* of the reaction and is often referred to as the *activation barrier*.

The progress of a reaction can be represented by an energy diagram, or energy profile. Figure 6.1 is an example that shows the conversion of a substrate (reactant) to a product in a single step. The *y* axis shows the free energies of the reacting species. The *x* axis, called the *reaction coordinate*, measures the progress of the reaction, beginning with the substrate on the left and proceeding to the product on the right. This axis is not time but rather the progress of bond breaking and bond formation of a particular molecule. The transition state occurs at the peak of the activation barrier—this is the energy level that must be exceeded for the reaction to proceed. The lower the barrier the more stable the transition state and the more often the reaction proceeds.

Intermediates, unlike transition states, can be sufficiently stable to be detected or isolated. When there is an intermediate in a reaction, the energy diagram has a trough that represents the free energy of the intermediate as shown in Figure 6.2. This reaction has two transition states, one preceding formation of the intermediate and one preceding its conversion to product. The slowest step, the rate-determining or rate-limiting step, is the step with the highest energy transition state. In Figure 6.2, the rate-determining step is the formation of the intermediate. The intermediate is metastable because relatively little energy is required for the intermediate either to continue to product or to revert to the original reactant. Proposed intermediates that are too short-lived to be isolated or detected are often enclosed in square brackets like transition states, which they presumably closely resemble.

Catalysts create reaction pathways that have lower activation energies than those of uncatalyzed reactions. Catalysts participate directly in reactions by stabilizing the transition states along the reaction pathways. Enzymes are catalysts that accelerate reactions by lowering the overall activation energy. They achieve rate enhancement by providing a multistep pathway (with one or several intermediates) in which each of the steps has lower activation energy than the corresponding stages in the nonenzymatic reaction.

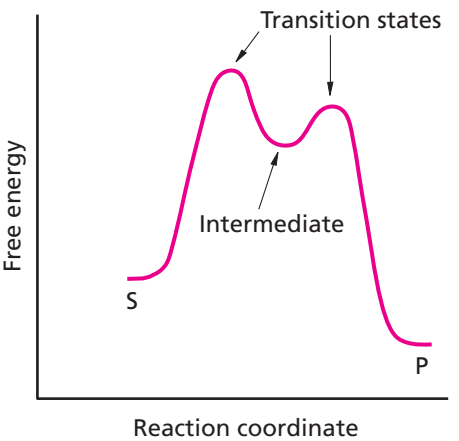
The first step in an enzymatic reaction is the formation of a noncovalent enzyme–substrate complex, ES. In a reaction between A and B, formation of the EAB complex collects and positions the reactants making the probability of reaction much higher for the enzyme-catalyzed reaction than for the uncatalyzed reaction. Figures 6.3a and 6.3b show a hypothetical case in which substrate binding is the only mode of catalysis by an enzyme. In this example, the activation energy is lowered by bringing the reactants together in the substrate binding site. Correct substrate binding accounts for a large part of the catalytic power of enzymes.

The active sites of enzymes bind substrates and products. They also bind transition states. In fact, transition states are likely to bind to active sites much more tightly than

KEY CONCEPT

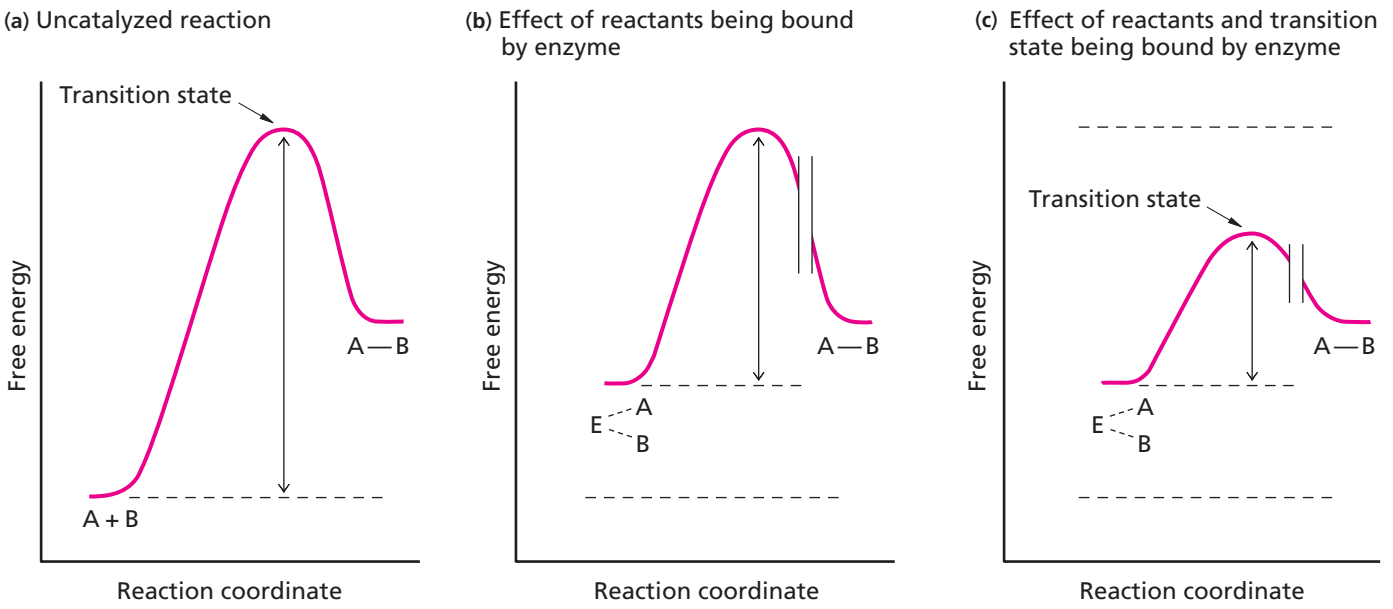
Transition states are unstable molecules with free energies higher than either the substrate or the product.

The meaning of activation energy is described in Section 1.4D.



▲ **Figure 6.2**
Energy diagram for a reaction with an intermediate. The intermediate occurs in the trough between the two transition states. The rate-determining step in the forward direction is formation of the first transition state, the step with the higher energy transition state. S represents the substrate, and P represents the product.

Downloaded From Vet-eBooks.com



▲ **Figure 6.3**
Enzymatic catalysis of the reaction A + B → A—B. (a) Energy diagram for an uncatalyzed reaction. **(b)** Effect of reactant binding. Collection of the two reactants in the EAB complex properly positions them for reaction, makes formation of the transition state more frequent, and hence lowers the activation energy. **(c)** Effect of transition-state stabilization. An enzyme binds the transition state more tightly than it binds substrates, further lowering the activation energy. Thus, an enzymatic reaction has a much lower activation energy than an uncatalyzed reaction. (The breaks in the reaction curves indicate that the enzymes provide multistep pathways.)

substrates do. The extra binding interactions stabilize the transition state, further lowering the activation energy (Figure 6.3c). We will see that the binding of substrates followed by the binding of transition states provides the greatest rate acceleration in enzyme catalysis.

We return to binding phenomena later in this chapter after we examine the chemical processes that underlie enzyme function. (Note that enzyme-catalyzed reactions are usually reversible. The same principles apply to the reverse reaction. The activation energy is lowered by binding the "products" and stabilizing the transition state.)

6.3 Chemical Modes of Enzymatic Catalysis

In addition to reactive amino acid residues, there may be metal ions or coenzymes in the active site. The role of these cofactors in enzyme catalysis is described in Chapter 7.

The formation of an ES complex places reactants in proximity to reactive amino acid residues in the enzyme active site. Ionizable side chains participate in two kinds of chemical catalysis; acid–base catalysis and covalent catalysis. These are the two major chemical modes of catalysis.

A. Polar Amino Acid Residues in Active Sites

The active site cavity of an enzyme is generally lined with hydrophobic amino acid residues. However, a few polar, ionizable residues (and a few molecules of water) may also be present in the active site. Polar amino acid residues (or sometimes coenzymes) undergo chemical changes during enzymatic catalysis. These residues make up much of the catalytic center of the enzyme.

Table 6.1 lists the ionizable residues found in the active sites of enzymes. Histidine, which has a pK_a of about 6 to 7 in proteins, is often an acceptor or a donor of protons. Aspartate, glutamate, and occasionally lysine can also participate in proton transfer. Certain amino acids, such as serine and cysteine, are commonly involved in group-transfer reactions. At neutral pH, aspartate and glutamate usually have negative charges, and lysine and arginine have positive charges. These anions and cations can serve as sites for electrostatic binding of oppositely charged groups on substrates.

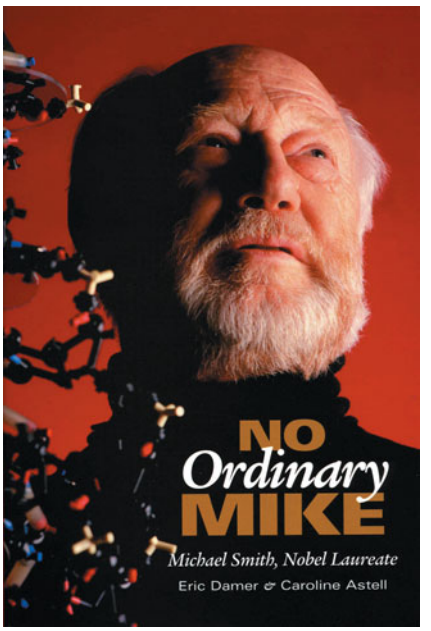
Downloaded From Vet-e-Books.com

BOX 6.1 SITE-DIRECTED MUTAGENESIS MODIFIES ENZYMES

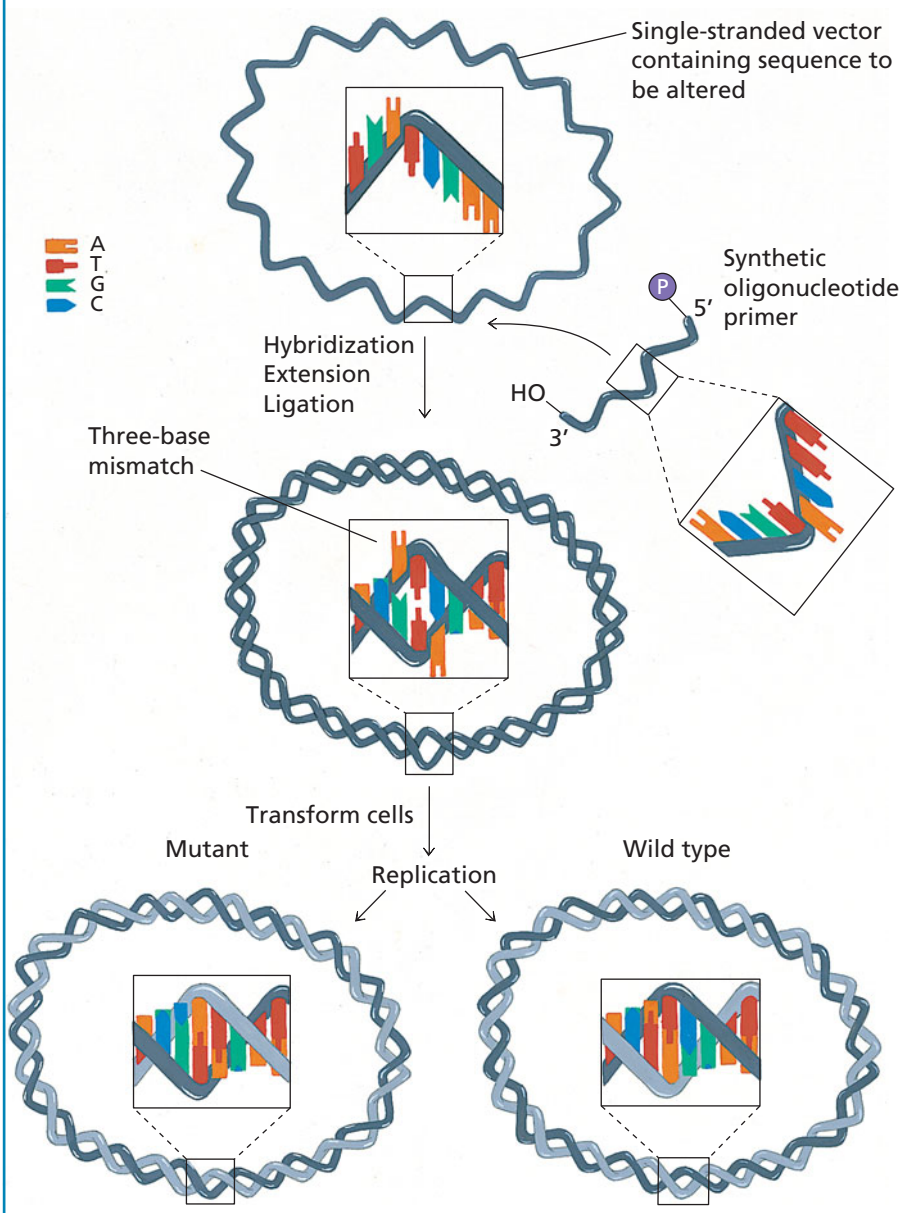
It is possible to test the functions of the amino acid side chains of an enzyme using the technique of *site-directed mutagenesis* (see Section 23.10). This technique has had a huge impact on our understanding of structure–function relationships of enzymes.

In site-directed mutagenesis, a desired mutation is engineered directly into a gene by synthesizing an oligonucleotide that contains the mutation flanked by sequences identical to the target gene. When this oligonucleotide is used as a primer for DNA replication *in vitro*, the new copy of the gene contains the desired mutation. Since alterations can be made at any position in a gene, specific changes in proteins can be engineered allowing direct testing of hypotheses about the functional role of key amino acid residues. Site-directed mutagenesis is commonly used to introduce single codon mutations into genes, resulting in single amino acid substitutions.

The mutated gene can be introduced into bacterial cells where modified enzymes are synthesized from the gene. The structure and activity of the mutant protein can then be analyzed to see the effect of changing an individual amino acid.



▲ Michael Smith (1932–2000), received the Nobel Prize in Chemistry in 1993 for inventing site-directed mutagenesis.



◄ Oligonucleotide-directed, site-specific mutagenesis. A synthetic oligonucleotide containing the desired change (3 bp) is annealed to the single-stranded vector containing the sequence to be altered. The synthetic oligonucleotide serves as a primer for the synthesis of a complementary strand. The double-stranded, circular heteroduplex is transformed into *E. coli* cells where replication produces mutant and wild-type DNA molecules.

Downloaded From Vet-eBooks.com

Table 6.1 Catalytic functions of reactive groups of ionizable amino acids

Amino acid	Reactive group	Net charge at pH 7	Principal functions
Aspartate	$-COO^{\ominus}$	−1	Cation binding; proton transfer
Glutamate	$-COO^{\ominus}$	−1	Cation binding; proton transfer
Histidine	Imidazole	Near 0	Proton transfer
Cysteine	$-CH_2SH$	Near 0	Covalent binding of acyl groups
Tyrosine	Phenol	0	Hydrogen bonding to ligands
Lysine	$NH_3^{\oplus}$	+1	Anion binding; proton transfer
Arginine	Guanidinium	+1	Anion binding
Serine	$-CH_2OH$	0	Covalent binding of acyl groups

Table 6.2 Typical pK_a values of ionizable groups of amino acids in proteins

Group	pK_a
Terminal α-carboxyl	3–4
Side-chain carboxyl	4–5
Imidazole	6–7
Terminal α-amino	7.5–9
Thiol	8–9.5
Phenol	9.5–10
ε-Amino	~10
Guanidine	~12
Hydroxymethyl	~16

Table 6.3 Frequency distribution of catalytic residues in enzymes

	% of catalytic residues	% of all residues
His	18	3
Asp	15	6
Arg	11	5
Glu	11	6
Lys	9	6
Cys	6	1
Tyr	6	4
Asn	5	4
Ser	4	5
Gly	4	8

The pK_a values of the ionizable groups of amino acid residues in proteins may differ from the values of the same groups in free amino acids (Section 3.4). Table 6.2 lists the typical pK_a values of ionizable groups of amino acid residues in proteins. Compare these ranges to the exact values for free amino acids in Table 3.2. A given ionizable group can have different pK_a values within a protein because of differing microenvironments. These differences are usually small but can be significant.

Occasionally, the side chain of a catalytic amino acid residue exhibits a pK_a quite different from the one shown in Table 6.2. Bearing in mind that pK_a values may be perturbed, one can test whether particular amino acids participate in a reaction by examining the effect of pH on the reaction rate. If the change in rate correlates with the pK_a of a certain ionic amino acid (Section 6.3D), a residue of that amino acid may take part in catalysis.

Only a small number of amino acid residues participate directly in catalyzing reactions. Most residues contribute in an indirect way by helping to maintain the correct three-dimensional structure of a protein. As we saw in Chapter 4, the majority of amino acid residues are not evolutionarily conserved.

In vitro mutagenesis studies of enzymes have confirmed that most amino acid substitutions have little effect on enzyme activity. Nevertheless, every enzyme has a few key residues that are absolutely essential for catalysis. Some of these residues are directly involved in the catalytic mechanism, often by acting as an acid or base catalyst or a nucleophile. Other residues act indirectly to assist or enhance the role of a key residue. Other roles for key catalytic residues include substrate binding, stabilization of the transition state, and interacting with essential cofactors.

Enzymes usually have between two and six key catalytic residues. The top ten catalytic residues are listed in Table 6.3. The charged residues, His, Asp, Arg, Glu, and Lys account for almost two-thirds of all catalytic residues. This makes sense since charged side chains are more likely to act as acids, bases, and nucleophiles. They are also more likely to play a role in binding substrates or transition states. The number one catalytic residue is histidine. Histidine is 6 times more likely to be involved in catalysis than its abundance in proteins would suggest.

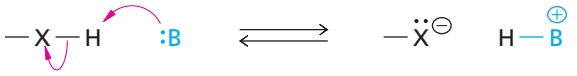
B. Acid–Base Catalysis

In acid–base catalysis, the acceleration of a reaction is achieved by catalytic transfer of a proton. Acid–base catalysis is the most common form of catalysis in organic chemistry and it's also common in enzymatic reactions. Enzymes that employ acid–base catalysis rely on amino acid side chains that can donate and accept protons under the nearly neutral pH conditions of cells. This type of acid–base catalysis, involving proton-transferring agents, is termed *general acid–base catalysis*. (Catalysis by $H^{\oplus}$ or $OH^{\ominus}$ is termed *specific acid* or *specific base catalysis*.) In effect, the active sites of these enzymes provide the biological equivalent of a solution of acid or base.

It is convenient to use B: to represent a base, or proton acceptor, and $BH^{\oplus}$ to represent its conjugate acid, a proton donor. (This acid–base pair can also be written as

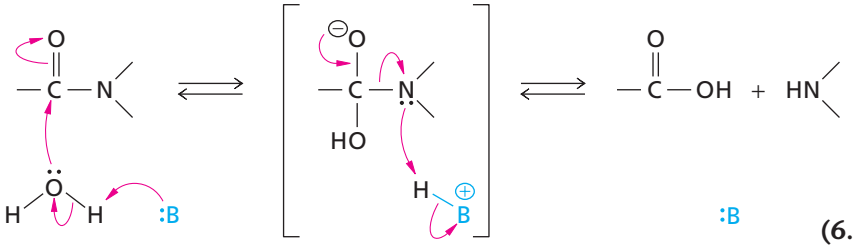
Downloaded From Vet-e-Books.com

$HA/A^{\ominus}$.) A proton acceptor can assist reactions in two ways: (1) it can cleave O—H, N—H, or even some C—H bonds by removing a proton



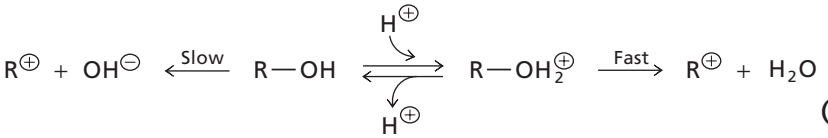
(6.6)

and (2) the general base B: can participate in the cleavage of other bonds involving carbon, such as a C—N bond, by generating the equivalent of $OH^{\ominus}$ in neutral solution through removal of a proton from a molecule of water.



(6.7)

The general acid $BH^{\oplus}$ can also assist in bond cleavage. A covalent bond may break more easily if one of its atoms is protonated. For example,



(6.8)

$BH^{\oplus}$ catalyzes bond cleavage by donating a proton to an atom (such as the oxygen of R—OH in Equation 6.8), thereby making bonds to that atom more labile. In all reactions involving $BH^{\oplus}$ the reverse reaction is catalyzed by B:, and vice versa.

Histidine is an ideal group for proton transfer at neutral pH values because the imidazole/imidazolium of the side chain has a pK_a of about 6 to 7 in most proteins. We have seen that histidine is a common catalytic residue. In the following sections, we will examine some specific roles of histidine side chains.

KEY CONCEPT

In acid–base catalysis, the reaction requires specific amino acid side chains that can donate and accept protons.

C. Covalent Catalysis

In covalent catalysis, a substrate is bound covalently to the enzyme to form a reactive intermediate. The reacting side chain of the enzyme can be either a nucleophile or an electrophile. Nucleophilic catalysis is more common. In the second step of the reaction, a portion of the substrate is transferred from the intermediate to a second substrate. For example, the group X can be transferred from molecule A—X to molecule B in the following two steps via the covalent ES complex X—E:

$$A—X + E \rightleftharpoons X—E + A \qquad (6.9)$$

and

$$X—E + B \rightleftharpoons B—X + E \qquad (6.10)$$

This is a common mechanism for coupling two different reactions in biochemistry. Recall that the ability to couple reactions is one of the important properties of enzymes (Chapter 5; "Introduction"). Transferases, one of the six classes of enzymes (Section 5.1), catalyze group-transfer reactions in this manner and hydrolases catalyze a special kind of group-transfer reaction where water is the acceptor. Transferases and hydrolases together make up more than half of known enzymes.

The reaction catalyzed by bacterial sucrose phosphorylase is an example of group transfer by covalent catalysis. (Sucrose is composed of one glucose residue and one fructose residue.)

$$\text{Sucrose} + P_i \rightleftharpoons \text{Glucose 1-phosphate} + \text{Fructose} \qquad (6.11)$$

Downloaded From Vet-e-Books.com

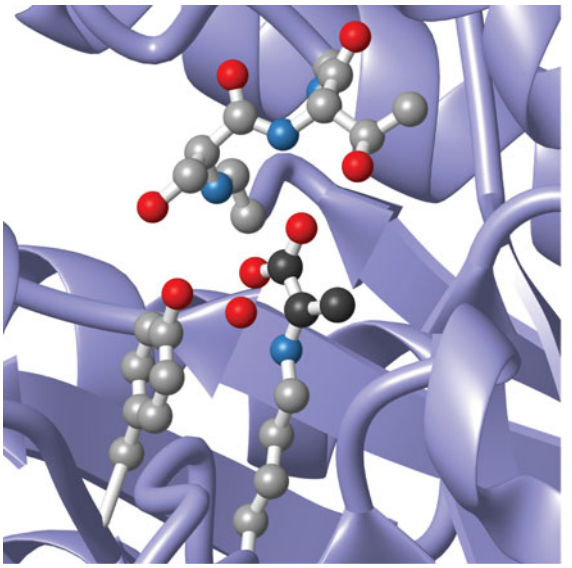


Figure 6.4 ►
Covalent catalysis. The enzyme *N*-acetyl-D-neuraminic acid lyase from *Escherichia coli* catalyzes the condensation of pyruvate and *N*-acetyl-D-mannosamine to form *N*-acetyl-D-neuraminic acid (see Section 8.7C). One of the intermediates in the reaction is a Schiff base (see Fig. 5.15) between pyruvate (black carbon atoms) and a lysine reside. The intermediate is stabilized by hydrogen bonds with other amino acid side chains. [PDB 2WKJ]

KEY CONCEPT

In covalent catalysis mechanisms, the enzyme participates directly in the reaction. It reacts with a substrate and an intermediate containing the enzyme is produced. The reaction is not complete until free enzyme is regenerated.

The first chemical step in the reaction is formation of a covalent glucosyl–enzyme intermediate. In this case, sucrose is equivalent to A—X and glucose is equivalent to X in Reaction 6.9.

$$\text{Sucrose} + \text{Enzyme} \rightleftharpoons \text{Glucosyl}-\text{Enzyme} + \text{Fructose} \qquad \textbf{(6.12)}$$

The covalent ES intermediate can donate the glucose unit either to another molecule of fructose, in the reverse of Reaction 6.12, or to phosphate (which is equivalent to B in Reaction 6.10).

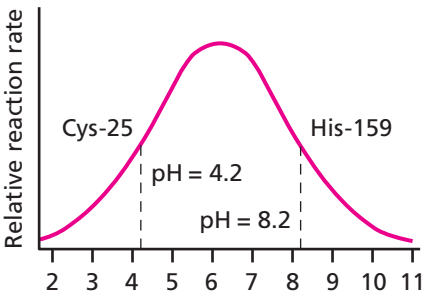
$$\text{Glucosyl}-\text{Enzyme} + P_i \rightleftharpoons \text{Glucose 1}-\text{phosphate} + \text{Enzyme} \qquad \textbf{(6.13)}$$

Proof that an enzyme mechanism relies on covalent catalysis often requires the isolation or detection of an intermediate and demonstration that it is sufficiently reactive. In some cases, the covalently bound intermediate is seen in the crystal structure of an enzyme, and this is direct proof of covalent catalysis (Figure 6.4).

D. pH Affects Enzymatic Rates

The effect of pH on the reaction rate of an enzyme can suggest which ionizable amino acid residues are in its active site. Sensitivity to pH usually reflects an alteration in the ionization state of one or more residues involved in catalysis, although occasionally substrate binding is affected. A plot of reaction velocity versus pH most often yields a bell-shaped curve provided the enzyme is not denatured when the pH is altered.

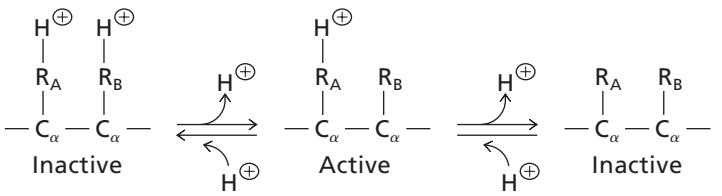
A good example is the pH versus rate profile for papain, a protease isolated from papaya fruit (Figure 6.5). The bell-shaped pH profile can be explained by assuming that the ascending portion of the curve represents the deprotonation of an active-site amino acid residue (B) and the descending portion represents the deprotonation of a second active-site amino acid residue (A). The two inflection points approximate the pK_a values of the two ionizable residues. A simple bell-shaped curve is the result of two overlapping



◄ Figure 6.5
pH vs rate profile for papain. The left and right segments of the bell-shaped curve represent the titrations of the side chains of active-site amino acids. The inflection point at pH 4.2 reflects the pK_a of Cys-25, and the inflection point at pH 8.2 reflects the pK_a of His-159. The enzyme is active only when these ionic groups are present as the thiolate–imidazolium ion pair.

Downloaded From Vet-e-Books.com

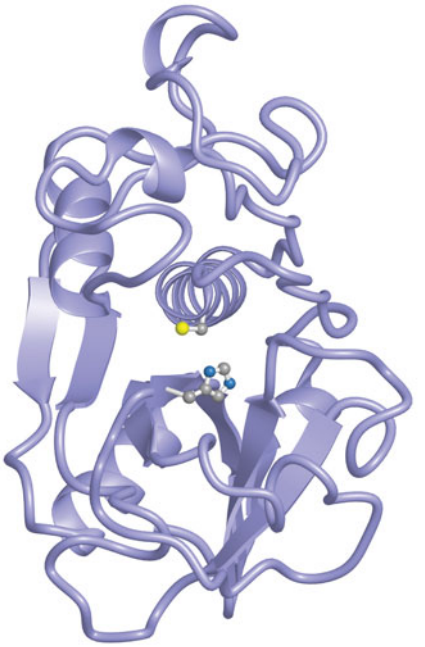
titrations. The side chain of A (R_A) must be protonated for activity and the side chain of B (R_B) must be unprotonated.



(6.14)

At the pH optimum, midway between the two pK_a values, the greatest number of enzyme molecules is in the active form with residue A protonated. Not all pH profiles are bell-shaped. A pH profile is a sigmoidal curve if only one ionizable amino acid residue participates in catalysis and it can have a more complicated shape if more than two ionizable groups participate. Enzymes are routinely assayed near their optimal pH, which is maintained using appropriate buffers.

The pH versus rate graph for papain has inflection points at pH 4.2 and pH 8.2, suggesting that the activity of papain depends on two active-site amino acid residues with pK_a values of about 4 and 8. These ionizable residues are a nucleophilic cysteine (Cys-25) and a proton-donating imidazolium group of histidine (His-159) (Figure 6.6). The side chain of cysteine normally has a pK_a value of 8 to 9.5 but in the active site of papain the pK_a of Cys-25 is greatly perturbed to 3.4. The pK_a of the His-159 residue is perturbed to 8.3. The inflection points on the pH profile do not correspond exactly to the pK_a values of Cys-25 and His-159 because the ionization of additional groups contributes slightly to the overall shape of the curve. Three ionic forms of the catalytic center of papain are shown in Figure 6.7. The enzyme is active only when the thiolate group and the imidazolium group form an ion pair (as in the upper tautomer of the middle pair).



▲ **Figure 6.6 Ionizable residues in papain.** Model of papain, showing ball-and-stick models of the active-site histidine and cysteine side chain. The imidazole nitrogen atoms are blue, and the sulfur atom is yellow.

6.4 Diffusion-Controlled Reactions

A few enzymes catalyze reactions at rates approaching the upper physical limit of reactions in solution. This theoretical upper limit is the rate of diffusion of reactants into the active site. A reaction that occurs with every collision between reactant molecules is termed a **diffusion controlled reaction** or a diffusion-limited reaction. Under physiological conditions the diffusion-controlled rate is about 10^8 to 10^9 M^{-1} s^{-1}. Compare this theoretical maximum to the apparent second-order rate constants (k_{cat}/K_m) for five very fast enzymes listed in Table 6.4.

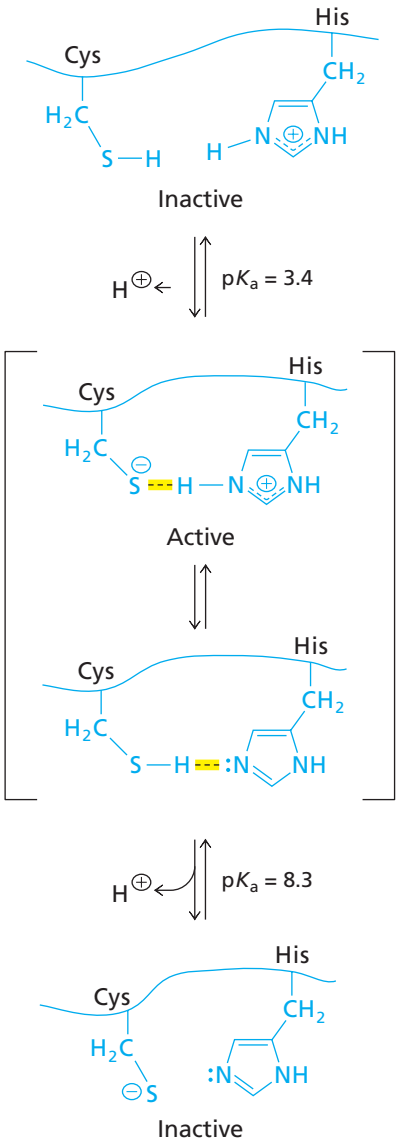
The binding of a substrate to an enzyme is a rapid reaction. If the rest of the reaction is simple and fast, the binding step may be the rate-determining step and the overall rate of the reaction may approach the upper limit for catalysis. Only a few types of chemical reactions can proceed this quickly. These include association reactions, some proton transfers, and electron transfers. The reactions catalyzed by all the enzymes listed in Table 6.4 are so simple that the rate-determining steps are roughly as fast as

Table 6.4 Enzymes with second-order rate constants near the upper limit

Enzyme	Substrate	k_{cat}/K_m (M^{-1} s^{-1})*
Catalase	H_2O_2	4×10^7
Acetylcholinesterase	Acetylcholine	2×10^8
Triose phosphate isomerase	D-Glyceraldehyde 3-phosphate	4×10^8
Fumarase	Fumarate	10^9
Superoxide dismutase	$\cdot O_2^{\ominus}$	2×10^9

*The ratio k_{cat}/K_m is the apparent second-order rate constant for the enzyme-catalyzed reaction E + S → E + P. For these enzymes, the formation of the ES complex can be the slowest step.

Downloaded From Vet-e-Books.com

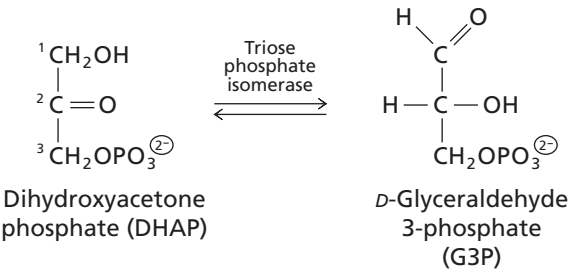


▲ **Figure 6.7 The activity of papain depends on two ionizable residues, histidine (His-159) and cysteine (Cys-25), in the active site.** Three ionic forms of these residues are shown. Only the upper tautomer of the middle pair is active.

binding of substrates to the enzymes. They catalyze diffusion-controlled reactions. We will now look at two of these enzymes in detail: triose phosphate isomerase and superoxide dismutase.

A. Triose Phosphate Isomerase

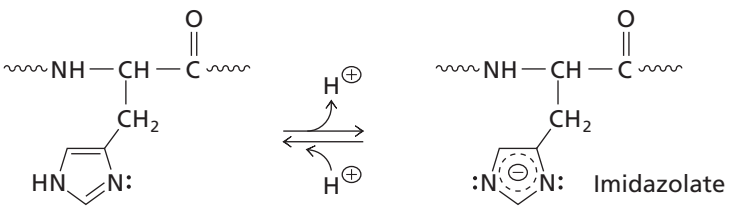
Triose phosphate isomerase catalyzes the rapid interconversion of dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P) in the glycolysis and gluconeogenesis pathways (Chapters 11 and 12).



(6.15)

The reaction proceeds by shifting protons from the carbon atom 1 of DHAP to the carbon atom 2 (Figure 6.8). Triose phosphate isomerase has two ionizable active-site residues: glutamate that acts as a general acid–base catalyst, and histidine that shuttles a proton between oxygen atoms of an enzyme-bound intermediate. When dihydroxyacetone phosphate (DHAP) binds, the carbonyl oxygen forms a hydrogen bond with the imidazole group of His-95. The carboxylate group of Glu-165 removes a proton from C-1 of the substrate to form an enoldiolate transition state (Figure 6.8, top). The transition-state molecule is rapidly converted to a stable enediol intermediate (middle, Figure 6.8). This intermediate is then converted via a second enediolate transition state to D-glyceraldehyde 3-phosphate (G3P).

In this reaction, the proton-donating form of histidine appears to be the neutral species and the proton-accepting species appears to be the imidazolate. The hydrogen bonds formed between histidine and the intermediates in this mechanism appear to be unusually strong.



(6.16)

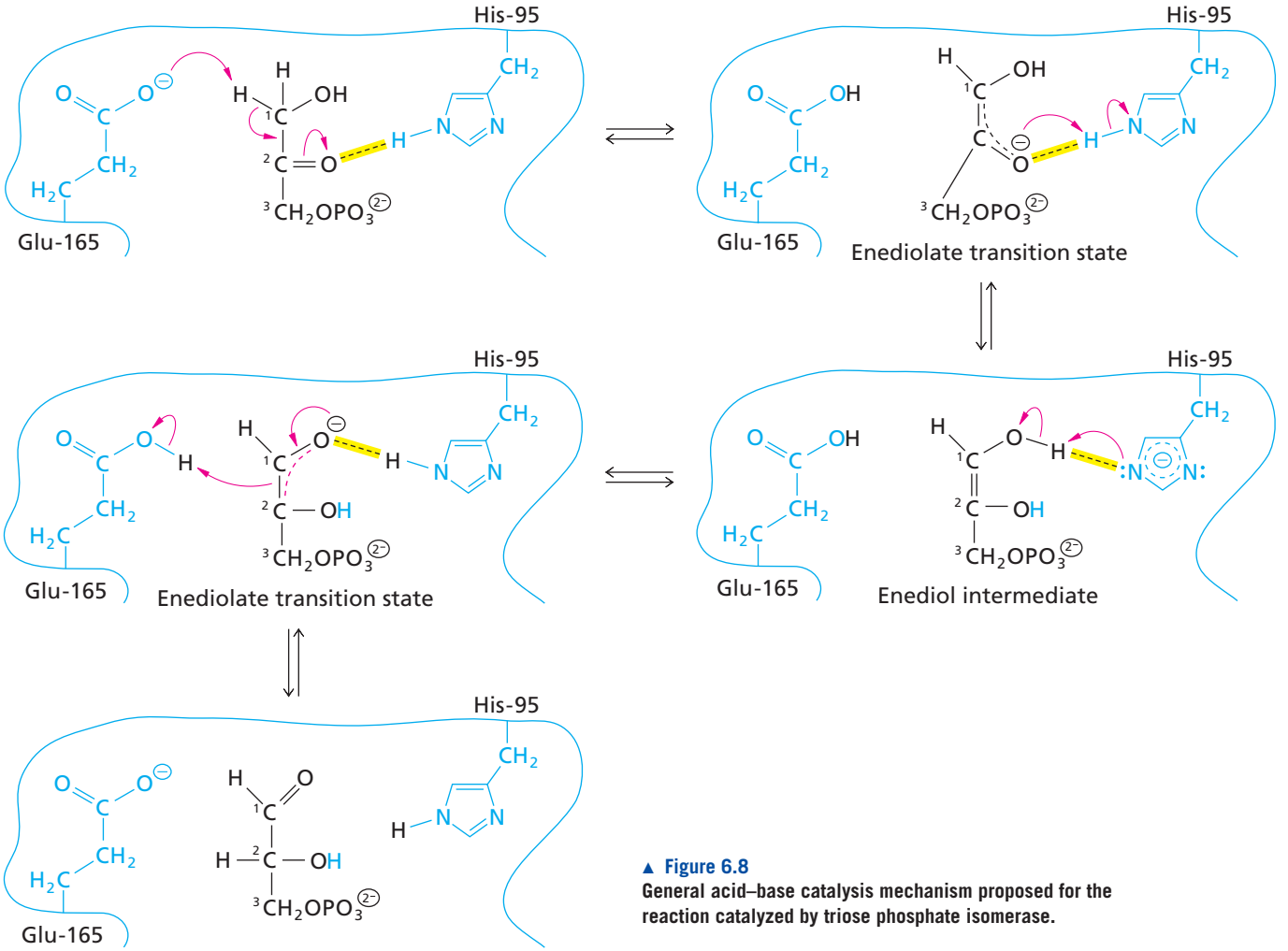
The imidazolate form of a histidine residue is unusual; the triose phosphate isomerase mechanism was the first enzymatic mechanism in which this form was implicated.

The enediol intermediate is stable and in order to prevent it from diffusing out of the active site, triose phosphate isomerase has evolved a "locking" mechanism to seal the active site until the reaction is complete. When substrate binds, a flexible loop of the protein moves to cover the active site and prevent release of the enediol intermediate (Figure 6.9).

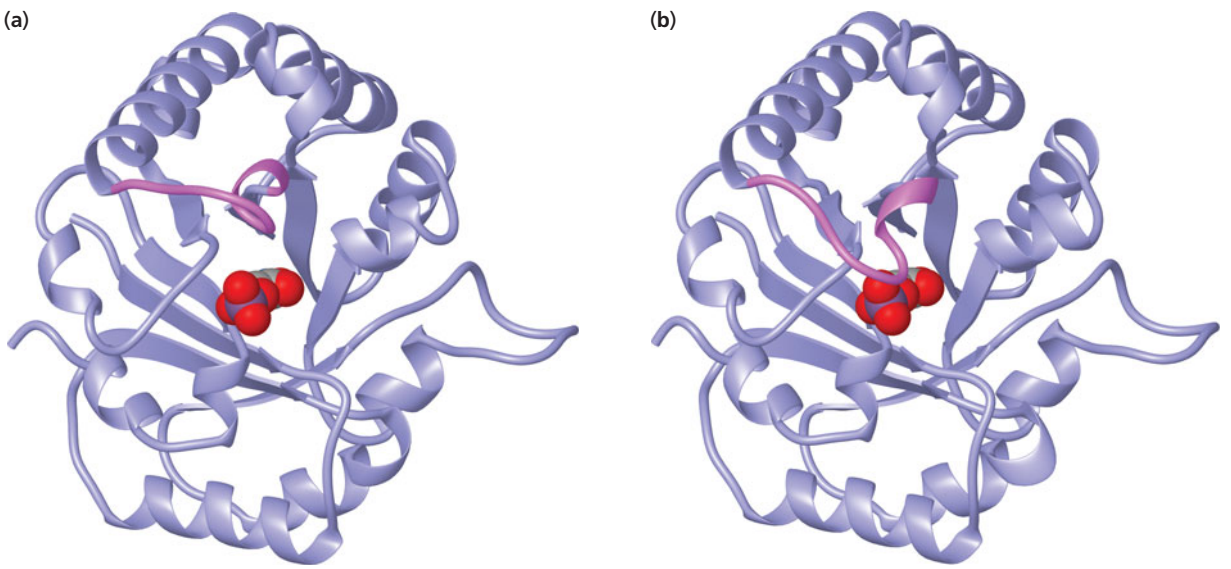
The rate constants of all four kinetically measurable enzymatic steps have been determined.

$$\text{E} + \text{DHAP} \overset{(1)}{\rightleftharpoons} \text{E-DHAP} \overset{(2)}{\rightleftharpoons} \text{E-Intermediate} \overset{(3)}{\rightleftharpoons} \text{E-G3P} \overset{(4)}{\rightleftharpoons} \text{E} + \text{G3P} \qquad \textbf{(6.17)}$$

Downloaded From Vet-e-Books.com



▲ **Figure 6.8**
General acid–base catalysis mechanism proposed for the reaction catalyzed by triose phosphate isomerase.



▲ **Figure 6.9**
Structure of yeast (*Saccharomyces cerevisiae*) triose phosphate isomerase. The location of the substrate is indicated by the space-filling model of a substrate analog. **(a)** The structure of the "open loop" form of the enzyme when the active site is unoccupied. **(b)** The structure when the loop has closed over the active site to prevent release of the enediol intermediate before the reaction is completed.

Downloaded From Vet-e-Books.com

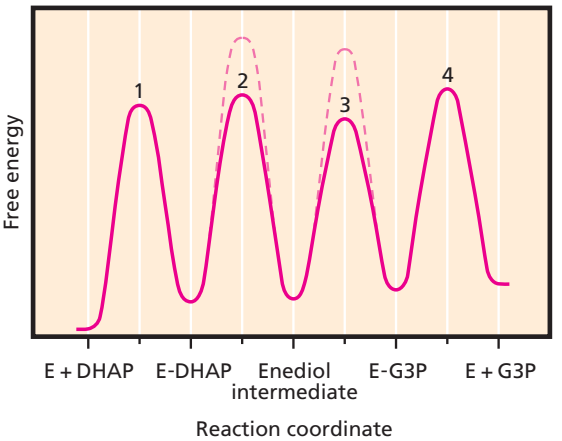


Figure 6.10 ►
Energy diagram for the reaction catalyzed by triose phosphate isomerase. [Adapted from Raines, R. T., Sutton, E. L., Strauss, D. R., Gilbert, W., and Knowles, J. R. (1986). Reaction energetics of a mutant triose phosphate isomerase in which the active-site glutamate has been changed to aspartate. *Biochem.* 25:7142–7154.]

The energy diagram constructed from these rate constants is shown in Figure 6.10. Note that all the barriers for the enzyme are approximately the same height. This means that the steps are balanced, and no single step is rate-limiting. The physical step of S binding to E is rapid but not much faster than the subsequent chemical steps in the reaction sequence. The value of the second-order rate constant k_{cat}/K_m for the conversion of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate is 4×10^8 M^{-1} s^{-1}, which is close to the theoretical rate of a diffusion-controlled reaction. It appears that this isomerase has achieved its maximum possible efficiency as a catalyst.

BOX 6.2 THE "PERFECT ENZYME"?

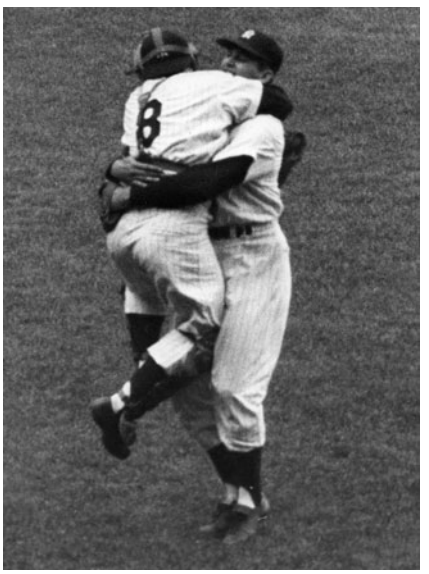
Much of our understanding of the mechanism of triose phosphate isomerase (TPI) comes from the lab of Jeremy Knowles at Harvard University (Cambridge, MA, USA). He points out that the enzyme has achieved catalytic perfection because the overall rate of the reaction is limited only by the rate of diffusion of substrate into the active site. TPI can't work any faster than this!

This has led many people to declare that TPI is the "perfect enzyme" because it has evolved to be so efficient. However, as Knowles and his coworkers have explained, the "perfect enzyme" isn't necessarily one that has evolved the maximum reaction rate. Most enzymes are not under selective pressure to increase their rate of reaction because they are part of a metabolic pathway that meets the cell's needs at less than optimal rates.

Even if it would be beneficial to increase the overall flux in a pathway (i.e., produce more of the end product per second), an individual enzyme need only keep up with the slowest enzyme in the pathway in order to achieve "perfection." The slowest enzyme might be catalyzing a very complicated reaction and might be very efficient. In this case, there will be no selective pressure on the other enzymes to evolve faster mechanisms and they are all "perfect enzymes."

In all species, triose phosphate isomerase is part of the gluconeogenesis pathway leading to the synthesis of glucose. In most species, it also plays a role in the reverse pathway where glucose is degraded (glycolysis). The enzyme is very ancient, and all versions—bacterial and eukaryotic—have achieved catalytic perfection. The two enzymes on either side of the reaction pathway, aldolase and glyceraldehyde 3-phosphate dehydrogenase (Section 11.2), are much slower. Thus, it is by no means obvious why TPI works as fast as it does.

The important point to keep in mind is that the vast majority of enzymes have not evolved catalytic perfection because their *in vivo* rates are "perfectly" adequate for the needs of the cell.



▲ **The Perfect Game.** New York Yankees catcher Yogi Berra congratulates Don Larson for pitching a perfect game in the 1956 World Series against the Brooklyn Dodgers. Perfect games are rare in baseball but there are many "perfect enzymes."

Downloaded From Vet-e-Books.com

B. Superoxide Dismutase

Superoxide dismutase is an even faster catalyst than triose phosphate isomerase. Superoxide dismutase catalyzes the very rapid removal of the toxic superoxide radical anion, $\cdot O_2^{\ominus}$, a by-product of oxidative metabolism. The enzyme catalyzes the conversion of superoxide to molecular oxygen and hydrogen peroxide, which is rapidly removed by the subsequent action of enzymes such as catalase.

$$4\,\cdot O_2^{\ominus} \xrightarrow[\text{Superoxide dismutase}]{4\,H^{\oplus}\ \ 2\,O_2} 2\,H_2O_2 \xrightarrow{\text{Catalase}} 2\,H_2O + O_2 \qquad \textbf{(6.18)}$$

The reaction proceeds in two steps during which an atom of copper bound to the enzyme is reduced and then oxidized.

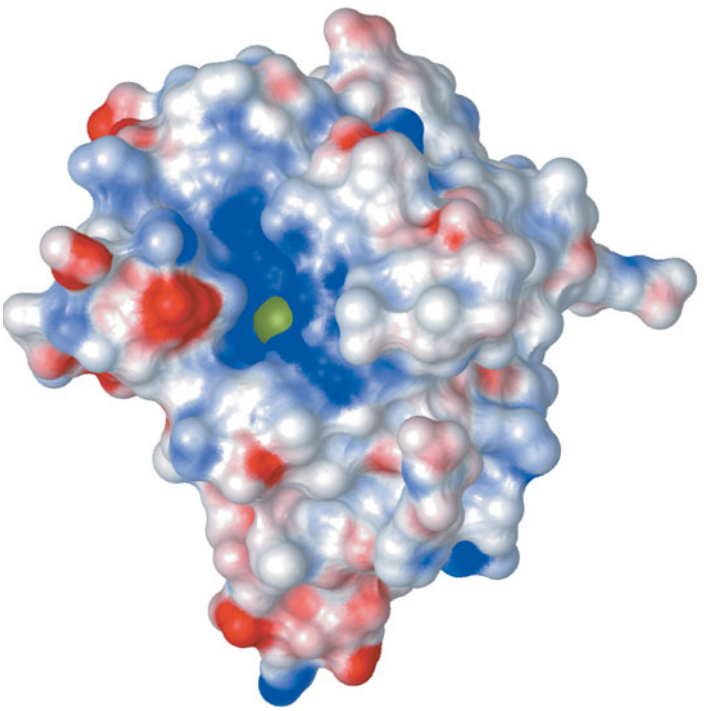
$$\text{E-Cu}^{2+} + \cdot O_2^{\ominus} \longrightarrow \text{E-Cu}^{\oplus} + O_2 \qquad \textbf{(6.19)}$$

$$\text{E-Cu}^{\oplus} + \cdot O_2^{\ominus} + 2H^{\oplus} \longrightarrow \text{E-Cu}^{2+} + H_2O_2 \qquad \textbf{(6.20)}$$

The overall reaction includes binding of the anionic substrate molecules, transfer of electrons and protons, and release of the uncharged products—all very rapid reactions with this enzyme. The k_{cat}/K_m value for superoxide dismutase at 25°C is near $2 \times 10^9\ M^{-1}\ s^{-1}$ (Table 6.4). This rate is even faster than that expected for association of the substrate with the enzyme based on typical diffusion rates.

How can the rate exceed the rate of diffusion? The explanation was revealed when the structure of the enzyme was examined. An electric field around the superoxide dismutase active site enhances the rate of formation of the ES complex about 30-fold. As shown in Figure 6.11, the active-site copper atom lies at the bottom of a deep channel in the protein. Hydrophilic amino acid residues at the rim of the active-site pocket guide negatively charged $\cdot O_2^{\ominus}$ to the positively charged region surrounding the active site. Electrostatic effects allow superoxide dismutase to bind and remove superoxide (radicals) much faster than expected from random collisions of enzyme and substrate.

There are probably many enzymes with enhanced rates of binding due to electrostatic effects. In most cases, the rate-limiting step is catalysis so the overall rate (k_{cat}/K_m) is slower than the maximum for a diffusion-controlled reaction. For those enzymes with fast catalytic reactions, natural selection might favor rapid binding to enhance the overall rate. Similarly, an enzyme with rapid binding might evolve a mechanism that favored a faster reaction. However, most biochemical reactions proceed at rates that are more than sufficient to meet the needs of the cell.

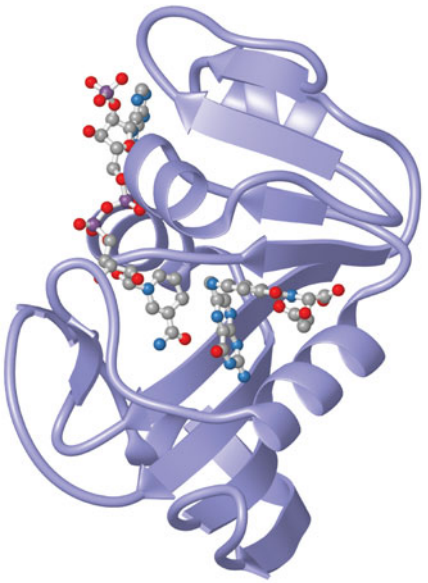


▲ **Figure 6.11**
Surface charge on human superoxide dismutase. The structure of the enzyme is shown as a model that emphasizes the surface of the protein. Positively charged regions are colored blue and negatively charged regions are colored red. The copper atom at the active site is green. Note that the channel leading to the binding site is lined with positively charged residues. [PDB 1HL5]

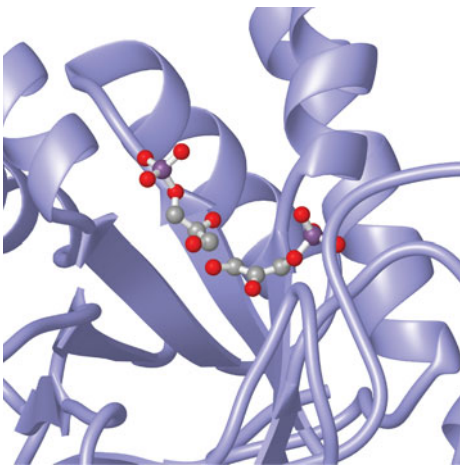
6.5 Modes of Enzymatic Catalysis

The quantitative effects of various catalytic mechanisms are difficult to assess. We have already seen two chemical mechanisms of enzymatic catalysis, acid–base catalysis and covalent catalysis. From studies of nonenzymatic catalysts it is estimated that acid–base catalysis can accelerate a typical enzymatic reaction by a factor of 10 to 100. Covalent catalysis can provide about the same rate acceleration.

Downloaded From Vet-e-Books.com



▲ **Figure 6.12**
Substrate binding. Dihydrofolate reductase binds $NADP^+$ (left) and folate (right), positioning them in the active site in preparation for the reductase reaction. Most of the catalytic rate enhancement is due to binding effects. [PDB 7DFR]



▲ **Figure 6.13**
The proximity effect. The enzyme fructose-1,6-*bis*phosphate aldolase catalyzes the biosynthesis of fructose-1,6-*bis*phosphate from DHAP and G3P during gluconeogenesis and the cleavage of fructose-1,6-*bis*phosphate to dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (G3P) during glycolysis (see Section 11.2#4). In the biosynthesis reaction, the two substrates DHAP and G3P must be positioned close together in the active site in an orientation that promotes their joining to form the larger fructose-1,6-*bis*phosphate. This proximity effect is illustrated for the aldolase from *Mycobacterium tuberculosis*. [PDB 2EKZ]

As important as these chemical modes are, they account for only a small portion of the observed rate accelerations achieved by enzymes (typically 10^8 to 10^{12}). The ability of proteins to specifically bind and orient ligands explains the remainder. The proper binding of reactants in the active sites of enzymes provides not only substrate and reaction specificity but also most of the catalytic power of enzymes (Figure 6.12).

There are two catalytic modes based on binding phenomena. First, for multisubstrate reactions the collecting and correct positioning of substrate molecules in the active site raises their effective concentrations over their concentrations in free solution. In the same way, binding of a substrate near a catalytic active-site residue decreases the activation energy by reducing the entropy while increasing the effective concentrations of these two reactants. High effective concentrations favor the more frequent formation of transition states. This phenomenon is called the *proximity effect*. Efficient catalysis requires fairly weak binding of reactants to enzymes since extremely tight binding would inhibit the reaction.

The second major catalytic mode arising from the ligand–enzyme interaction is the increased binding of transition states to enzymes compared to the binding of substrates or products. This catalytic mode is called *transition state stabilization*. There is an equilibrium (not the reaction equilibrium) between ES and the enzymatic transition state, $ES^‡$. Interaction between the enzyme and its ligands in the transition state shifts this equilibrium toward $ES^‡$ and lowers the activation energy.

The effects of proximity and transition-state stabilization were illustrated in Figure 6.3. Experiments suggest that proximity can increase reaction rates more than 10,000-fold, and transition-state stabilization can increase reaction rates at least that much. Enzymes can achieve extraordinary rate accelerations when both of these effects are multiplied by chemical catalytic effects.

The binding forces responsible for formation of ES complexes and for stabilization of $ES^‡$ are familiar from Chapters 2 and 4. These weak forces are charge–charge interactions, hydrogen bonds, hydrophobic interactions, and van der Waals forces. Charge–charge interactions are stronger in nonpolar environments than in water. Because active sites are largely nonpolar, charge–charge interactions in the active sites of enzymes can be quite strong. The side chains of aspartate, glutamate, histidine, lysine, and arginine residues provide negative and positive groups that form ion pairs with substrates in active sites. Next in bond strength are hydrogen bonds that often form between substrates and enzymes. The peptide backbone and the side chains of many amino acids can form hydrogen bonds. Highly hydrophobic amino acids, as well as alanine, proline, tryptophan, and tyrosine, can participate in hydrophobic interactions with the nonpolar groups of ligands. Many weak van der Waals interactions also help bind substrates. Keep in mind that both the chemical properties of the amino acid residues and the shape of the active site of an enzyme determine which substrates will bind.

A. The Proximity Effect

Enzymes are frequently described as entropy traps—agents that collect highly mobile reactants from dilute solution thereby decreasing their entropy and increasing the probability of their interaction. You can think of the reaction of two molecules positioned at the active site as an *intramolecular* (unimolecular) reaction. The correct positioning of two reacting groups in the active site reduces their degrees of freedom and produces a large loss of entropy sufficient to account for a large rate acceleration (Figure 6.13). The acceleration is expressed in terms of the enhanced relative concentration, called the *effective molarity*, of the reacting groups in the unimolecular reaction. The effective molarity can be obtained from the ratio

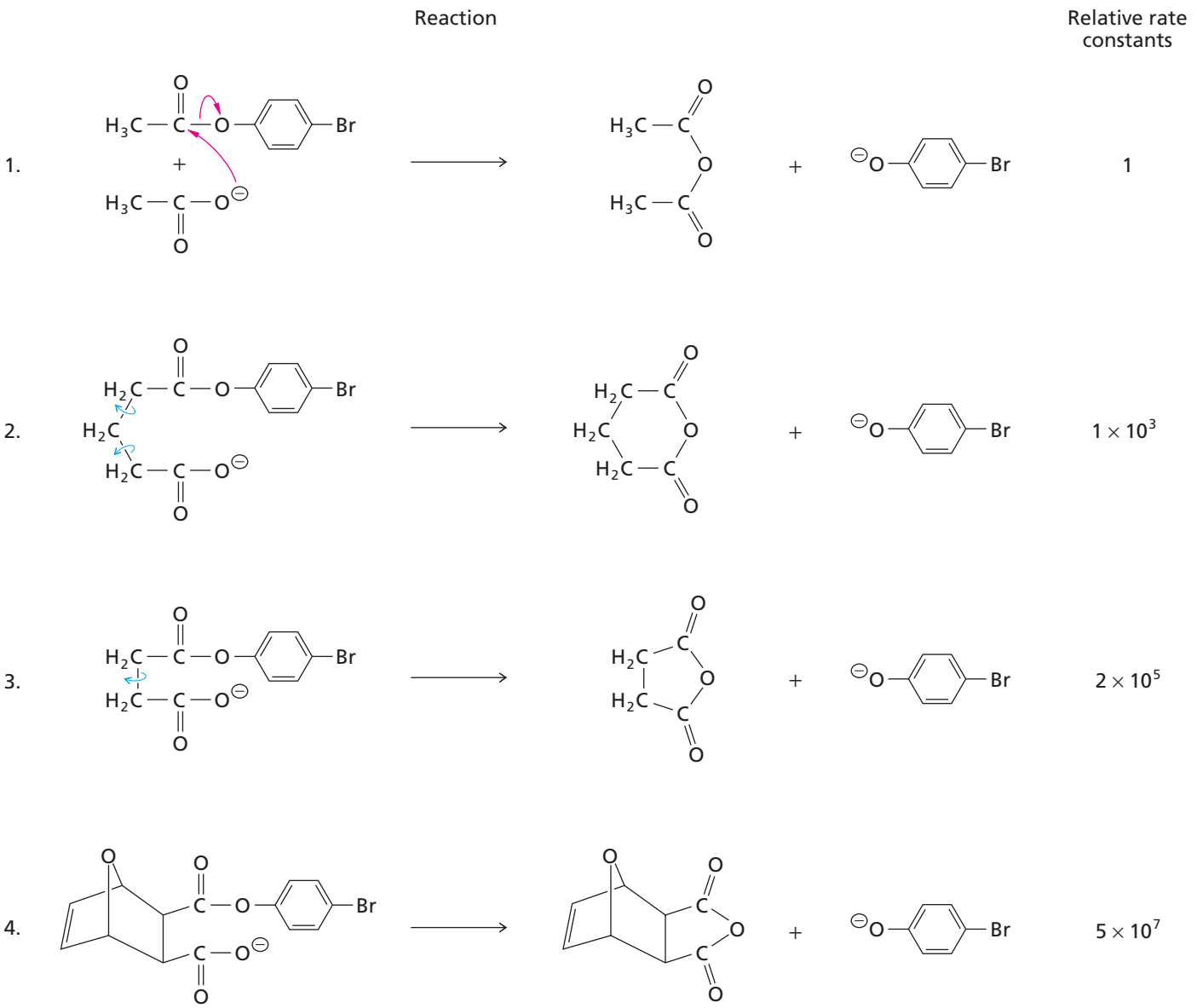
$$\text{Effective molarity} = \frac{k_1(s^{-1})}{k_2(M^{-1}\,s^{-1})} \qquad \textbf{(6.21)}$$

Downloaded From Vet-e-Books.com

where k_1 is the rate constant when the reactants are preassembled into a single molecule and k_2 is the rate constant of the corresponding bimolecular reaction. All the units in this equation cancel except M, so the ratio is expressed in molar units. Effective molarities are not real concentrations; in fact, for some reactions the values are impossibly high. Nevertheless, effective molarities indicate how favorably reactive groups are oriented.

The importance of the proximity effect is illustrated by experiments comparing a nonenzymatic bimolecular reaction to a series of chemically similar intramolecular reactions (Figure 6.14). The bimolecular reaction was the two-step hydrolysis of *p*-bromophenyl acetate, catalyzed by acetate and proceeding via the formation of acetic anhydride. (The second step, hydrolysis of acetic anhydride, is not shown in Figure 6.14.) In the unimolecular version, reacting groups were connected by a bridge with progressively greater restriction of rotation. With each restriction placed on the substrate molecules, the relative rate constant (k_1/k_2) increased markedly. The glutarate ester (compound 2) has two bonds that allow rotational freedom whereas the succinate ester (compound 3) has only one. The most restricted compound, the rigid bicyclic compound 4, has no rotational freedom. In this compound, the carboxylate is

▼ Figure 6.14
Reactions of a series of carboxylates with substituted phenyl esters. The proximity effect is illustrated by the increase in rate observed when the reactants are held more rigidly in proximity. Reaction 4 is 50 million times faster than Reaction 1, the bimolecular reaction. [Based on Bruice and Pandit (1960). Intramolecular models depicting the kinetic importance of "fit" in enzymatic catalysis. *Biochem.* 46:402–404.]



Downloaded From Vet-e-Books.com

KEY CONCEPT

The correct binding and positioning of specific substrates in the active site of an enzyme produces a large acceleration in the rate of a reaction.

close to the ester and the reacting groups are properly aligned. The effective molarity of the carboxylate group is 5×10^7 M. Compound 4 has an extremely high probability of reaction because very little entropy must be lost to reach the transition state. Theoretical considerations suggest that the greatest rate acceleration that can be expected from the proximity effect is about 10^8. This entire rate acceleration can be attributed to the loss of entropy that occurs when two reactants are properly positioned for reaction. These intramolecular reactions can serve as a model of the positioning of two substrates bound in the active site of an enzyme.

B. Weak Binding of Substrates to Enzymes

Reactions of ES complexes are analogous to unimolecular reactions even when two substrates are involved. Although the correct positioning of substrates in an active site produces a large rate acceleration, enzymes do not achieve the maximum 10^8 acceleration theoretically generated by the proximity effect. Typically, the loss in entropy on binding of the substrate allows an acceleration of only 10^4. That's because in ES complexes the reactants are brought toward, but not extremely close to, the transition state. This conclusion is based on both mechanistic reasoning and measurements of the tightness of binding of substrates and inhibitors to enzymes. One major limitation is that binding of substrates to enzymes cannot be extremely tight; that is, K_m values cannot be extremely low.

Figure 6.15 shows energy diagrams for a nonenzymatic unimolecular reaction and the corresponding multistep enzyme-catalyzed reaction. As we will see in the next section, an enzyme increases the rate of a reaction by stabilizing (i.e., tightly binding) the transition state. Therefore, the energy required for ES to reach the transition state ($ES^‡$) in the enzymatic reaction is less than the energy required for S to reach $S^‡$, the transition state in the nonenzymatic reaction.

Recall that the substrate must be bound fairly weakly in the ES complex. If a substrate were bound extremely tightly, it could take just as much energy to reach $ES^‡$ from ES (the arrow labeled 2) as is required to reach $S^‡$ from S in the nonenzymatic reaction (the arrow labeled 1). In other words, extremely tight binding of the substrate would mean little or no catalysis. Excessive ES stability is a thermodynamic pit. The role of enzymes is to bind and position substrates before the transition state is reached but not so tightly that the ES complex is too stable.

The K_m values (representing dissociation constants) of enzymes for their substrates show that enzymes avoid the thermodynamic pit. Most K_m values are on the order of 10^{-4} M, a number that indicates weak binding of the substrate. Enzymes specific for small substrates, such as urea, carbon dioxide, and superoxide anion, exhibit relatively high K_m values for these compounds (10^{-3} to 10^{-4} M) because these molecules can form few noncovalent bonds with enzymes. Enzymes typically have low K_m values

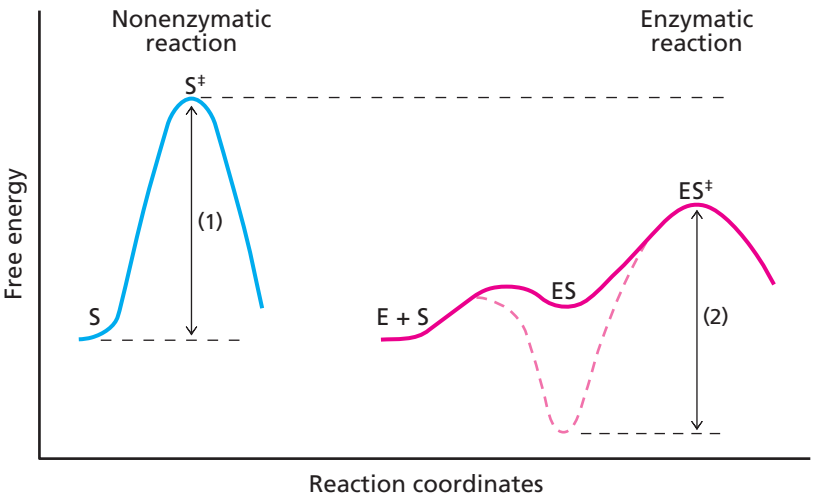


Figure 6.15 ►
Energy of substrate binding. In this hypothetical reaction, the enzyme accelerates the rate of the reaction by stabilizing the transition state. In addition, the activation barrier for formation of the transition state $ES^‡$ from ES must be relatively low. If the enzyme bound the substrate too tightly (dashed profile), the activation barrier (2) would be comparable to the activation barrier of the nonenzymatic reaction (1).

Downloaded From Vet-eBooks.com

(10^{-6} to 10^{-5} M) for coenzymes, which are bulkier than many substrates. The K_m values for the binding of ATP to most ATP-requiring enzymes are about 10^{-4} M or greater but the muscle-fiber protein myosin (which is not an enzyme) binds ATP a billionfold more avidly. This large difference in binding reflects the fact that in an ES complex not all parts of the substrate are bound.

When the concentration of a substrate inside a cell is below the K_m value of its corresponding enzyme, the equilibrium of the binding reaction E + S $\rightleftharpoons$ ES favors E + S. In other words, the formation of the ES complex is slightly uphill energetically (Figures 6.3 and 6.15), and the ES complex is closer to the energy of the transition state than the ground state is. This weak binding of substrates accelerates reactions. K_m values appear to be optimized by evolution for effective catalysis—low enough that proximity is achieved, but high enough that the ES complex is not too stable. The weak binding of substrates is an important feature of another major force that drives enzymatic catalysis—increased binding of reactants in the $ES^{\ddagger}$ transition state.

The meaning of K_m is discussed in Section 5.3C. In most cases, it represents a good approximation of the *dissociation* constant for the reaction E + S $\rightleftharpoons$ ES. Thus, a K_m of 10^{-4} M means that at equilibrium the concentration of ES will be approximately 10,000-fold higher than the concentration of free substrate.

C. Induced Fit

Enzymes resemble solid catalysts by having limited flexibility but they are not entirely rigid molecules. The atoms of proteins are constantly making small, rapid motions, and small conformational adjustments occur on binding of ligands. An enzyme is most effective if it is in the active form initially so no binding energy is consumed in converting it to an active conformation. In some cases, however, enzymes undergo major shape alterations when substrate molecules bind. The enzyme shifts from an inactive to an active form. Activation of an enzyme by a substrate-initiated conformation change is called **induced fit**. Induced fit is not a catalytic mode but primarily a substrate specificity effect.

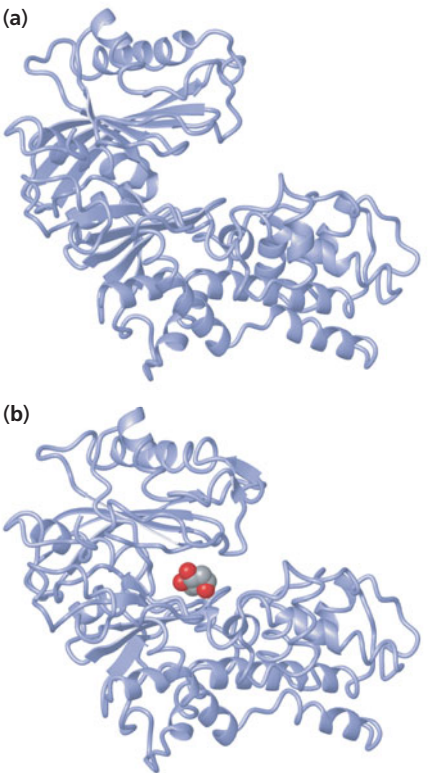
One example of induced fit is seen with hexokinase, an enzyme that catalyzes the phosphorylation of glucose by ATP:

$$\text{Glucose} + \text{ATP} \rightleftharpoons \text{Glucose 6-phosphate} + \text{ADP} \qquad \textbf{(6.22)}$$

Water (HOH), which resembles the alcoholic group at C-6 of glucose (ROH), is small enough and of the proper shape to fit into the active site of hexokinase and therefore it should be a good substrate. If water entered the active site, hexokinase would quickly catalyze the hydrolysis of ATP. However, hexokinase-catalyzed hydrolysis of ATP was shown to be 40,000 times slower than phosphorylation of glucose.

How does the enzyme avoid nonproductive hydrolysis of ATP in the absence of glucose? Structural experiments with hexokinase show that the enzyme exists in two conformations: an open form when glucose is absent, and a closed form when glucose is bound. The angle between the two domains of hexokinase changes considerably when glucose binds, closing the cleft in the enzyme–glucose complex (Figure 6.16). Productive hydrolysis of ATP can only take place in the closed form of the enzyme where the newly formed active site is already occupied by glucose. Water is not a large enough substrate to induce a change in the conformation of hexokinase and this explains why water does not stimulate ATP hydrolysis. Thus, sugar-induced closure of the hexokinase active site prevents wasteful hydrolysis of ATP. A number of other kinases follow induced fit mechanisms.

The substrate specificity that occurs with the induced fit mechanism of hexokinase economizes cellular ATP but exacts a catalytic price. The binding energy consumed in moving the protein molecule into the closed shape—a less-favored conformation—is energy that cannot be used for catalysis. Consequently, an enzyme that uses an induced fit mechanism is less effective as a catalyst than a hypothetical enzyme that is always in an active shape and catalyzes the same reaction. The catalytic cost of induced fit slows kinases so that their k_{cat} values are approximately 10^3 s^{-1} (Table 5.1). We will see another example of induced fit and how it economizes metabolic energy in Section 13.3#1 when we describe citrate synthase. The loop-closing reaction of triose phosphate isomerase is also an example of an induced fit binding mechanism.



▲ **Figure 6.16**
Yeast hexokinase. Yeast hexokinase contains two structural domains connected by a hinge region. On binding of glucose, these domains close, shielding the active site from water. **(a)** Open conformation. **(b)** Closed conformation. [PDB 2YHX and 1HKG].

Downloaded From Vet-e-Books.com

KEY CONCEPT

Most enzymes exhibit some form of induced fit binding mechanism.

Hexokinase, citrate synthase, and triose phosphate isomerase are extreme examples of induced fit mechanisms. Recent advances in the study of enzyme structures reveal that almost all enzymes undergo some conformational change when substrate binds. The simple concept of a rigid lock and a rigid key is being replaced by a more dynamic interaction where both the "lock" (enzyme) and the "key" (substrate) adjust to each other to form a perfect match.

D. Transition-State Stabilization

Enzymes catalyze reactions by physically or electronically distorting the structures of substrates making them similar to the transition state of the reaction. Transition-state stabilization—the increased interaction of the enzyme with the substrate in the transition state—explains a large part of the rate acceleration of enzymes.

Recall Emil Fischer's lock-and-key theory of enzyme specificity described in Section 5.2B. Fischer proposed that enzymes were rigid templates that accepted only certain substrates as keys. This idea has been replaced by a more dynamic model where both enzyme and substrate change conformations when they interact. Furthermore, the classic lock-and-key model dealt with the interaction between enzyme and substrate but we now think of it in terms of enzyme and transition state—the "key" in the "lock" is the transition state and not the substrate molecule. When a substrate binds to an enzyme the enzyme distorts the structure of the substrate forcing it toward the transition state. Maximal interaction with the substrate molecule occurs only in $ES^{\ddagger}$. A portion of this binding in $ES^{\ddagger}$ can be between the enzyme and nonreacting portions of the substrate.

KEY CONCEPT

The catalytic power of enzymes is explained by binding effects (positioning the substrates together in the correct orientation) and stabilization of the transition state. The result is a lower activation energy and an increased rate of reaction.

An enzyme must be complementary to the transition state in shape and in chemical character. The graph in Figure 6.15 shows that tight binding of the transition state to an enzyme can lower the activation energy. Because the energy difference between E + S and $ES^{\ddagger}$ is significantly less than the energy difference between S and $S^{\ddagger}$, k_{cat} is greater than k_n (the rate constant for the nonenzymatic reaction). The enzyme–substrate transition state ($ES^{\ddagger}$) is lower in absolute energy—and therefore more stable—than the transition state of the reactant in the uncatalyzed reaction. Some transition states may bind to their enzymes 10^{10} to 10^{15} times more tightly than their substrates do. The affinity of other enzymes for their transition states need not be that extreme. A major task for biochemists is to show how transition state stabilization occurs.

The comparative stabilization of $ES^{\ddagger}$ could occur if an enzyme has an active site with a shape and an electrostatic structure that more closely fits the transition state than the substrate. An undistorted substrate molecule would not be fully bound. For example, an enzyme could have sites that bind the partial charges present only in the unstable transition state.

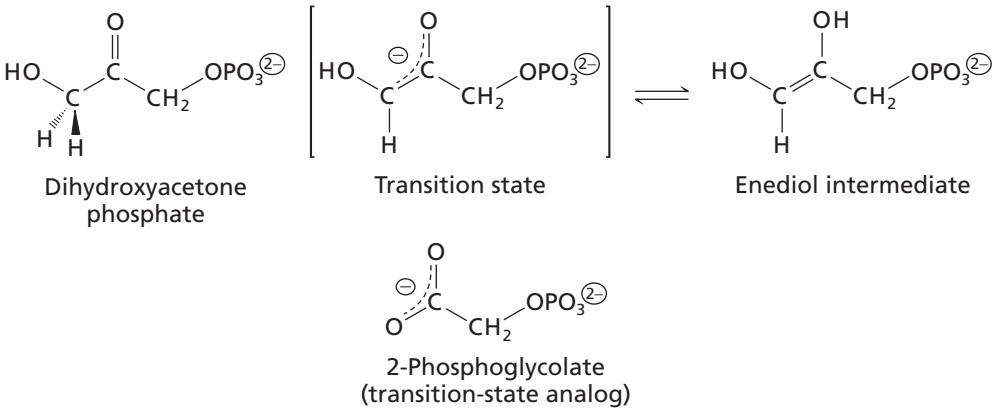
Transition-state molecules are ephemeral—they have very short half-lives and are difficult to detect. One way in which biochemists can study transition states is to create stable analogs that can bind to the enzyme. These **transition-state analogs** are molecules whose structures resemble presumed transition states. If enzymes prefer to bind to transition states, then a transition-state analog should bind extremely tightly to the appropriate enzyme—much more tightly than substrate—and thus be a potent inhibitor. The dissociation constant for a transition state analog should be about 10^{-13} M or less.

The role of adenosine deaminase is described in Section 18.8.

One of the first examples of a transition-state analog was 2-phosphoglycolate (Figure 6.17), whose structure resembles the first enediolate transition state in the reaction catalyzed by triose phosphate isomerase (Section 6.4A). This transition-state analog binds to the isomerase at least 100 times more tightly than either of the substrates of the enzyme (Figure 6.18). Tighter binding results from a partially negative oxygen atom in the carboxylate group of 2-phosphoglycolate, a feature shared with the transition state but not with the substrates.

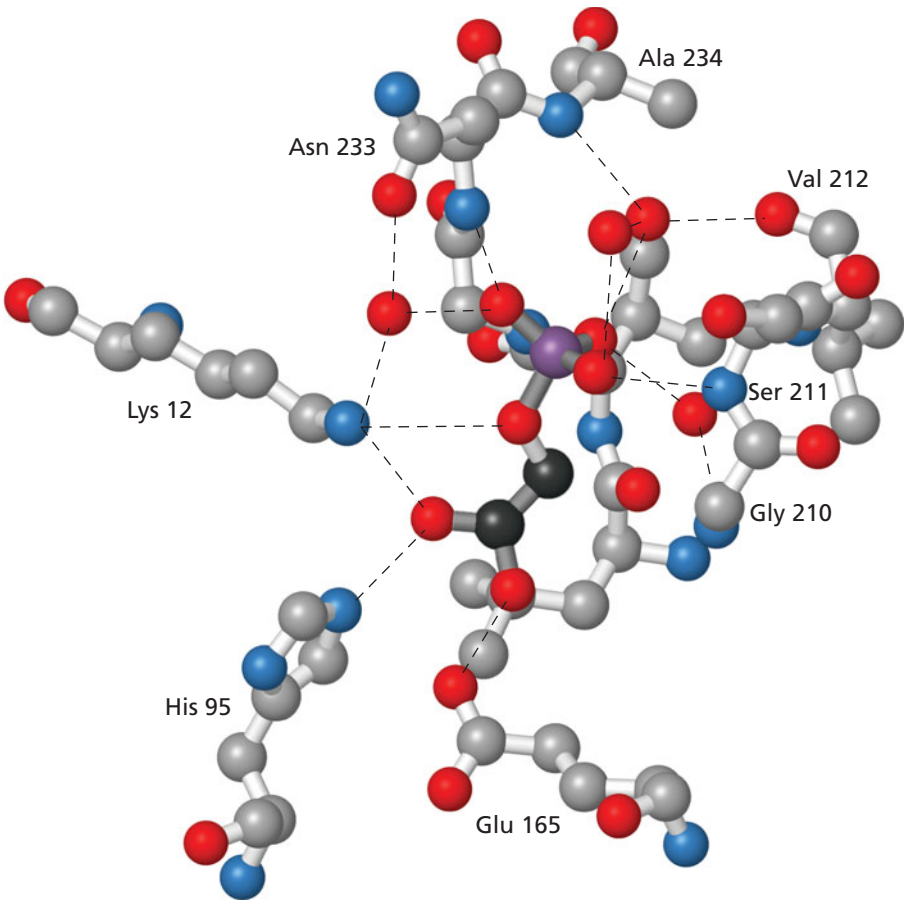
Experiments with adenosine deaminase have identified a transition-state analog that binds to the enzyme with amazing affinity because it resembles the transition state very closely. Adenosine deaminase catalyzes the hydrolytic conversion of the purine nucleoside adenosine to inosine. The first step of this reaction is the addition of a molecule

Downloaded From Vet-e-Books.com



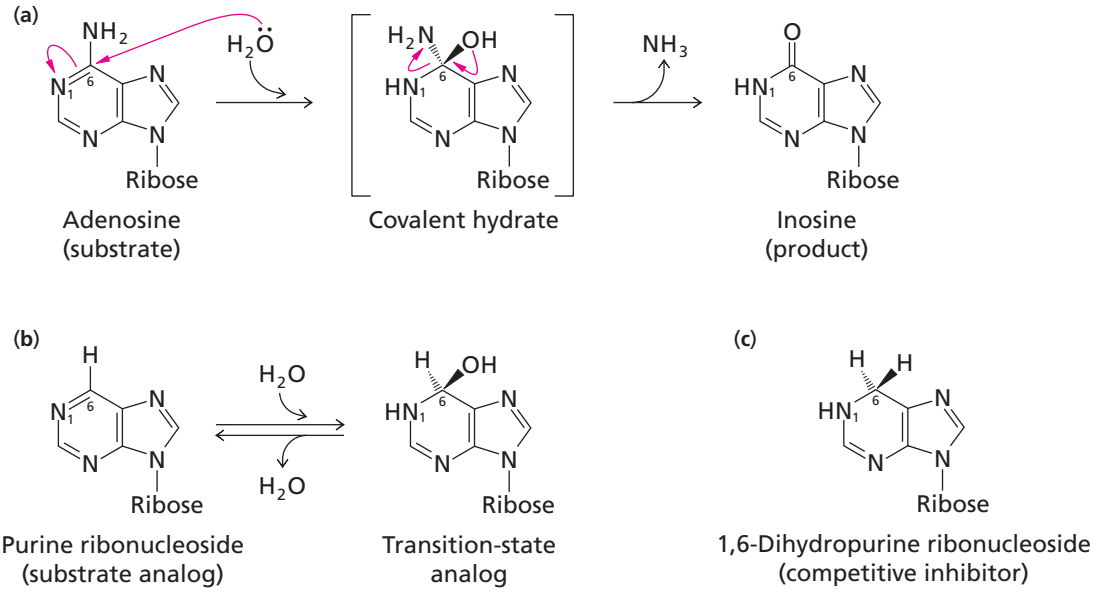
◀ **Figure 6.17**
2-Phosphoglycolate, a transition-state analog for the enzyme triose phosphate isomerase. 2-Phosphoglycolate is presumed to be an analog of C-2 and C-3 of the transition state (center) between dihydroxyacetone phosphate (right) and the initial enediolate intermediate in the reaction.

of water (Figure 6.19a). The complex with water, called a covalent hydrate, forms as soon as adenosine is bound to the enzyme and quickly decomposes to products. Adenosine deaminase has broad substrate specificity and catalyzes the hydrolytic removal of various groups from position 6 of purine nucleosides. However, the inhibitor purine ribonucleoside (Figure 6.19b) has just hydrogen at position 6 and undergoes only the first enzymatic step of hydrolysis, addition of the water molecule. The covalent hydrate that's formed is a transition-state analog, a competitive inhibitor having a K_i of 3×10^{-13} M. (For comparison, the affinity constant of adenosine deaminase for its true transition state is expected to be 3×10^{-17} M.). The binding of this analog exceeds the binding of either the substrate or the product by a factor of more than 10^8. A very similar reduced inhibitor, 1,6-dihydropurine ribonucleoside (Figure 6.19c), lacks the hydroxyl group at C-6, and it has a K_i of only 5×10^{-6} M. We can conclude from these studies that adenosine

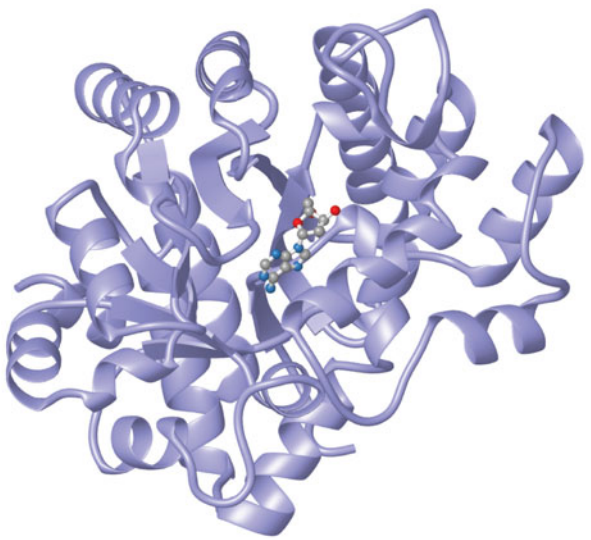


◀ **Figure 6.18**
Binding of 2-phosphoglycolate to triose phosphate isomerase. The transition state analogue, 2-phosphoglycolate is bound at the active site of *Plasmodium falciparum* triose phosphate isomerase. The molecule is held in position by many hydrogen bonds between the phosphate group and surrounding amino acid side chains. Some of the hydrogen bonds are formed through bridged "frozen" water molecules in the active site. The catalytic residues, Glu-165 and His-95, form hydrogen bonds with the carboxylate group of 2-phosphoglycolate as expected in the transition state. [PDB 1LYZ]

Downloaded From Vet-e-Books.com



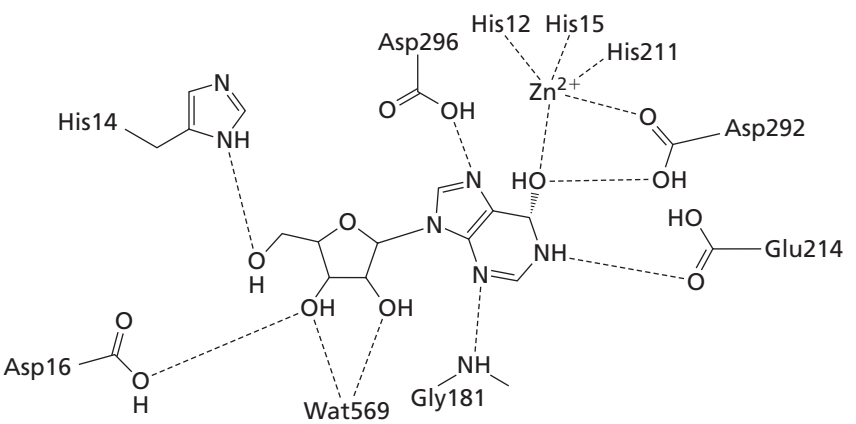
▲ **Figure 6.19**
Inhibition of adenosine deaminase by a transition-state analog. (a) In the deamination of adenosine, a proton is added to N-1 and a hydroxide ion is added to C-6 to form an unstable covalent hydrate, which decomposes to produce inosine and ammonia. **(b)** The inhibitor purine ribonucleoside also rapidly forms a covalent hydrate, 6-hydroxy-1,6-dihydropurine ribonucleoside. This covalent hydrate is a transition-state analog that binds more than a million times more avidly than another competitive inhibitor, 1,6-dihydropurine ribonucleoside **(c)**, which differs from the transition-state analog only by the absence of the 6-hydroxyl group.



▲ **Figure 6.20**
Adenosine deaminase with bound transition-state analog.

deaminase must specifically and avidly bind the transition-state analog—and also the transition state—through interaction with the hydroxyl group at C-6.

The structure of adenosine deaminase with the bound transition-state analog is shown in Figure 6.20 and the interactions between the analog and amino acid side chains in the active site are depicted in Figure 6.21. Notice the hydrogen bonds between Asp-292 and the hydroxyl group on C-6 of 6-hydroxy-1,6-dihydropurine and the interaction between this hydroxyl group and a bound zinc ion in the active site. This confirms the hypothesis that the enzyme specifically binds the transition state in the normal reaction.



▲ **Figure 6.21**
Transition-state analog binding to adenosime deaminase. The interactions between the transition state analog, 6-hydroxy-1,6-dihydropurine, and amino acid side chains in the active site of adenosine deaminase confirms that the enyme recognizes the hydroxyl group at C-6. [PDB 1KRM]

Downloaded From Vet-e-Books.com

6.6 Serine Proteases

Serine proteases are a class of enzymes that cleave the peptide bond of proteins. As the name implies, they are characterized by the presence of a catalytic serine residue in their active sites. The best-studied serine proteases are the related enzymes trypsin, chymotrypsin, and elastase. These enzymes provide an excellent opportunity to explore the relationship between protein structure and catalytic function. They have been intensively studied for 50 years and form an important part of the history of biochemistry and the elucidation of enzyme mechanisms. In this section, we see how the activity of serine proteases is regulated by zymogen activation and examine a structural basis for the substrate specificity of different serine proteases.

A. Zymogens Are Inactive Enzyme Precursors

Mammals digest food in the stomach and intestines. During this process, food proteins undergo a series of hydrolytic reactions as they pass through the digestive tract. Following mechanical disruption by chewing and moistening with saliva, foods are swallowed and mixed with hydrochloric acid in the stomach. The acid denatures proteins and pepsin (a protease that functions optimally in an acidic environment) catalyzes hydrolysis of these denatured proteins to a mixture of peptides. The mixture passes into the intestine where it is neutralized by sodium bicarbonate and digested by the action of several proteases to amino acids and small peptides that can be absorbed into the bloodstream.

Pepsin is initially secreted as an inactive precursor called pepsinogen. When pepsinogen encounters HCl in the stomach it is activated to cleave itself forming the more active protease, pepsin. The stomach secretions are stimulated by food—or even the anticipation of food—as shown by Ivan Pavlov in his experiments with dogs over 100 years ago. (Pavlov was awarded a Nobel Prize in 1904.) The inactive precursor is called a zymogen. Pavlov was the first to show that zymogens could be converted to active proteases in the stomach and intestines.

The main serine proteases are trypsin, chymotrypsin, and elastase. Together, they catalyze much of the digestion of proteins in the intestine. Like pepsin, these enzymes are also synthesized and stored as inactive precursors called zymogens. The zymogens, are called trypsinogen, chymotrypsinogen, and proelastase. They are synthesized in the pancreas. It's important to store these hydrolytic enzymes as inactive precursors within the cell since the active proteases would kill the pancreatic cells by cleaving cytoplasmic proteins.

BOX 6.3 KORNBERG'S TEN COMMANDMENTS

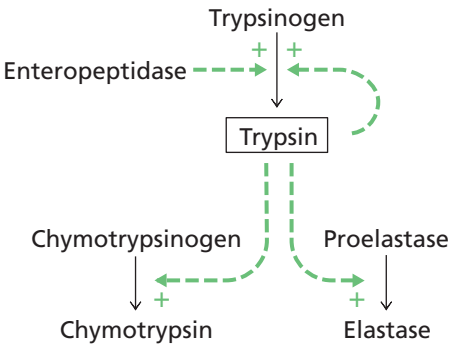
1. Rely on enzymology to clarify biologic questions
2. Trust the universality of biochemistry and the power of microbiology
3. Do not believe something because you can explain it
4. Do not waste clean thinking on dirty enzymes
5. Do not waste clean enzymes on dirty substrates
6. Depend on viruses to open windows
7. Correct for extract dilution with molecular crowding
8. Respect the personality of DNA
9. Use reverse genetics and genomics
10. Employ enzymes as unique reagents

Arthur Kornberg, Nobel Laureate in Physiology or Medicine 1959



Kornberg, A. (2000). Ten commandments: lessons from the enzymology of DNA replication. *J. Bacteriol.* 182:3613–3618.
Kornberg, A. (2003). Ten commandments of enzymology, amended. *Trends Biochem. Sci.* 28:515–517.

Downloaded From Vet-e-Books.com

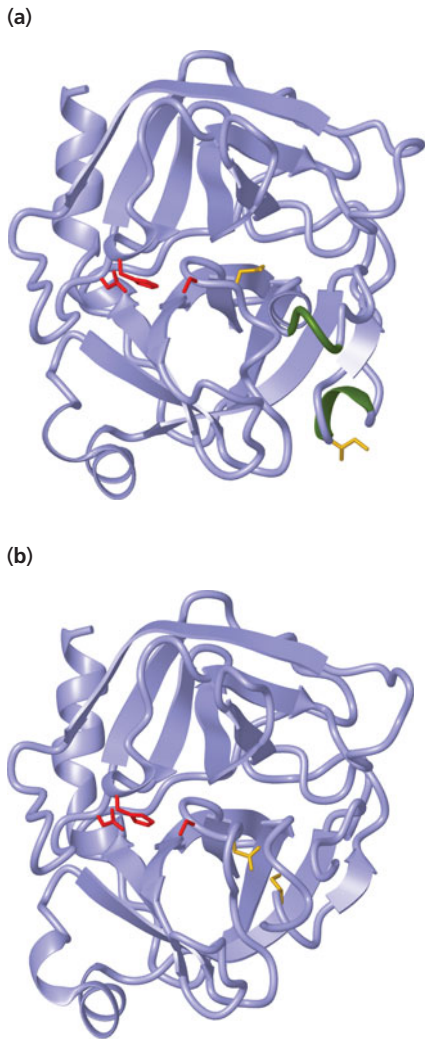


▲ **Figure 6.22**
Activation of some pancreatic zymogens. Initially, enteropeptidase catalyzes the activation of trypsinogen to trypsin. Trypsin then activates chymotrypsinogen, proelastase, and additional trypsinogen molecules.

The enzymes are activated by selective proteolysis—enzymatic cleavage of one or a few specific peptide bonds—when they are secreted from the pancreas into the small intestine. A protease called enteropeptidase specifically activates trypsinogen to trypsin by catalyzing cleavage of the bond between Lys-6 and Ile-7. Once activated by the removal of its N-terminal hexapeptide, trypsin proteolytically cleaves the other pancreatic zymogens, including additional trypsinogen molecules (Figure 6.22).

The activation of chymotrypsinogen to chymotrypsin is catalyzed by trypsin and by chymotrypsin itself. Four peptide bonds (between residues 13 and 14, 15 and 16, 146 and 147, and 148 and 149) are cleaved resulting in the release of two dipeptides. The resulting chymotrypsin retains its three-dimensional shape, despite two breaks in its backbone. This stability is partly due to the presence of five disulfide bonds in the protein.

X-ray crystallography has revealed one major difference between the conformation of chymotrypsinogen and chymotrypsin—the lack of a hydrophobic substrate-binding pocket in the zymogen. The differences are shown in Figure 6.23 where the structures of chymotrypsinogen and chymotrypsin are compared. On zymogen activation, the newly generated α-amino group of Ile-16 turns inward and interacts with the β-carboxyl group of Asp-194 to form an ion pair. This local conformational change generates a relatively hydrophobic substrate-binding pocket near the three catalytic residues with ionizable side chains (Asp-102, His-57, and Ser-195).



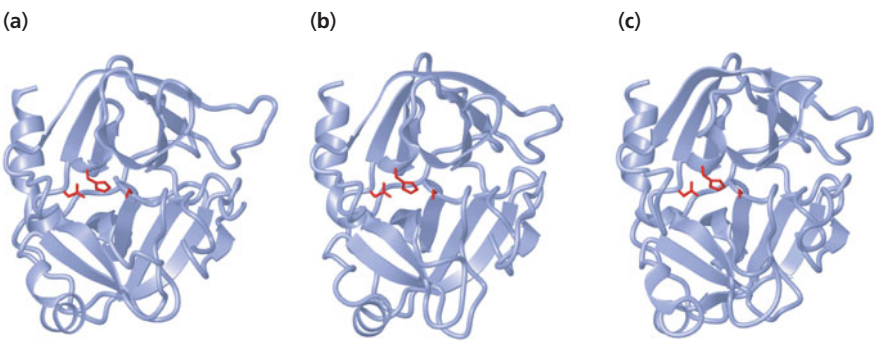
▲ **Figure 6.23**
Polypeptide chains of chymotrypsinogen (left) [PDB 2CGA] and α-chymotrypsin (right) [PDB 5CHA]. Ile-16 and Asp-194 in both zymogen and the active enzyme are shown in yellow. The catalytic-site residues (Asp-102, His-57, and Ser-195) are shown in red. The residues that are removed by processing the zymogen are colored green.

B. Substrate Specificity of Serine Proteases

Chymotrypsin, trypsin, and elastase are similar enzymes that share a common ancestor; in other words, they are homologous. Each enzyme has a two-lobed structure with the active site located in a cleft between the two domains. The positions of the catalytically active side chains of the serine, histidine, and aspartate residues in the active sites are almost identical in the three enzymes (Figure 6.24).

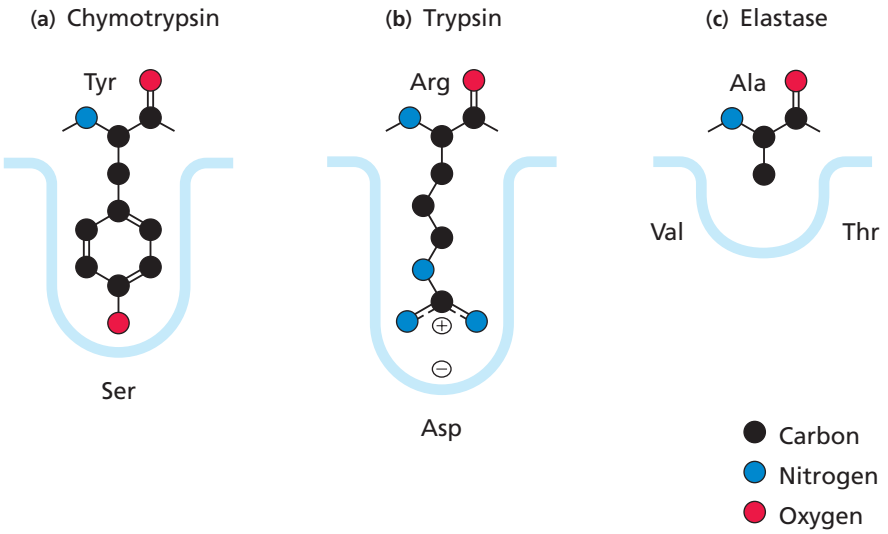
The substrate specificities of chymotrypsin, trypsin, and elastase have been explained by relatively small structural differences in the enzymes. Recall that trypsin catalyzes the hydrolysis of peptide bonds whose carbonyl groups are contributed by arginine or lysine (Section 3.10). Both chymotrypsin and trypsin contain a binding pocket that correctly positions the substrates for nucleophilic attack by an active-site serine residue. Each protease has a similar extended region into which polypeptides fit but the so-called specificity pocket near the active-site serine is markedly different for each enzyme. Trypsin differs from chymotrypsin because in chymotrypsin there is an uncharged serine residue at the base of the hydrophobic binding pocket. In trypsin this residue is an aspartate residue (Figure 6.25). This negatively charged aspartate residue forms an ion pair with the positively charged side chain of an arginine or lysine residue of the substrate in the ES complex. Experiments with specifically mutated trypsin indicate that the aspartate residue at the base of its specificity pocket is a major factor in substrate specificity but other parts of the molecule also affect specificity.

Elastase catalyzes the degradation of elastin, a fibrous protein that is rich in glycine and alanine residues. Elastase is similar in tertiary structure to chymotrypsin except that



▲ **Figure 6.24**
Serine proteases. Comparison of the polypeptide backbones of (a) chymotrypsin [PDB 5CHA], (b) trypsin [PDB 1TLD], and (c) elastase [PDB 3EST]. Residues at the catalytic center are shown in red.

Downloaded From Vet-eBooks.com



◀ **Figure 6.25**
Binding sites of chymotrypsin, trypsin, and elastase. The differing binding sites of these three serine proteases are primary determinants of their substrate specificities. **(a)** Chymotrypsin has a hydrophobic pocket that binds the side chains of aromatic or bulky hydrophobic amino acid residues. **(b)** A negatively charged aspartate residue at the bottom of the binding pocket of trypsin allows trypsin to bind the positively charged side chains of lysine and arginine residues. **(c)** In elastase, the side chains of a valine and a threonine residue at the binding site create a shallow binding pocket. Elastase binds only amino acid residues with small side chains, especially glycine and alanine residues.

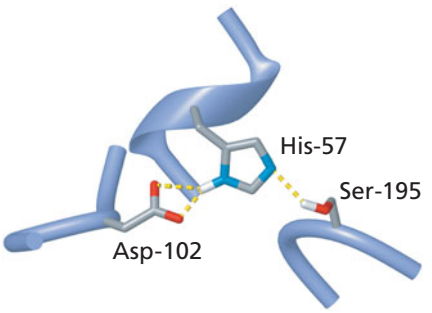
its binding pocket is much shallower. Two glycine residues found at the entrance of the binding site of chymotrypsin and trypsin are replaced in elastase by much larger valine and threonine residues (Figure 6.25c). These residues keep potential substrates with large side chains away from the catalytic center. Thus, elastase specifically cleaves proteins that have small residues such as glycine and alanine.

C. Serine Proteases Use Both the Chemical and the Binding Modes of Catalysis

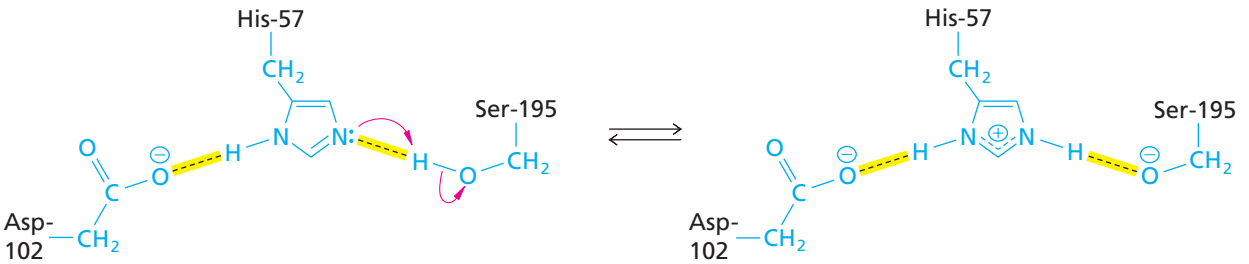
Let's examine the mechanism of chymotrypsin and the roles of three catalytic residues: His-57, Asp-102, and Ser-195. Many enzymes catalyze the cleavage of amide or ester bonds by the same process so study of the chymotrypsin mechanism can be applied to a large family of hydrolases.

Asp-102 is buried in a rather hydrophobic environment. It is hydrogen-bonded to His-57 that in turn is hydrogen-bonded to Ser-195 (Figure 6.26). This group of amino acid residues is called the *catalytic triad.* The reaction cycle begins when His-57 abstracts a proton from Ser-195 (Figure 6.27). This creates a powerful nucleophile (Ser-195) that will eventually attack the peptide bond. Initiation of this part of the reaction is favored because Asp-102 stabilizes the histidine promoting its ability to deprotonate the serine residue.

The discovery that Ser-195 is a catalytic residue of chymotrypsin was surprising because the side chain of serine is usually not sufficiently acidic to undergo deprotonation in order to serve as a strong nucleophile. The hydroxymethyl group of a serine residue generally has a pK_a of about 16 and is similar in reactivity to the hydroxyl group of ethanol. You may recall from organic chemistry that although ethanol can ionize to



▲ **Figure 6.26**
The catalytic site of chymotrypsin. The active-site residues Asp-102, His-57, and Ser-195 are arrayed in a hydrogen-bonded network. The conformation of these three residues is stabilized by a hydrogen bond between the carbonyl oxygen of the carboxylate side chain of Asp-102 and the peptide-bond nitrogen of His-57. Oxygen atoms of the active-site residues are red, and nitrogen atoms are dark blue. [PDB 5CHA].



▲ **Figure 6.27**
Catalytic triad of chymotrypsin. The imidazole ring of His-57 removes the proton from the hydroxymethyl side chain of Ser-195 (to which it is hydrogen-bonded), thereby making Ser-195 a powerful nucleophile. This interaction is facilitated by interaction of the imidazolium ion with its other hydrogen-bonded partner, the buried β-carboxylate group of Asp-102. The residues of the triad are drawn in an arrangement similar to that shown in Figure 6.24.

Downloaded From Vet-e-Books.com

BOX 6.4 CLEAN CLOTHES

It's a little-known fact that 75% of all laundry detergents contain proteases that are used in helping to remove stubborn protein-based stains from dirty clothes.

All protease additives are based on serine proteases isolated from various *Bacillus* species. These enzymes have been extensively modified in order to be active under the harsh conditions of a detergent solution at high temperature. A successful example of site-directed mutagenesis is the alteration of the serine protease subtilisin from *Bacillus subtilis* (Box 6.4) to make it more resistant to chemical oxidation. It has a methionine residue in the active-site cleft (Met-222) that is readily oxidized leading to inactivation of the enzyme. Resistance to oxidation increases the suitability of subtilisin as a detergent additive. Met-222 was systematically replaced by each of the other common amino acids in a series of mutagenic experiments. All 19 possible mutant subtilisins were isolated and tested and most had greatly diminished peptidase activity. The Cys-222 mutant had high activity but was also subject to oxidation. The Ala-222 and Ser-222 mutants, with nonoxidizable side chains, were not inactivated by oxidation and had relatively high activity. They were the only active, oxygen-stable mutant subtilisin variants.

Site-directed mutagenesis has been performed to alter eight of the 319 amino acid residues of a bacterial protease. The wild-type protease is moderately stable when heated but the suitably mutated enzyme is stable and can function at 100°C. Its denaturation in detergent is prevented by groups, such as a disulfide bridge, that stabilize its conformation.

Recently there has been a trend to lower wash temperatures in order to save energy. The older group of enzymes are not effective at lower wash temperatures so a whole new round of bioengineering has begun creating modified enzymes that can be effective in a modern energy-conscious household.



form an ethoxide this reaction requires the presence of an extremely strong base or treatment with an alkali metal. We see below how the active site of chymotrypsin, achieves this ionization in the presence of a substrate.

A proposed mechanism for chymotrypsin and related serine proteases includes covalent catalysis (by a nucleophilic oxygen) and general acid–base catalysis (donation of a proton to form a leaving group). The steps of the proposed mechanism are illustrated in Figure 6.28.

Binding of the peptide substrate causes a slight conformation change in chymotrypsin, sterically compressing Asp-102 and His-57. A low-barrier hydrogen bond is formed between these side chains and the pK_a of His-57 rises from about 7 to about 11. (Formation of this strong, almost covalent, bond drives electrons toward the second N atom of the imidazole ring of His-57 making it more basic.) This increase in basicity makes His-57 an effective general base for abstracting a proton from the $-CH_2OH$ of Ser-195. This mechanism explains how the normally unreactive alcohol group of serine becomes a potent nucleophile.

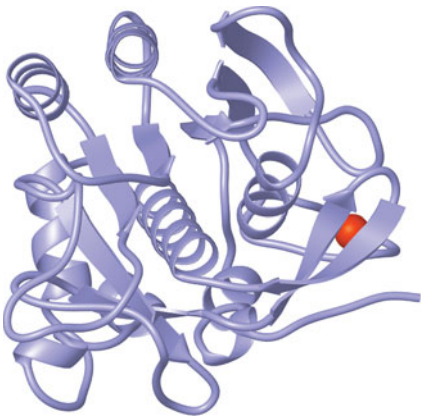
All the catalytic modes described in this chapter are used in the mechanisms of serine proteases. In the reaction scheme shown in Figure 6.28, steps 1 and 4 in the forward direction use the proximity effect, the gathering of reactants. For example, when a water molecule replaces the amine (P_1) in step 4, it is held by histidine, providing a proximity effect. Acid–base catalysis by histidine lowers the energy barriers for steps 2 and 4. Covalent catalysis using the $-CH_2OH$ of serine occurs in steps 2 through 5. The unstable tetrahedral intermediates at steps 2 and 4 (E-TI_1 and E-TI_2) are believed to be similar to the transition states for these steps. Hydrogen bonds in the oxyanion hole stabilize these intermediates, which are oxyanion forms of the substrate, by binding them more tightly to the enzyme than the substrate was bound. The chemical modes of catalysis (acid–base and covalent catalysis) and the binding modes of catalysis (the proximity effect and transition-state stabilization) all contribute to the enzymatic activity of serine proteases.

Downloaded From Vet-e-Books.com

BOX 6.5 CONVERGENT EVOLUTION

The protease subtilisin from the bacterium *Bacillus subtilis* is another example of a serine protease. It possesses a catalytic triad consisting of Asp-32, His-64, and Ser-221 at its active site. These are arranged in an alignment similar to the Asp-102, His-57, and Ser-195 residues in chymotrypsin (Figure 6.27). However, as you might deduce from the residue numbers, the structures of subtilisin and chymotrypsin are very different and there is no significant sequence similarity.

This is a remarkable example of convergent evolution. The mammalian intestinal serine proteases and the bacterial subtilisins have independently discovered the catalytic Asp-His-Ser triad.



► **Subtilisin from *Bacillus subtilis*.** The structure of this enzyme is very different from that of serine proteases shown in Figure 6.24. [PDB 1SBC]

6.7 Lysozyme

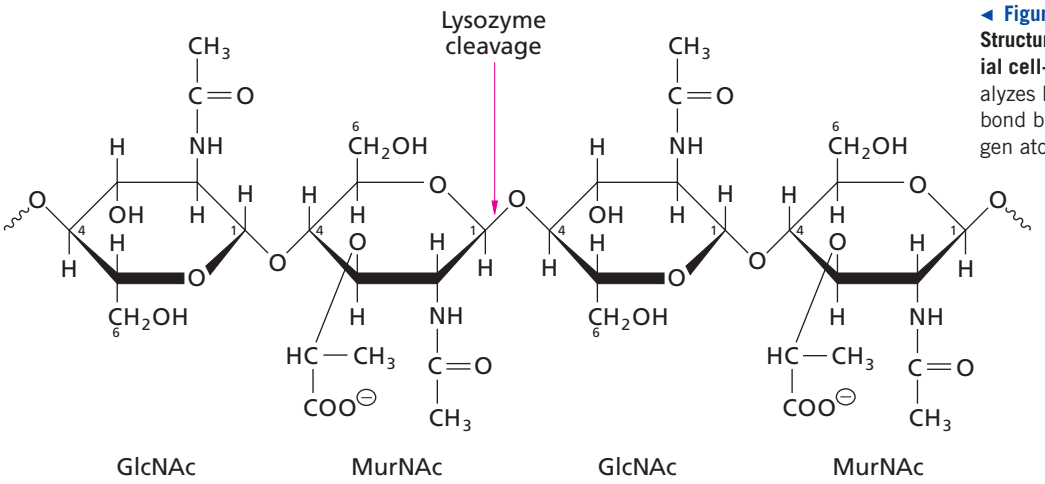
Lysozyme catalyzes the hydrolysis of some polysaccharides, especially those that make up the cell walls of bacteria. It is the first enzyme whose structure was solved and for this reason there has been a long-term interest in working out its precise mechanism of action. Many secretions, such as tears, saliva, and nasal mucus, contain lysozyme activity to help prevent bacterial infection. (Lysozyme causes *lysis*, or disruption, of bacterial cells.) The best-studied lysozyme is from chicken egg white.

The substrate of lysozyme is a polysaccharide composed of alternating residues of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) connected by glycosidic bonds (Figure 6.29). Lysozyme specifically catalyzes hydrolysis of the glycosidic bond between C-1 of a MurNAc residue and the oxygen atom at C-4 of a GlcNAc residue.

The structure of bacterial cell walls is described in Seciton 8.7B.

Models of lysozyme and its complexes with saccharides have been obtained by X-ray crystallographic analysis (Figure 6.30). The substrate-binding cleft of lysozyme accommodates six saccharide residues. Each of the residues binds to a particular part of the active cleft at sites A through E.

Sugar molecules fit easily into all but one site of the structural model. At site D a sugar molecule such as MurNAc does not fit into the model unless it is distorted into a



◄ **Figure 6.29**
Structure of a four-residue portion of a bacterial cell-wall polysaccharide. Lysozyme catalyzes hydrolytic cleavage of the glycosidic bond between C-1 of MurNAc and the oxygen atom involved in the glycosidic bond.

Downloaded From Vet-e-Books.com

The noncovalent enzyme-substrate complex is formed, orienting the substrate for reaction. Interactions holding the substrate in place include binding of the R_1 group in the specificity pocket (shaded). The binding interactions position the carbonyl carbon of the scissile peptide bond (the bond susceptible to cleavage) next to the oxygen of Ser-195.

E + S

Binding of the substrate compresses Asp-102 and His-57. This strain is relieved by formation of a low-barrier hydrogen bond. The raised pK_a of His-57 enables the imidazole ring to remove a proton from the hydroxyl group of Ser-195. The nucleophilic oxygen of Ser-195 attacks the carbonyl carbon of the peptide bond to form a tetrahedral intermediate (E-TI_1), which is believed to resemble the transition state.

E-S

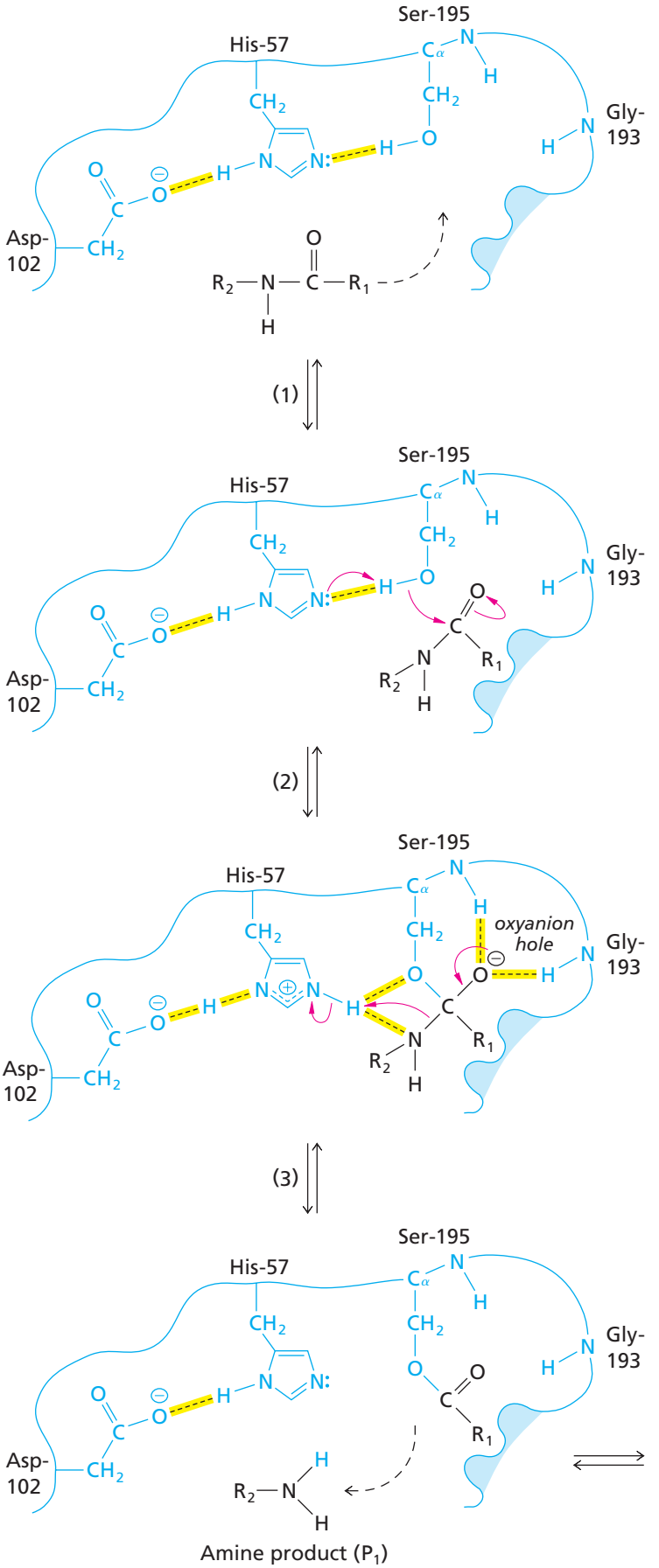
When the tetrahedral intermediate is formed, the substrate C—O bond changes from a double bond to a longer single bond. This allows the negatively charged oxygen (the oxyanion) of the tetrahedral intermediate to move to a previously vacant position, called the oxyanion hole, where it can form hydrogen bonds with the peptide-chain —NH groups of Gly-193 and Ser-195.

The imidazolium ring of His-57 acts as an acid catalyst, donating a proton to the nitrogen of the scissile peptide bond, thus facilitating its cleavage.

E-TI_1

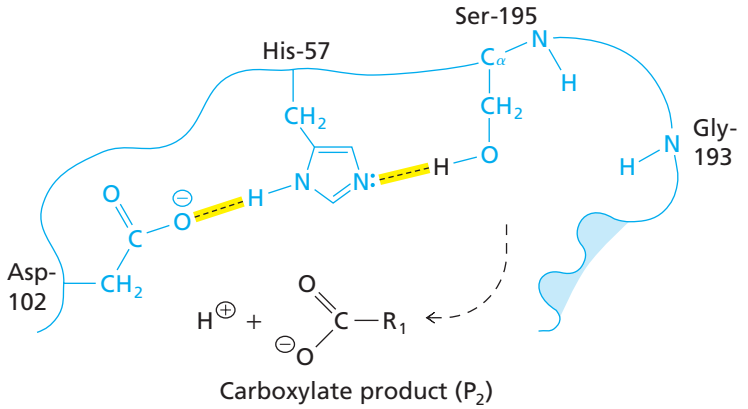
The carbonyl group from the peptide forms a covalent bond with the enzyme, producing an acyl-enzyme intermediate. After the peptide product (P_1) with the new amino terminus leaves the active site, water enters.

Acyl E + P_1



▲ **Figure 6.28**
Mechanism of chymotrypsin-catalyzed cleavage of a peptide bond.

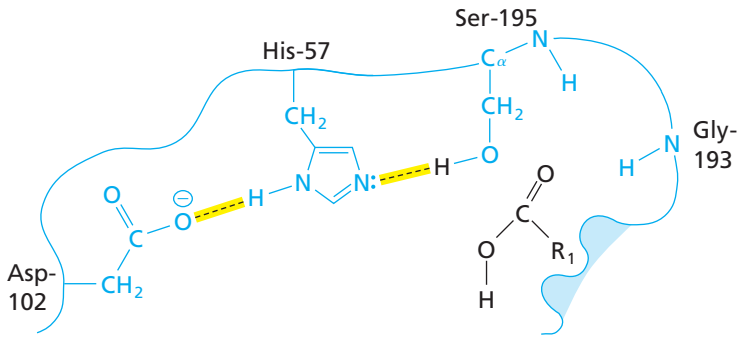
Downloaded From Vet-e-Books.com



$E + P_2$

The carboxylate product is released from the active site, and free chymotrypsin is regenerated.

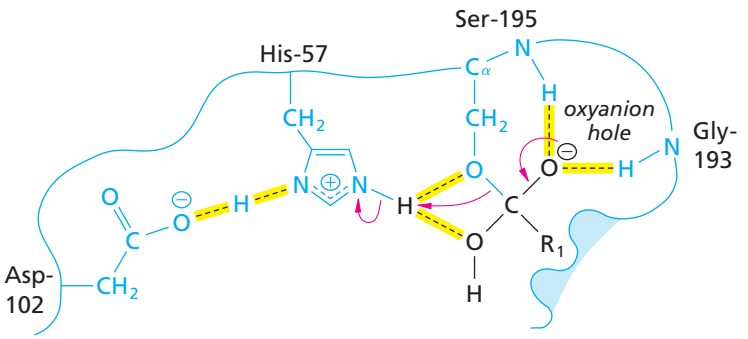
(6)



$E\text{-}P_2$

The second product (P_2)—a polypeptide with a new carboxy terminus—is formed.

(5)

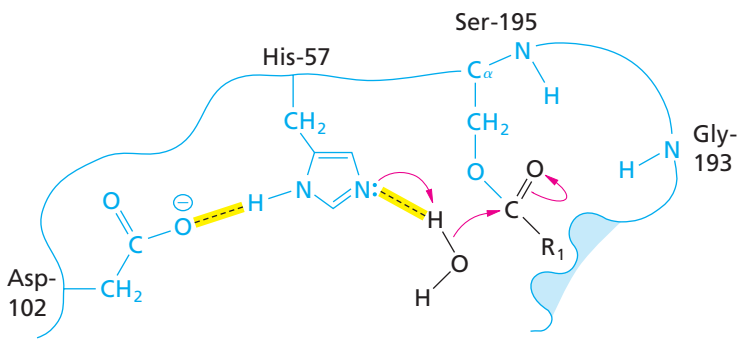


$E\text{-}TI_2$

His-57, once again an imidazolium ion, donates a proton, leading to the collapse of the second tetrahedral intermediate.

A second tetrahedral intermediate ($E\text{-}TI_2$) is formed and stabilized by the oxyanion hole.

(4)

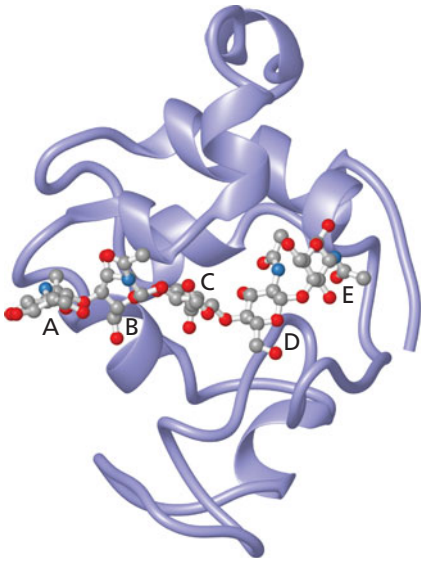


Acyl E + H_2O

Hydrolysis (deacylation) of the acyl-enzyme intermediate starts when Asp-102 and His-57 again form a low-barrier hydrogen bond and His-57 removes a proton from the water molecule to provide an $OH^{\ominus}$ group to attack the carbonyl group of the ester.

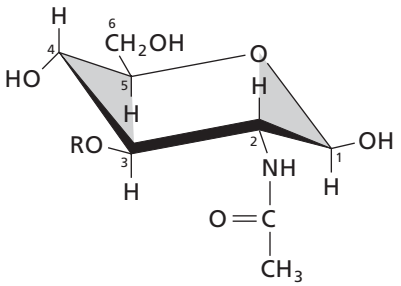
▲ **Figure 6.28** (***continued***)

Downloaded From Vet-e-Books.com

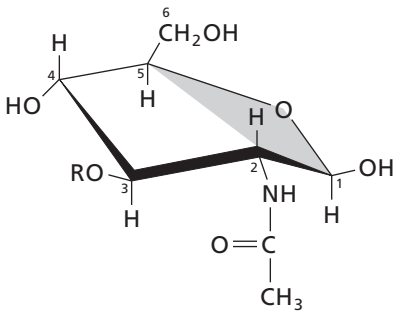


▲ **Figure 6.30**
Lysozyme from chicken with a pentasaccharide molecule (pink). The ligand is bound in sites A, B, C, D and E. Site F is not occupied in this structure. The active site for bond cleavage is between sites D and E. [PDB 1SFB].

(a) Chair conformation



(b) Half-chair conformation



▲ **Figure 6.31**
Conformations of *N*-acetylmuramic acid. **(a)** Chair conformation. **(b)** Half-chair conformation proposed for the sugar bound in site D of lysozyme. R represents the lactyl group of MurNAc.

half-chair conformation (Figure 6.31). Two ionic amino acid residues, Glu-35 and Asp-52, are located close to C-1 of the distorted sugar molecule in the D binding site. Glu-35 is in a nonpolar region of the cleft and has a perturbed pK_a near 6.5. Asp-52, in a more polar environment, has a pK_a near 3.5. The pH optimum of lysozyme is near 5—between these two pK_a values. Recall that the pK_a value of individual amino acid side chains may not be the same as the pK_a value of the free amino acid in solution (Section 3.4).

The proposed mechanism of lysozyme is shown in Figure 6.32. When a molecule of polysaccharide binds to lysozyme, MurNAc residues bind to sites B, D, and F (there is no cavity for the lactyl side chain of MurNAc in site A, C, or E). The extensive binding of the oligosaccharide forces the MurNAc residue in the D site into the half-chair conformation. A near covalent bond forms between Asp-52 and the postulated intermediate (an unstable oxocarbocation). Recent evidence suggests that this interaction might be more like a covalent bond than a strong ion pair but there is much controversy over this point. It's interesting that there are still details of the lysozyme mechanism to be worked out after almost 50 years of effort.

Lysozyme is only one representative of a large group of glycoside hydrolases. Recently, the structures of a bacterial cellulase and its complexes with substrate, intermediate, and product have been determined. This glycosidase has a slightly different mechanism than lysozyme—it forms a covalent glycosyl–enzyme intermediate rather than the strong ion pair postulated for lysozyme. Other aspects of its mechanism, such as distortion of a sugar residue and interaction with active-site —COOH and —$COO^{\ominus}$ side chains, resemble those of the lysozyme mechanism. The structures of the enzyme complexes show that distortion of the substrate forces it toward the transition state.

6.8 Arginine Kinase

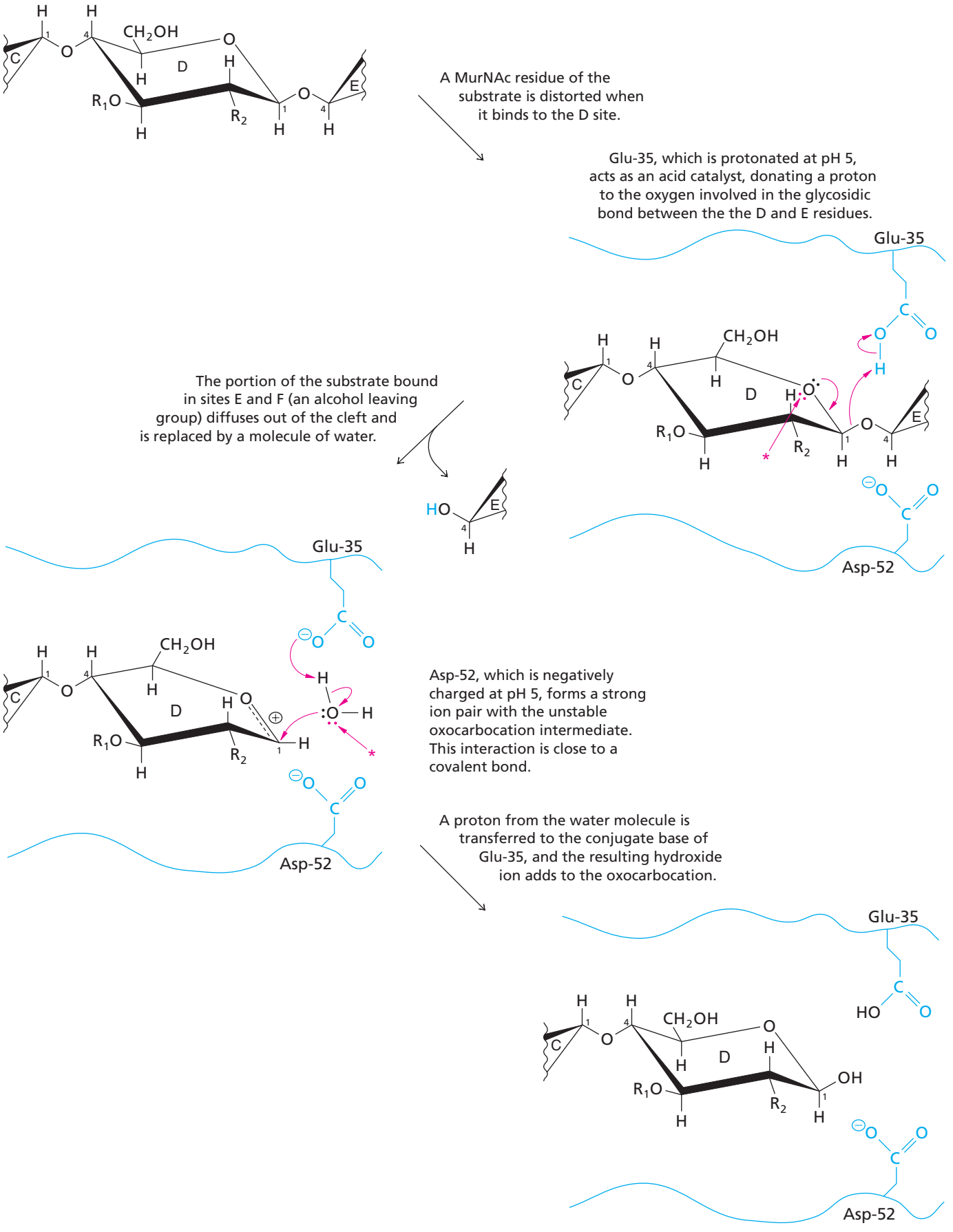
Most enzymatic reactions for which detailed mechanisms have been elucidated involve fairly simple reactions, such as isomerizations, cleavage reactions, or reactions with water as the second reactant. Therefore, in order to assess proximity effects and the extent of transition state stabilization, it's worthwhile looking at a more complicated reaction, such as that catalyzed by arginine kinase:

$$\text{Arginine} + \text{MgATP} \rightleftharpoons \text{Arginine Phosphate} + \text{MgADP} + H^{\oplus}$$

The structure of a transition-state analog–enzyme complex of arginine kinase has been determined at high resolution (Figure 6.33). However, rather than studying the usual type of transition-state analog in which reactants are fused by covalent bonds, the scientists used three separate components: arginine, nitrate (to model the phosphoryl group transferred between arginine and ADP), and ADP. X-ray crystallographic examination of the active site containing these three compounds led to the proposal of a structure for the transition state and a mechanism for the reaction (see Figure 6.33). The crystallographic results showed that the enzyme has greatly restricted the movement of the bound species (and presumably also of the transition state). For example, the terminal pyrophosphoryl group of ATP is held in place by four arginine side chains and a bound Mg^{2+} ion and the guanidinium group of the arginine substrate molecule is held firmly by two glutamate side chains. The components are precisely and properly aligned by the enzyme.

Arginine kinase, like other kinases, is an induced-fit enzyme (Section 6.5C). It assumes the closed shape when it is crystallized in the presence of arginine, nitrate, and ADP. This enzyme has a k_{cat} of about 2×10^2 s^{-1} and K_m values above 10^{-4} M for both arginine and ATP—values that are quite typical for kinases. The movement that occurs during the induced-fit binding of substrates has precisely aligned the substrates, which had previously been bound fairly weakly, as shown by their moderate K_m values. At least four interrelated catalytic effects participate in this enzymatic reaction: proximity

Downloaded From Vet-e-Books.com



▲ **Figure 6.32**
Mechanism of lysozyme. R_1 represents the lactyl group, and R_2 represents the *N*-acetyl group of MurNAc.

Downloaded From Vet-e-Books.com

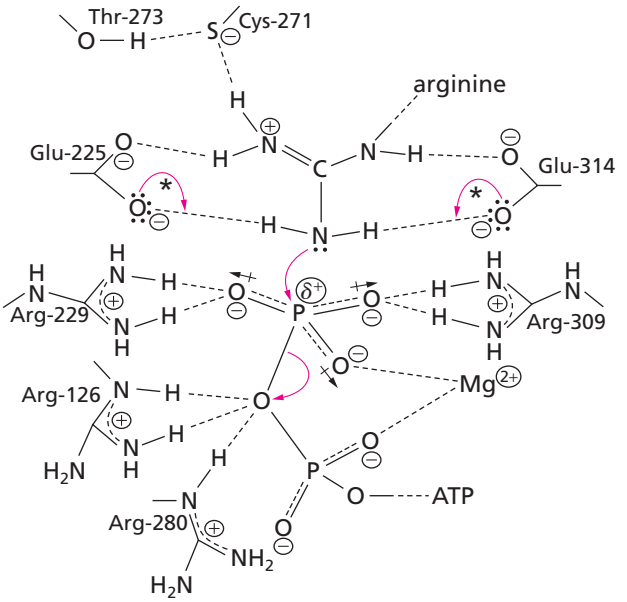


Figure 6.33 ►
Proposed structure of the active site of arginine kinase in the presence of ATP and arginine. The substrate molecules are held firmly and aligned toward the transition state, as shown by the dashed lines. The asterisks (*) show that either Glu-225 or Glu-314 could act as a general acid–base catalyst.

{Adapted from Zhov, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. S. (1998). Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. Proc. Natl. Acad. Sci. USA. 95:8453.)

(collection and alignment of substrate molecules), fairly weak initial binding of substrates, acid–base catalysis, and transition-state stabilization (strain of substrates toward the shape of the transition state).

Having gained insight into the general mechanisms of enzymes, we can now go on to examine reactions that include coenzymes. These reactions require groups not supplied by the side chains of amino acids.

Summary

1. The four major modes of enzymatic catalysis are acid–base catalysis and covalent catalysis (chemical modes) and proximity and transition-state stabilization (binding modes). The atomic details of reactions are described by reaction mechanisms, which are based on the analysis of kinetic experiments and protein structures.
2. For each step in a reaction, the reactants pass through a transition state. The energy difference between stable reactants and the transition state is the activation energy. Catalysts allow faster reactions by lowering the activation energy.
3. Ionizable amino acid residues in active sites form catalytic centers. These residues may participate in acid–base catalysis (proton addition or removal) or covalent catalysis (covalent attachment of a portion of the substrate to the enzyme). The effects of pH on the rate of an enzymatic reaction can suggest which residues participate in catalysis.
4. The catalytic rates for a few enzymes are so high that they approach the upper physical limit of reactions in solution, the rate at which reactants approach each other by diffusion.
5. Most of the rate acceleration achieved by an enzyme arises from the binding of substrates to the enzyme.
6. The proximity effect is acceleration of the reaction rate due to the formation of a noncovalent ES complex that collects and orients reactants resulting in a decrease in entropy.
7. An enzyme binds its substrates fairly weakly. Excessively strong binding would stabilize the ES complex and slow the reaction.
8. An enzyme binds a transition state with greater affinity than it binds substrates. Evidence for transition state stabilization is provided by transition-state analogs that are enzyme inhibitors.
9. Some enzymes use induced fit (substrate-induced activation that involves a conformation change) to prevent wasteful hydrolysis of a reactive substrate.
10. Many serine proteases are synthesized as inactive zymogens that are activated extracellularly under appropriate conditions by selective proteolysis. The examination of serine proteases by X-ray crystallography shows how the three-dimensional structures of proteins can reveal information about the active sites, including the binding of specific substrates.
11. The active sites of serine proteases contain a hydrogen-bonded Ser–His–Asp catalytic triad. The serine residue serves as a covalent catalyst, and the histidine residue serves as an acid–base catalyst. Anionic tetrahedral intermediates are stabilized by hydrogen bonds with the enzyme.
12. The proposed mechanism for lysozyme, an enzyme that catalyzes the hydrolysis of bacterial cell walls, includes substrate distortion and stabilization of an unstable oxocarbocation intermediate.

Downloaded From Vet-e-Books.com

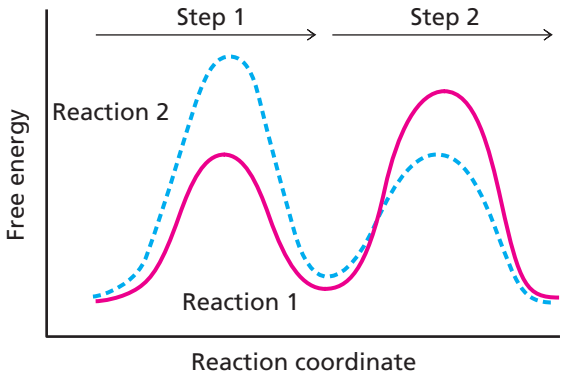
Problems

1. (a) What forces are involved in binding substrates and intermediates to the active sites of enzymes?

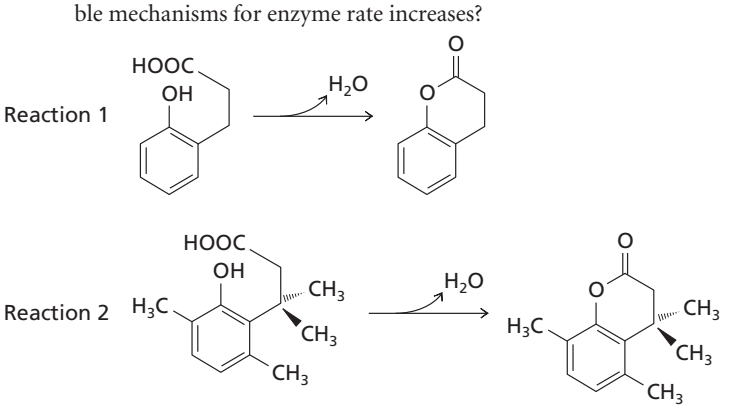
(b) Explain why very tight binding of a substrate to an enzyme is not desirable for enzyme catalysis, whereas tight binding of the transition state is desirable.

2. The enzyme orotodine 5-phosphate decarboxylase is one of the most proficient enzymes known, accelerating the rate of decarboxylation of orotidine 5′ monophosphate by a factor of 10^{23} (Section 5.4). Nitrogen-15 isotope effect studies have shown that two major participating mechanisms are (1) destabilization of the ground state ES complex by electrostatic repulsion between the enzyme and substrate, and (2) stabilization of the transition state by favorable electrostatic interactions between the enzyme and ES‡. Draw an energy diagram that shows how these two effects promote catalysis.

3. The energy diagrams for two multistep reactions are shown below. What is the rate-determining step in each of these reactions?



4. Reaction 2 below occurs 2.5×10^{11} times faster than Reaction 1. What is likely to be a major reason for this enormous rate increase in Reaction 2? How is this model relevant for interpreting possible mechanisms for enzyme rate increases?



5. List three major catalytic effects for lysozyme and explain how each is used during the enzyme-catalyzed hydrolysis of a glycosidic bond.

6. There are multiple serine residues in α-chymotrypsin but only serine 195 reacts rapidly when the enzyme is treated with active phosphate inhibitors such as diisopropyl fluorophosphate (DFP). Explain.

7. (a) Identify the residues in the catalytic triad of α-chymotrypsin and indicate the type of catalysis mediated by each residue.

(b) What additional amino acid groups are found in the oxyanion hole and what role do they play in catalysis?

(c) Explain why site-directed mutagenesis of aspartate to asparagine in the active site of trypsin decreases the catalytic activity 10,000-fold.

8. Catalytic triad groupings of amino acid residues increase the nucleophilic character of active-site serine, threonine, or cysteine residues present in many enzymes involved in catalyzing the cleavage of substrate amide or ester bonds. Using α-chymotrypsin as a model system, diagram the expected arrangements of the catalytic triads in the enzymes below.

(a) Human cytomegalovirus protease: His, His, Ser

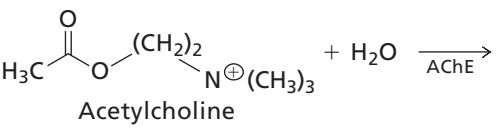
(b) β-lactamase: Glu, Lys, Ser

(c) Asparaginase: Asp, Lys, Thr

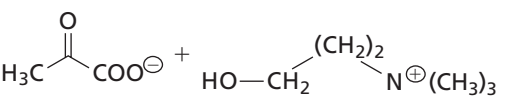
(d) Hepatitis A protease: Asp, (H_2O), His, Cys (a water molecule is situated between the Asp and His residues)

9. Human dipeptidyl peptidase IV (DDP-IV) is a serine protease that catalyzes hydrolysis of prolyl peptide bonds at the next-to-last position at the N terminus of a protein. Many physiological peptides have been identified as substrates, including proteins involved in the regulation of glucose metabolism. DDP-IV contains a catalytic triad at the active site (Glu-His-Ser) and a tyrosine residue in the oxyanion hole. Site-directed mutagenesis of this tyrosine residue in DPP-IV was performed, and the ability of the enzyme to cleave a peptide substrate was compared to that of the wild-type enzyme. The tyrosine residue found in the oxyanion hole was changed to a phenylalanine. The phenylalanine mutant had less than 1% of the activity of the wild-type enzyme (Bjelke, J. R., Christensen, J., Branner, S., Wagtmann, N., Olsen, C. Kanstrup, A. B., and Rasmussen, H. B. (2004). Tyrosine 547 constitutes an essential part of the catalytic mechanism of dipeptidyl peptidase IV. *J. Biol. Chem.* 279:34691–34697). Is this tyrosine required for activity of DDP-IV? Why does the replacement of a tyrosine with a phenylalanine abolish the enzyme activity?

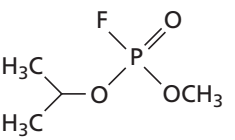
10. Acetylcholinesterase (AChE) catalyzes the breakdown of the neurotransmitter acetylcholine to acetate and choline. This enzyme contains a catalytic triad with the residues His, Glu, and Ser. The catalytic triad enhances the nucleophilicity of the serine residue. The nucleophilic oxygen of serine attacks the carbonyl carbon of acetylcholine to form a tetrahedral intermediate.



Acetylcholine



The nerve agent sarin is an extremely potent inactivator of AChE. Sarin is an irreversible inhibitor that covalently modifies the serine residue in the active site of AChE.



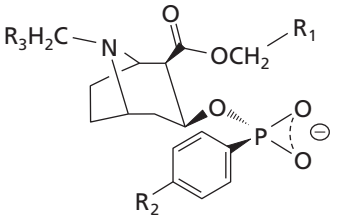
Sarin

(a) Diagram the expected arrangement of the amino acids in the catalytic triad.

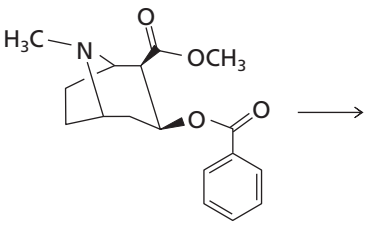
(b) Propose a mechanism for the covalent modification of AChE by sarin.

Downloaded From Vet-e-Books.com

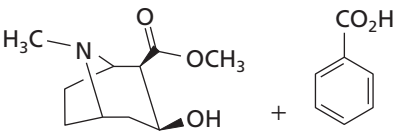
11. Catalytic antibodies are potential therapeutic agents for drug overdose and addiction. For example, a catalytic antibody that catalyzes the breakdown of cocaine before it reached the brain would be an effective detoxification treatment for drug abuse and addiction. The phosphonate analog below was used to raise an anticocaine antibody that catalyzes the rapid hydrolysis of cocaine. Explain why this phosphonate ester was chosen to produce a catalytic antibody.



Phosphonate analog



(−) - Cocaine



Ecgonine methyl ester

Benzoic acid

12. In the chronic lung disease emphysema, the lung's air sacs (alveoli), where oxygen from the air is exchanged for carbon dioxide in the blood, degenerate. α1-Proteinase inhibitor deficiency is a genetic condition that runs in certain families and results from mutations in critical amino acids in the sequence of α1-proteinase inhibitor. The individuals with mutations are more likely to develop emphysema. α1-Proteinase inhibitor is produced by the liver and then circulates in the blood. α1-Proteinase inhibitor is a protein that serves as the major inhibitor of neutrophil elastase, a serine protease present in the lung. Neutrophil elastase cleaves the protein elastin, which is an important component for lung function. The increased rate of elastin breakdown in lung tissue is believed to cause emphysema. One treatment for α1-proteinase inhibitor deficiency is to give the patient human wild-type α1-proteinase inhibitor (derived from large pools of human plasma) intravenously by injecting the protein directly into the bloodstream.

(a) Explain the rational for the treatment with wild-type α1-proteinase inhibitor.

(b) This treatment involves the intravenous administration of the wild-type α1-proteinase inhibitor. Explain why α1-proteinase inhibitor cannot be taken orally.

Selected Readings

General

Fersht, A. (1985). *Enzyme Structure and Mechanism*, 2nd ed. (New York: W. H. Freeman).

Binding and Catalysis

Bartlett, G. J., Porter, C. T., Borkakoti, N. and Thornton, J. M. (2002). Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* 324:105–121.

Bruice, T. C. and Pandrit, U. K. (1960). Intramolecular models depicting the kinetic importance of "fit" in enzymatic catalysis. *Proc. Natl. Acad. Sci. USA.* 46:402–404.

Hackney, D. D. (1990). Binding energy and catalysis. In *The Enzymes*, Vol. 19, 3rd ed., D. S. Sigman and P. D. Boyer, eds. (San Diego: Academic Press), pp. 1–36.

Jencks, W. P. (1987). Economics of enzyme catalysis. *Cold Spring Harbor Symp. Quant. Biol.* 52:65–73.

Kraut, J. (1988). How do enzymes work? *Science* 242:533–540.

Neet, K. E. (1998). Enzyme catalytic power mini-review series. *J. Biol. Chem.* 273:25527–25528, and related papers on pages 25529–25532, 26257–26260, and 27035–27038.

Pauling, L. (1948) Nature of forces between large molecules of biological interest. *Nature* 161:707–709.

Schiøtt, B., Iversen, B. B., Madsen, G. K. H., Larsen, F. K., and Bruice, T. C. (1998). On the electronic nature of low-barrier hydrogen bonds in enzymatic reactions. *Proc. Natl. Acad. Sci. USA* 95:12799–12802.

Shan, S.-U., and Herschlag, D. (1996). The change in hydrogen bond strength accompanying charge rearrangement: implications for enzymatic catalysis. *Proc. Natl. Acad. Sci. USA* 93:14474–14479.

Transition-State Analogs

Schramm, V. L. (1998). Enzymatic transition states and transition state analog design. *Annu. Rev. Biochem.* 67:693–720.

Wolfenden, R., and Radzicka, A. (1991). Transition-state analogues. *Curr. Opin. Struct. Biol.* 1:780–787.

Specific Enzymes

Cassidy, C. S., Lin, J., and Frey, P. A. (1997). A new concept for the mechanism of action of chymotypsin: the role of the low-barrier hydrogen bond. *Biochem.* 36:4576–4584.

Blacklow, S. C., Raines, R. T., Lim, W. A., Zamore, P. D., and Lnowles, J. R. (1988). Triosephosphate isomerase catalysis is diffusion controlled. *Biochem.* 27:1158–1167.

Downloaded From Vet-e-Books.com

Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M., and Withers, S. G. (1998). Snapshots along an enzymatic reaction coordinate: analysis of a retaining β-glycoside hydrolase. *Biochem.* 37:11707–11713.

Dodson, G., and Wlodawer, A. (1998). Catalytic triads and their relatives. *Trends Biochem. Sci.* 23:347–352.

Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994). A low-barrier hydrogen bond in the catalytic triad of serine proteases. *Science.* 264:1927–1930.

Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., and Hallewell, R. A. (1992). Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature.* 358:347–351.

Harris, T. K., Abeygunawardana, C., and Mildvan, A. S. (1997). NMR studies of the role of hydrogen bonding in the mechanism of triosephosphate isomerase. *Biochem.* 36:14661–14675.

Huber, R., and Bode, W. (1978). Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* 11:114–122.

Kinoshita, T., Nishio, N., Nakanishi, I., Sato, A., and Fujii, T. (2003). Structure of bovine adenosine deaminase complexed with 6-hydroxy-1,6-dihydropurine riboside. *Acta Cryst.* D59:299–303.

Kirby, A. J. (2001). The lysozyme mechanism sorted—after 50 years. *Nature Struct. Biol.* 8:737–739.

Knolwes, J. R. (1991) Enzyme catalysis: not different, just better. *Nature.* 350:121–124.

Knowles, J. R., and Albery, W. J. (1977). Perfection in enzyme catalysis: the energetics of triosephosphate isomerase. *Acc. Chem. Res.* 10:105–111.

Kuser, P., Cupri, F., Bleicher, L., and Polikarpov, I. (2008). Crystal structure of yeast hexokinase P1 in complex with glucose: a classical "induced fit" example revisited. *Proteins.* 72:731–740.

Lin, J., Cassidy, C. S., and Frey, P. A. (1998). Correlations of the basicity of His-57 with transition state analogue binding, substrate reactivity, and the strength of the low-barrier hydrogen bond in chymotrypsin. *Biochem.* 37:11940–11948.

Lodi, P. J., and Knowles, J. R. (1991). Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. *Biochem.* 30:6948–6956.

Parthasarathy, S., Ravinda, G., Balaram, H., Balaram, P., and Murthy, M. R. N. (2002). Structure of the plasmodium falciparum triosephosphate isomerase—phosphoglycolate complex in two crystal forms: characterization of catalytic open and closed conformations in the ligand-bound state. *Biochem.* 41:13178–13188.

Paetzel, M., and Dalbey, R. E. (1997). Catalytic hydroxyl/amine dyads within serine proteases. *Trends Biochem. Sci.* 22:28–31.

Perona, J. J., and Craik, C. S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* 272:29987–29990.

Schäfer T., Borchert T. W., Nielsen V. S., Skagerlind P., Gibson K., Wenger K., Hatzack F., Nilsson L. D., Salmon S., Pedersen S., Heldt-Hansen H. P., Poulsen P. B., Lund H., Oxenbøll K. M., Wu, G. F., Pedersen H. H., Xu, H. (2007). Industrial enzymes. *Adv. Biochem. Eng. Biotechnol. 2007* 105:59–131.

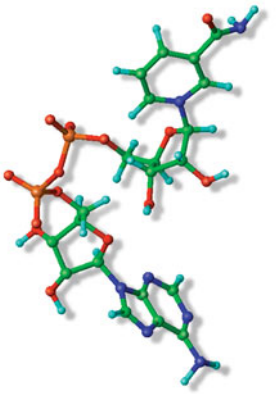
Steitz, T. A., and Shulman, R. G. (1982). Crystallographic and NMR studies of the serine proteases. *Annu. Rev. Biophys. Bioeng.* 11:419–444.

Von Dreele, R. B. (2005). Binding of *N*-acetylglucosamine oligosaccharides to hen egg-white lysozyme: a powder diffraction study. *Acta Crystallographic.* D61:22–32.

Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. S. (1998). Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc. Natl. Acad. Sci. USA* 95:8449–8454.

Downloaded From Vet-e-Books.com





Coenzymes and Vitamins

Evolution has produced a spectacular array of protein catalysts but the catalytic repertoire of an organism is not limited by the reactivity of amino acid side chains. Other chemical species, called **cofactors**, often participate in catalysis. Cofactors are required by inactive apoenzymes (proteins only) to convert them to active holoenzymes. There are two types of cofactors: **essential ions** (mostly metal ions) and organic compounds known as **coenzymes** (Figure 7.1). Both inorganic and organic cofactors become essential portions of the active sites of certain enzymes.

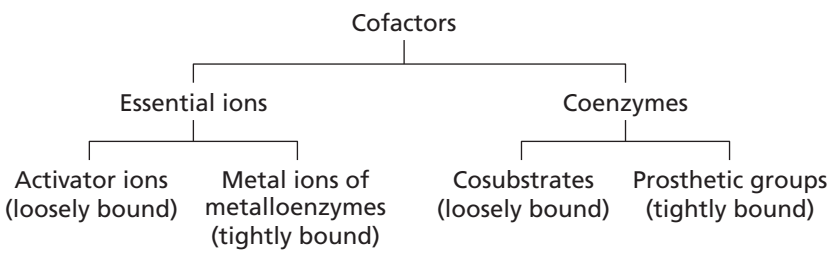
Many of the minerals required by all organisms are essential because they are cofactors. Some essential ions, called *activator ions*, are reversibly bound and often participate in the binding of substrates. In contrast, some cations are tightly bound and frequently participate directly in catalytic reactions.

Coenzymes act as group-transfer reagents. They accept and donate specific chemical groups. For some coenzymes, the group is simply hydrogen or an electron but other coenzymes carry larger, covalently attached chemical groups. These mobile metabolic groups are attached at the **reactive center** of the coenzyme. (Either the mobile metabolic group or the reactive center is shown in red in the structures presented in this chapter.) We can simplify our study of coenzymes by focusing on the chemical properties of their reactive centers. The two classes of coenzymes are described in Section 7.2.

We begin this chapter with a discussion of essentialion cofactors. Much of the rest of the chapter is devoted to the more complex organic cofactors. In mammals, many of these coenzymes are derived from dietary precursors called vitamins. We therefore discuss vitamins in this chapter. We conclude with a look at a few proteins that are coenzymes. Most of the structures and reactions presented here will be encountered in later chapters when we discuss particular metabolic pathways.

Finally, we come to a group of compounds which have only been known for a relatively short time, but which during this short time have attracted very considerable attention, both from chemists and from the public at large. Who today is unacquainted with vitamins, these mysterious substances which are of such immense significance for life, vita, itself and which have thus justifiably taken their name from it?

—H.G. Söderbaum Presentation speech for the Nobel Prize in chemistry to Adolf Windaus, 1928



◀ **Figure 7.1**
Types of cofactors. Essential ions and coenzymes can be further distinguished by the strength of interaction with their apoenzymes.

Top: Nicotinamide adenine dinucleotide ($NAD^{\oplus}$), a coenzyme derived from the vitamin nicotinic acid (niacin). $NAD^{\oplus}$ is an oxidizing agent.

Downloaded From Vet-e-Books.com

7.1 Many Enzymes Require Inorganic Cations

Over a quarter of all known enzymes require metallic cations to achieve full catalytic activity. These enzymes can be divided into two groups: metal-activated enzymes and metalloenzymes. **Metal-activated enzymes** either have an absolute requirement for added metal ions or are stimulated by the addition of metal ions. Some of these enzymes require monovalent cations such as $K^{\oplus}$ and others require divalent cations such as Ca^{2+} or Mg^{2+}. Kinases, for example, require magnesium ions for the magnesium-ATP complex they use as a phosphoryl group donating substrate. Magnesium shields the negatively charged phosphate groups of ATP making them more susceptible to nucleophilic attack (Section 10.6).

Metalloenzymes contain firmly bound metal ions at their active sites. The ions most commonly found in metalloenzymes are the transition metals, iron and zinc, and less often, copper and cobalt. Metal ions that bind tightly to enzymes are usually required for catalysis. The cations of some metalloenzymes can act as electrophilic catalysts by polarizing bonds. For example, the cofactor for the enzyme carbonic anhydrase is an electrophilic zinc atom bound to the side chains of three histidine residues and to a molecule of water. Binding to Zn^{2+} causes the water to ionize more readily. A basic carboxylate group of the enzyme removes a proton from the bound water molecule, producing a nucleophilic hydroxide ion that attacks the substrate (Figure 7.2). This enzyme has a very high catalytic rate partly because of the simplicity of its mechanism (Section 6.4). Many other zinc metalloenzymes activate bound water molecules in this fashion.

The ions of other metalloenzymes can undergo reversible oxidation and reduction by transferring electrons from a reduced substrate to an oxidized substrate. For example, iron is part of the heme group of catalase, an enzyme that catalyzes the degradation of H_2O_2. Similar heme groups also occur in cytochromes, electron-transferring proteins found associated with specific metalloenzymes in mitochondria and chloroplasts. Nonheme iron is often found in metalloenzymes in the form of iron-sulfur clusters (Figure 7.3). The most common iron-sulfur clusters are the [2 Fe–2 S] and [4 Fe–4 S] clusters in which the iron atoms are complexed with an equal number of sulfide ions from H_2S and $—S^{\ominus}$ groups from cysteine residues. Iron-sulfur clusters mediate some oxidation-reduction reactions. Each cluster, whether it contains two or four iron atoms, can accept only one electron in an oxidation reaction.

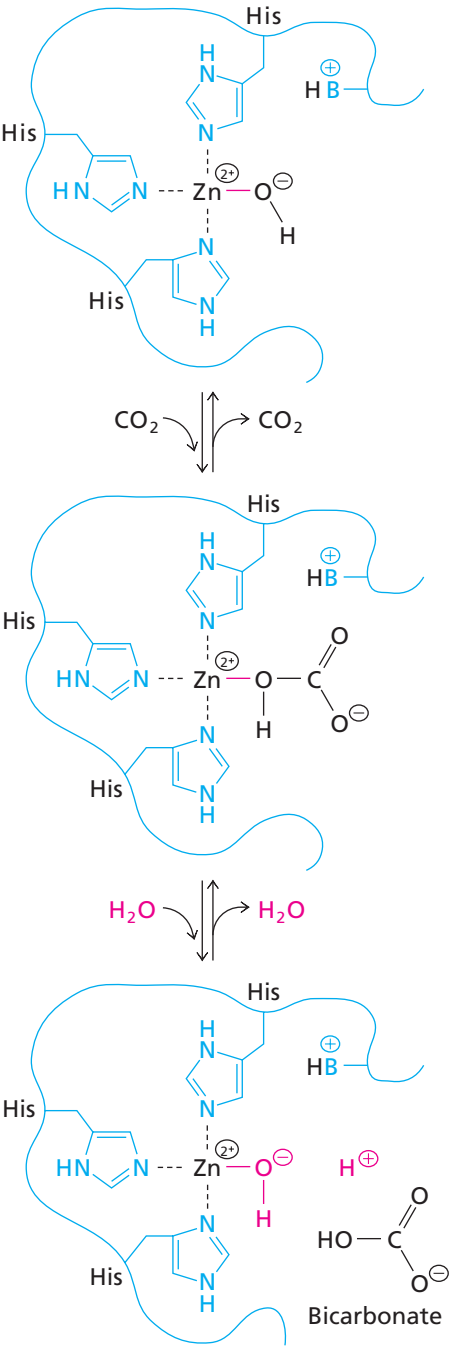
7.2 Coenzyme Classification

Coenzymes can be classified into two types based on how they interact with the apoenzyme (Figure 7.1). Coenzymes of one type—often called **cosubstrates**—are actually substrates in enzyme-catalyzed reactions. A cosubstrate is altered in the course of the reaction and dissociates from the active site. The original structure of the cosubstrate is regenerated in a subsequent reaction catalyzed by another enzyme. The cosubstrate is recycled repeatedly within the cell, unlike an ordinary substrate whose product typically undergoes further transformation. Cosubstrates shuttle mobile metabolic groups among different enzyme-catalyzed reactions.

The second type of coenzyme is called a **prosthetic group**. A prosthetic group remains bound to the enzyme during the course of the reaction. In some cases the prosthetic group is covalently attached to its apoenzyme, while in other cases it is tightly bound to the active site by many weak interactions. Like the ionic amino acid residues of the active site, a prosthetic group must return to its original form during each full catalytic event or the holoenzyme will not remain catalytically active. Cosubstrates and prosthetic groups are part of the active site of enzymes. They supply reactive groups that are not available on the side chains of amino acid residues.

Every living species uses coenzymes in a diverse number of important enzyme-catalyzed reactions. Most of these species are capable of synthesizing their coenzymes from simple precursors. This is especially true in four of the five kingdoms—prokaryotes, protists, fungi, and plants—but animals have lost the ability to synthesize some

Refer to Figure 1.1 for a table of the essential elements.

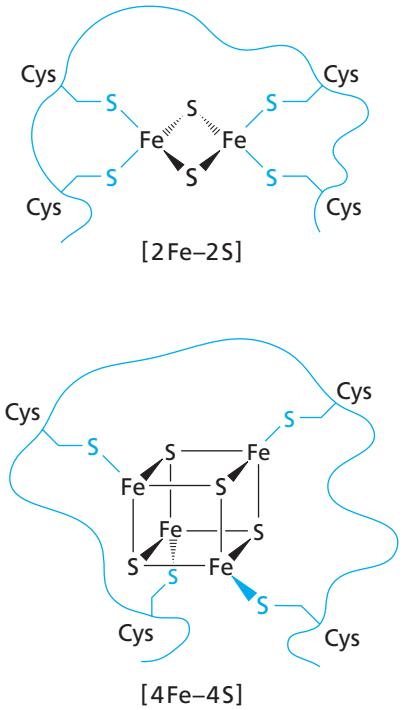


▲ **Figure 7.2**
Mechanism of carbonic anhydrase. The zinc ion in the active site promotes the ionization of a bound water molecule. The resulting hydroxide ion attacks the carbon atom of carbon dioxide, producing bicarbonate, which is released from the enzyme.

Review Section 4.12 for the structure of heme.

Cytochromes will be discussed in Section 7.16.

Downloaded From Vet-eBooks.com



▲ **Figure 7.3**
Iron-sulfur clusters. In each type of iron-sulfur cluster, the iron atoms are complexed with an equal number of sulfide ions (S^{2-}) and with the thiolate groups of the side chains of cysteine residues.

Table 7.1 Some vitamins and their associated deficiency diseases

Vitamin	Disease
Ascorbate (C)	Scurvy
Thiamine (B_1)	Beriberi
Riboflavin (B_2)	Growth retardation
Nicotinic acid (B_3)	Pellagra
Pantothenate (B_5)	Dermatitis in chickens
Pyridoxal (B_6)	Dermatitis in rats
Biotin (B_7)	Dermatitis in humans
Folate (B_9)	Anemia
Cobalamin (B_{12})	Pernicious anemia

The structure and chemistry of nucleotides is discussed in more detail in Chapter 19.



◀ **Brown rice and white rice.** Brown rice (top left) has been processed to remove the outer husks but it retains part of the outer skin or "bran." This skin contains thiamine (vitamin B_1). Further processing of the grain yields white rice (middle left), which lacks thiamine.

coenzymes. Mammals (including humans) need a source of coenzymes in order to survive. The ones they can't synthesize are supplied by nutrients, usually in small amounts (micrograms or milligrams per day). These essential compounds are called **vitamins** and animals rely on other organisms to supply these micronutrients. The ultimate sources of vitamins are usually plants and microorganisms. Most vitamins are coenzyme precursors—they must be enzymatically transformed to their corresponding coenzymes.

A vitamin-deficiency disease can result when a vitamin is deficient or absent in the diet of an animal. Such diseases can be overcome or prevented by consuming the appropriate vitamin. Table 7.1 lists nine vitamins and the diseases associated with their deficiencies. Each of these vitamins and their metabolic roles are discussed below. Most of them are converted to coenzymes, sometimes after a reaction with ATP.

The word *vitamin* (originally spelled "vitamine") was coined by Casimir Funk in 1912 to describe a "vital amine" from brown rice that cured beriberi, a nutritional-deficiency disease that results in neural degeneration. The term *vitamin* has been retained even though many vitamins proved not to be amines. Beriberi was first described in birds and then in humans whose diets consisted largely of polished rice. Christiaan Eijkman, a Dutch physician working in what was then the Dutch East Indies (now Indonesia), was the first to notice that chickens fed polished rice leftover from the local hospital developed beriberi but they recovered when they were fed brown rice. This discovery led eventually to isolation of an antiberiberi substance from the skin that covers brown rice. This substance became known as vitamin B_1 (thiamine).

Two broad classes of vitamins have since been identified: water-soluble (such as B vitamins) and fat-soluble (also called lipid vitamins). Water-soluble vitamins are required daily in small amounts because they are readily excreted in the urine and the cellular stores of their coenzymes are not stable. Conversely, lipid vitamins such as vitamins A, D, E, and K, are stored by animals and excessive intakes can result in toxic conditions known as hypervitaminoses. It's important to note that not all vitamins are coenzymes or their precursors (see Box 7.4 and Section 7.14).

The most common coenzymes are listed in Table 7.2 along with their metabolic role and their vitamin source. The following sections describe the structures and functions of these common coenzymes.

7.3 ATP and Other Nucleotide Cosubstrates

A number of nucleosides and nucleotides are coenzymes. Adenosine triphosphate (ATP) is by far the most abundant. Other common examples are GTP, *S*-adenosylmethionine, and nucleotide sugars such as uridine diphosphate glucose (UDP-glucose). ATP (Figure 7.4) is a versatile reactant that can donate its phosphoryl, pyrophosphoryl, adenylyl (AMP), or adenosyl groups in group-transfer reactions.

The most common reaction involving ATP is phosphoryl group transfer. In reactions catalyzed by kinases, for example, the γ-phosphoryl group of ATP is transferred to a nucleophile leaving ADP. The second most common reaction is nucleotidyl group transfer (transfer of the AMP moiety) leaving pyrophosphate (PP_i). ATP plays a central role in metabolism. Its role as a "high energy" cofactor is described in more detail in Chapter 10, "Introduction to Metabolism."

ATP is also the source of several other metabolite coenzymes. One, *S*-adenosylmethionine (Figure 7.5), is synthesized by the reaction of methionine with ATP.

$$\text{Methionine} + \text{ATP} \longrightarrow \text{S-Adenosylmethionine} + P_i + PP_i \qquad (7.1)$$

The normal thiomethyl group of methionine (—S—CH_3) is not very reactive but the positively charged sulfonium of *S*-adenosylmethionine is highly reactive. *S*-adenosylmethionine

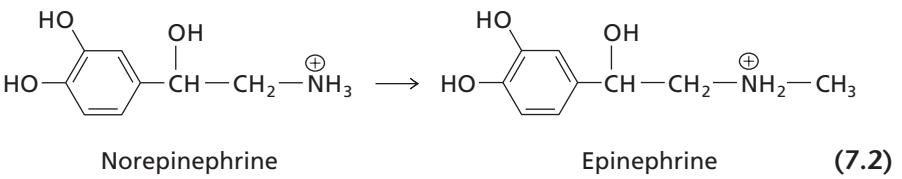
Downloaded From Vet-e-Books.com

Table 7.2 Major coenzymes

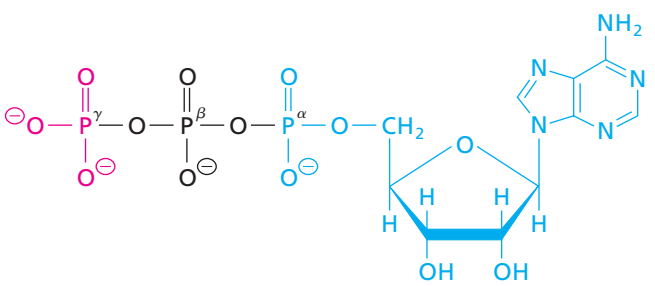
Coenzyme	Vitamin source	Major metabolic roles	Mechanistic role
Adenosine triphosphate (ATP)	—	Transfer of phosphoryl or nucleotidyl groups	Cosubstrate
S-Adenosylmethionine	—	Transfer of methyl groups	Cosubstrate
Uridine diphosphate glucose	—	Transfer of glycosyl groups	Cosubstrate
Nicotinamide adenine dinucleotide ($NAD^{\oplus}$) and nicotinamide adenine dinucleotide phosphate ($NADP^{\oplus}$)	Niacin (B_3)	Oxidation-reduction reactions involving two-electron transfer	Cosubstrate
Flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD)	Riboflavin (B_2)	Oxidation-reduction reactions involving one- and two-electron transfers	Prosthetic group
Coenzyme A (CoA)	Pantothenate (B_5)	Transfer of acyl groups	Cosubstrate
Thiamine pyrophosphate (TPP)	Thiamine (B_1)	Transfer of multi-carbon fragments containing a carbonyl group	Prosthetic group
Pyridoxal phosphate (PLP)	Pyridoxine (B_6)	Transfer of groups to and from amino acids	Prosthetic group
Biotin	Biotin (B_7)	ATP-dependent carboxylation of substrates or carboxyl-group transfer between substrates	Prosthetic group
Tetrahydrofolate	Folate	Transfer of one-carbon substituents, especially formyl and hydroxymethyl groups; provides the methyl group for thymine in DNA	Cosubstrate
Cobalamin	Cobalamin (B_{12})	Intramolecular rearrangements, transfer of methyl groups.	Prosthetic group
Lipoamide	—	Oxidation of a hydroxyalkyl group from TPP and subsequent transfer as an acyl group	Prosthetic group
Retinal	Vitamin A	Vision	Prosthetic group
Vitamin K	Vitamin K	Carboxylation of some glutamate residues	Prosthetic group
Ubiquinone (Q)	—	Lipid-soluble electron carrier	Cosubstrate
Heme Group	—	Electron transfer	Prosthetic group

reacts readily with nucleophilic acceptors and is the donor of almost all the methyl groups used in biosynthetic reactions. For example, it is required for conversion of the hormone norepinephrine to epinephrine.

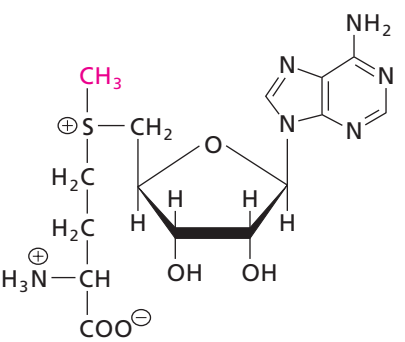
The thermodynamics of reactions involving ATP is explained in Section 10.6.



Norepinephrine Epinephrine **(7.2)**



▲ **Figure 7.4**
ATP. The nitrogenous base adenine is linked to a ribose bearing three phosphoryl groups. Transfer of a phosphoryl group (red) generates ADP, and transfer of a nucleotidyl group (AMP, blue) generates pyrophosphate.



▲ **Figure 7.5**
***S*-Adenosylmethionine.** The activated methyl group of this coenzyme is shown in red.

Downloaded From Vet-e-Books.com

BOX 7.1 MISSING VITAMINS

Whatever happened to vitamin B_4 and vitamin B_8? They are never listed in the textbooks but you'll often find them sold in stores that cater to the demand for supplements that might make you feel better and live longer.

Vitamin B_4 was adenine, the base found in DNA and RNA. We now know that it's not a vitamin. All species, including humans, can make copious quantities of adenine whenever it's needed (Sections 18.1 and 18.2). Vitamin B_8 was inositol, a precursor of several important lipids (Figure 8.16 and Section 9.12C). It's no longer considered a vitamin.

If you know anyone who is paying money for vitamin B_4 and B_8 supplements then here's your chance to be helpful. Tell them why they're wasting their money.



▲ **P.T. Barnum.** P.T. Barnum was a famous American showman. He's credited with saying, "There's a sucker born every minute." It's likely that the memorable phrase was coined by one of his rivals and later attributed to Barnum in order to discredit him.

Methylation reactions that require *S*-adenosylmethionine include methylation of phospholipids, proteins, DNA, and RNA. In plants, *S*-adenosylmethionine—as a precursor of the plant hormone ethylene—is involved in regulating the ripening of fruit.

Nucleotide-sugar coenzymes are involved in carbohydrate metabolism. The most common nucleotide sugar, uridine diphosphate glucose (UDP-glucose), is formed by the reaction of glucose 1-phosphate with uridine triphosphate (UTP) (Figure 7.6). UDP-glucose can donate its glycosyl group (shown in red) to a suitable acceptor, releasing UDP. UDP-glucose is regenerated when UDP accepts a phosphoryl group from ATP and the resulting UTP reacts with another molecule of glucose 1-phosphate.

Both the sugar and the nucleoside of nucleotide-sugar coenzymes may vary. Later on, we will encounter CDP, GDP, and ADP variants of this coenzyme.

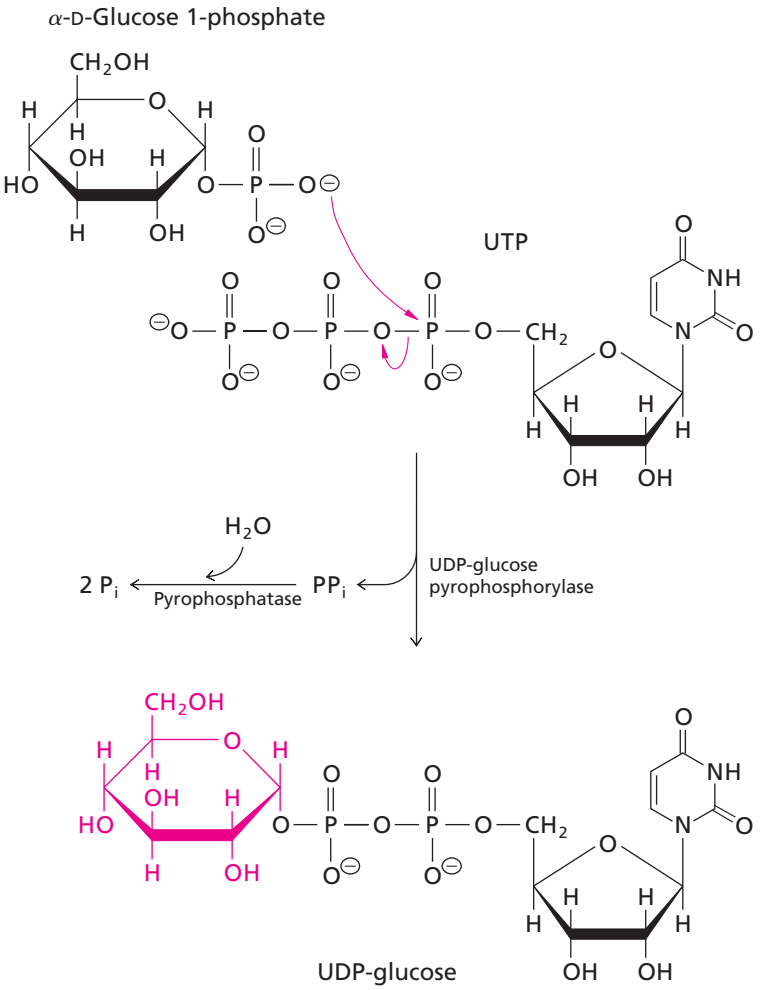
7.4 $NAD^{\oplus}$ and $NADP^{\oplus}$

The nicotinamide coenzymes are nicotinamide adenine dinucleotide ($NAD^{\oplus}$) and the closely related nicotinamide adenine dinucleotide phosphate ($NADP^{\oplus}$). These were the first coenzymes to be recognized. Both contain nicotinamide, the amide of nicotinic acid (Figure 7.7). Nicotinic acid (also called niacin) is the factor missing in the disease pellagra. Nicotinic acid or nicotinamide is essential as a precursor of $NAD^{\oplus}$ and $NADP^{\oplus}$. (In many species, tryptophan is degraded to nicotinic acid. Dietary tryptophan can therefore spare some of the requirement for niacin or nicotinamide.)

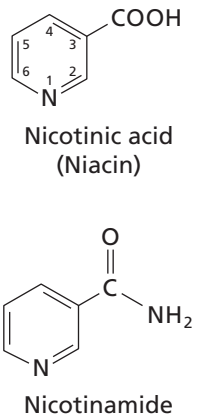
The nicotinamide coenzymes play a role in many oxidation–reduction reactions. They assist in the transfer of electrons to and from metabolites (Section 10.9). The oxidized forms, $NAD^{\oplus}$ and $NADP^{\oplus}$, are electron deficient and the reduced forms, NADH and NADPH, carry an extra pair of electrons in the form of a covalently bound hydride ion. The structures of these coenzymes are shown in Figure 7.8 . Both coenzymes contain a phosphoanhydride linkage that joins two 5′-nucleotides: AMP and the ribonucleotide of nicotinamide, called nicotinamide mononucleotide (NMN) (formed from nicotinic acid). In the case of $NADP^{\oplus}$, a phosphoryl group is present on the 2′-oxygen atom of the adenylate moiety.

Note that the $^{\oplus}$ sign in $NAD^{\oplus}$ simply indicates that the nitrogen atom carries a positive charge. This does not mean that the entire molecule is a positively charged ion; in fact, it is negatively charged due to the phosphates. A nitrogen atom normally has

Downloaded From Vet-e-Books.com



◀ **Figure 7.6**
Formation of UDP-glucose catalyzed by UDP-glucose pyrophosphorylase. An oxygen of the phosphate group of α-D-glucose 1-phosphate attacks the α-phosphorus of UTP. The PP_i released is rapidly hydrolyzed to $2P_i$ by the action of pyrophosphatase. This hydrolysis helps drive the pyrophosphorylase-catalyzed reaction toward completion. The mobile glycosyl group of UDP-glucose is shown in red.



▲ **Figure 7.7**
Nicotinic acid (niacin) and nicotinamide.

seven protons and seven electrons. The outer shell has five electrons that can participate in bond formation. In the oxidized form of the coenzyme ($NAD^⊕$ and $NADP^⊕$) the nicotinamide nitrogen is missing one of its electrons. It has only four electrons in the outer shell and those are shared with adjacent carbon atoms to form a total of four covalent bonds. (Each bond has a pair of electrons so the outer shell of the nitrogen atom is filled with eight shared electrons.) This is why we normally associate the positive charge with the ring nitrogen atom as shown in Figure 7.8. In fact, the charge is distributed over the entire aromatic ring.

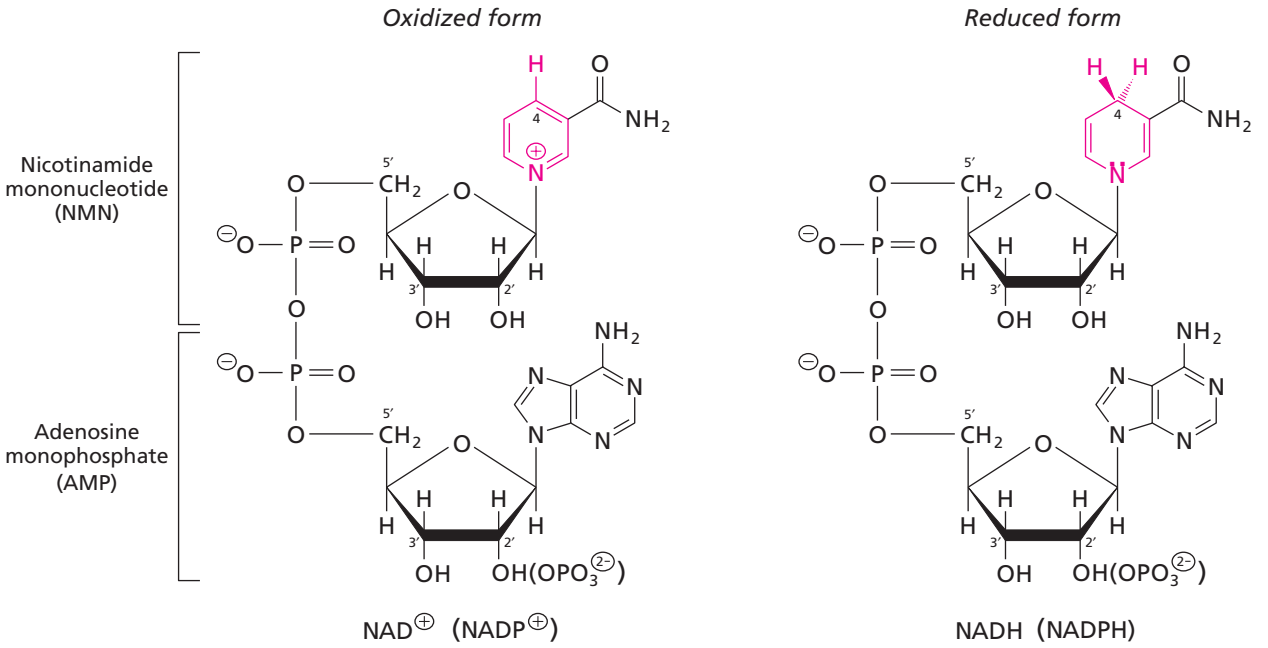
The reduced form of the nitrogen atom has its normal, full complement of electrons. In particular, the nitrogen atom has five electrons in its outer shell. Two of these electrons (represented by dots in Figure 7.8) are a free pair of electrons. The other three electrons participate in three covalent bonds.

$NAD^⊕$ and $NADP^⊕$ almost always act as cosubstrates for dehydrogenases. Pyridine nucleotide-dependent dehydrogenases catalyze the oxidation of their substrates by transferring two electrons and a proton in the form of a hydride ion ($H^⊖$) to C-4 of the nicotinamide group of $NAD^⊕$ or $NADP^⊕$. This generates the reduced form, NADH or NADPH, where a new C—H bond has formed at C-4 (one pair of electrons) and the electron previously associated with the ring double bond has delocalized to the ring nitrogen atom. Thus, oxidation by pyridine nucleotides (or reduction, the reverse reaction) always occurs two electrons at a time.

NADH and NADPH are said to possess reducing power (i.e., they are biological reducing agents). The stability of reduced pyridine nucleotides allows them to carry their reducing power from one enzyme to another, a property not shared by flavin

NADH and NADPH exhibit a peak of ultraviolet absorbance at 340 nm due to the dihydropyridine ring, whereas $NAD^⊕$ and $NADP^⊕$ do not absorb light at this wavelength. The appearance and disappearance of absorbance at 340 nm are useful for measuring the rates of oxidation and reduction reactions if they involve $NAD^⊕$ or $NADP^⊕$. (see Box 10.1).

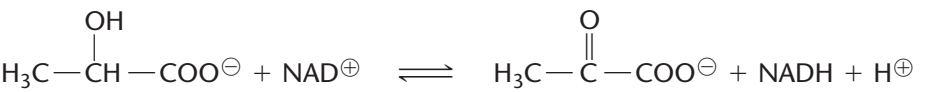
Downloaded From Vet-e-Books.com



▲ **Figure 7.8**
Oxidized and reduced forms of NAD (and NADP). The pyridine ring of $NAD^{\oplus}$ is reduced by the addition of a hydride ion to C-4 when $NAD^{\oplus}$ is converted to NADH (and when $NADP^{\oplus}$ is converted to NADPH). In $NADP^{\oplus}$, the 2′-hydroxyl group of the sugar ring of adenosine is phosphorylated. The reactive center of these coenzymes is shown in red.

coenzymes (Section 7.5). Most reactions forming NADH and NADPH are catabolic reactions and the subsequent oxidation of NADH by the membrane-associated electron transport system is coupled to the synthesis of ATP. Most NADPH is used as a reducing agent in biosynthetic reactions. The concentration of NADH is about ten times higher than that of NADPH.

Lactate dehydrogenase is an oxidoreductase that catalyzes the reversible oxidation of lactate. The enzyme is a typical NAD-dependent dehydrogenase. A proton is released from lactate when $NAD^{\oplus}$ is reduced.

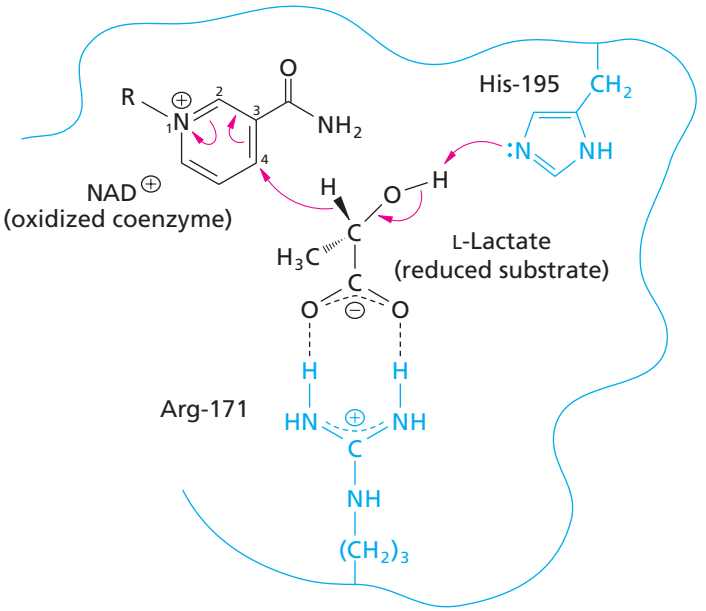


(7.3)

NADH is a cosubtrate, like ATP. When the reaction is complete, the structure of the cosubstrate is altered and the original form must be regenerated in a separate reaction. In this example, $NAD^{\oplus}$ is reduced to NADH and the reaction will soon reach equilibrium unless NADH is used up in a separate reaction where $NAD^{\oplus}$ is regenerated. We describe one example of how this is accomplished in Section 11.3B.

Figure 7.9 shows how both the enzyme and the coenzyme participate in the oxidation of lactate to pyruvate catalyzed by lactate dehydrogenase. In this mechanism, the coenzyme accepts a hydride ion at C-4 in the nicotinamide group. This leads to a rearrangement of bonds in the ring as electrons are shuffled to the positively charged nitrogen atom. The enzyme provides an acid–base catalyst and suitable binding sites for both the coenzyme and the substrate. Note that two hydrogens are removed from lactate to produce pyruvate (Equation 7.3). One of these hydrogens is transferred to $NAD^{\oplus}$ as a hydride ion carrying two electrons and the other is transferred to His-195 as a proton. The second hydrogen is subsequently released as $H^{\oplus}$ in order to regenerate the base catalyst (His-195). There are many examples of NAD-dependent reactions where the reduction of $NAD^{\oplus}$ is accompanied by release of a proton so it's quite common to see NADH + $H^{\oplus}$ on one side of the equation.

Downloaded From Vet-e-Books.com



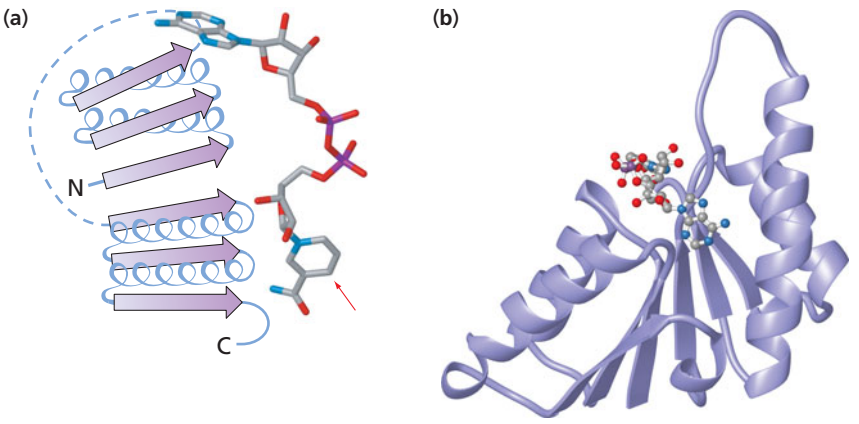
◀ **Figure 7.9**
Mechanism of lactate dehydrogenase. His-195, a base catalyst in the active site, abstracts a proton from the C-2 hydroxyl group of lactate, facilitating transfer of the hydride ion ($H^{\ominus}$) from C-2 of the substrate to C-4 of the bound $NAD^{\oplus}$. Arg-171 forms an ion pair with the carboxylate group of the substrate. In the reverse reaction, $H^{\ominus}$ is transferred from the reduced coenzyme, NADH, to C-2 of the oxidized substrate, pyruvate.

BOX 7.2 NAD BINDING TO DEHYDROGENASES

In the 1970s, structures were determined for four NAD-dependent dehydrogenases: lactate dehydrogenase, malate dehydrogenase, alcohol dehydrogenase, and glyceraldehyde 3-phosphate dehydrogenase. Each of these enzymes is oligomeric, with a chain length of about 350 amino acid residues. These chains all fold into two distinct domains—one to bind the coenzyme and one to bind the specific substrate. For each enzyme, the active site is in the cleft between the two domains.

As structures of more dehydrogenases were determined, several conformations of the coenzyme-binding motif were observed. Many of them possess one or more similar NAD- or NADP-binding structures consisting of a pair of $\beta\alpha\beta\alpha\beta$ units known as the Rossman fold after Michael Rossman, who first observed them in nucleotide-binding proteins (see figure). Each of the Rossman fold motifs binds to one half of the $NAD^{\oplus}$ dinucleotide. All of these enzymes bind the coenzyme in the same orientation and in a similar extended conformation.

Although many different dehydrogenases contain the Rossman fold motif, the rest of the structures may be very different and the dehydrogenases may not share significant sequence similarity. It's possible that all Rossman fold–containing enzymes descend from a common ancestor, but it's also possible that the motifs evolved independently in different dehydrogenases. That would be another example of convergent evolution.



◀ **NAD-binding region of some dehydrogenases.** **(a)** The coenzyme is bound in an extended conformation through interaction with two side-by-side motifs known as Rossman folds. The extended protein motifs form a β sheet of six parallel β strands. The arrow indicates the site where the hydride ion is added to C-4 of the nicotinamide group. **(b)** NADH bound to a Rossmann fold motif in rat lactate dehydrogenase [PDB 3H3F].

[Adapted from Rossman et al. (1975). *The Enzymes*, Vol. 11, Part A, 3rd ed., P. D., Boyer, ed. (New York: Academic Press), pp. 61–102.]

Downloaded From Vet-e-Books.com



▲ These yellow FADs are not flavins but Fish Aggregating Devices. They are buoys tethered to the sea floor in order to attract fish. This one has been deployed by the government of New South Wales off the east coast of Australia. The strong ocean current is threatening to carry it off.

Crystals of Old Yellow Enzyme, a typical flavoprotein, are shown in the introduction to Chapter 5.

7.5 FAD and FMN

The coenzymes flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) are derived from riboflavin, or vitamin B_2. Riboflavin is synthesized by bacteria, protists, fungi, plants, and some animals. Mammals obtain riboflavin from food. Riboflavin consists of the five-carbon alcohol ribitol linked to the N-10 atom of a heterocyclic ring system called isoalloxazine (Figure 7.10a). The riboflavin-derived coenzymes are shown in Figure 7.11b. Like $NAD^{\oplus}$ and $NADP^{\oplus}$, FAD contains AMP and a diphosphate linkage.

Many oxidoreductases require FAD or FMN as a prosthetic group. Such enzymes are called flavoenzymes or flavoproteins. The prosthetic group is very tightly bound, usually noncovalently. By binding the prosthetic groups tightly, the apoenzymes protect the reduced forms from wasteful reoxidation.

FAD and FMN are reduced to $FADH_2$ and $FMNH_2$ by taking up a proton and two electrons in the form of a hydride ion (Figure 7.11). The oxidized enzymes are bright yellow as a result of the conjugated double-bond system of the isoalloxazine ring system. The color is lost when the coenzymes are reduced to $FMNH_2$ and $FADH_2$.

$FMNH_2$ and $FADH_2$ donate electrons either one or two at a time, unlike NADH and NADPH that participate exclusively in two-electron transfers. A partially oxidized compound, FADH• or FMNH•, is formed when one electron is donated. These intermediates are relatively stable free radicals called semiquinones. The oxidation of $FADH_2$ and $FMNH_2$ is often coupled to reduction of a metalloprotein containing Fe^{3+} (in an [Fe–S] cluster). Because an iron–sulfur cluster can accept only one electron, the reduced flavin must be oxidized in two one-electron steps via the semiquinone intermediate. The ability of FMN to couple two-electron transfers with one-electron transfers is important in many electron transfer systems.

7.6 Coenzyme A and Acyl Carrier Protein

Many metabolic processes depend on coenzyme A (CoA, or HS-CoA) including the oxidation of fuel molecules and the biosynthesis of some carbohydrates and lipids. This coenzyme is involved in acyl-group–transfer reactions in which simple carboxylic acids and fatty acids are the mobile metabolic groups. Coenzyme A has three major components: a 2-mercaptoethylamine unit that bears a free —SH group, the vitamin pantothenate (vitamin B_5, an amide of β-alanine and pantoate), and an ADP moiety

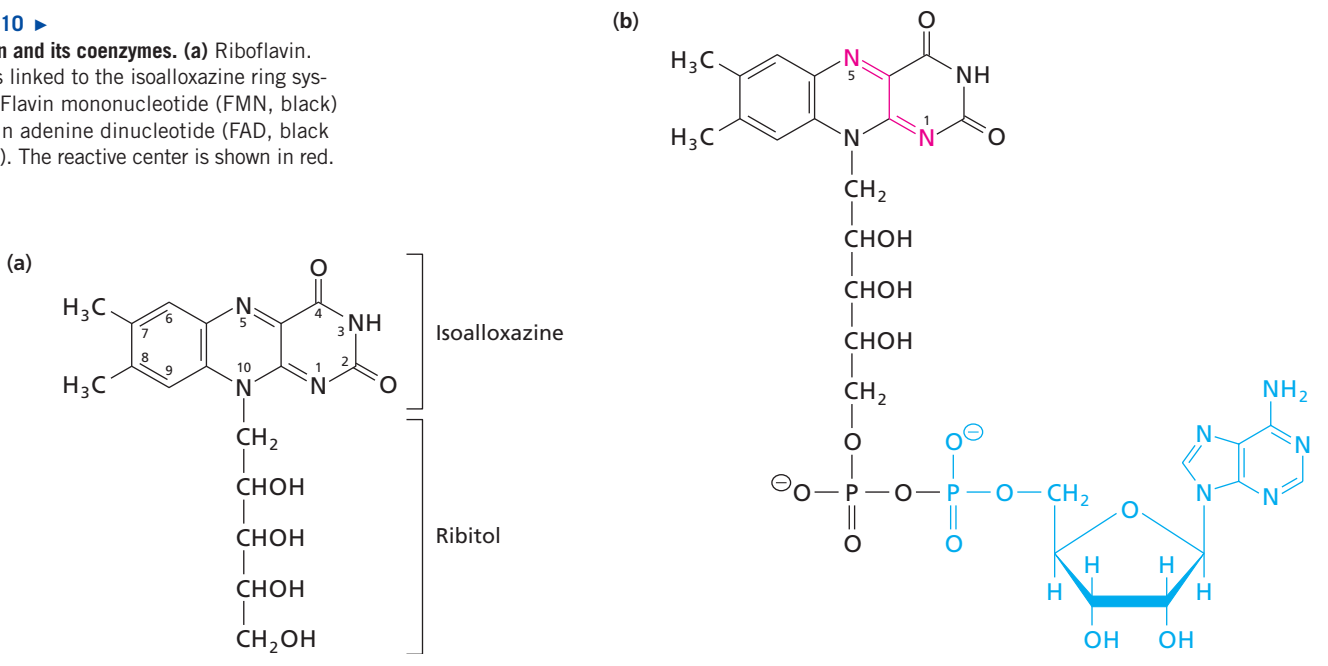
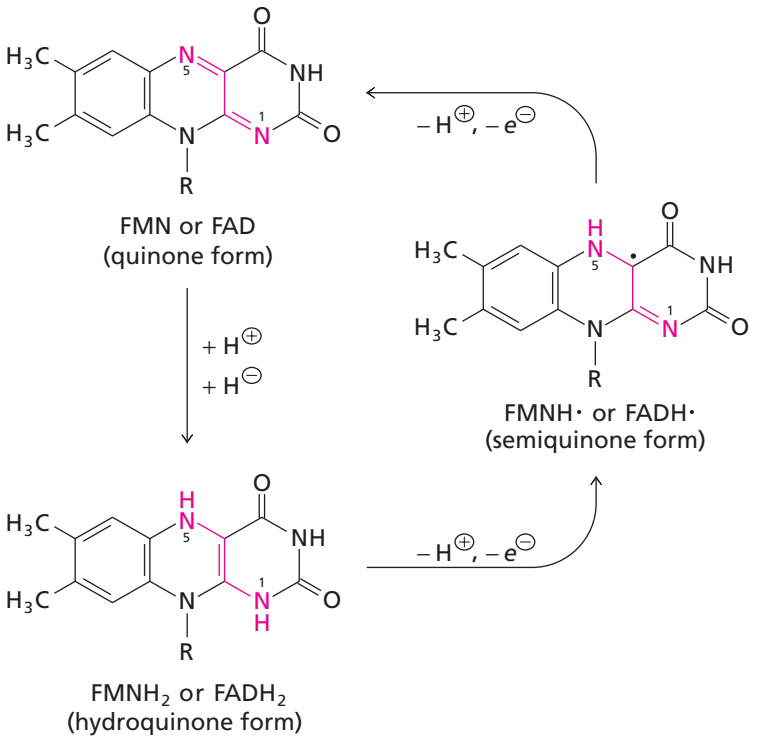


Figure 7.10 ►
Riboflavin and its coenzymes. (a) Riboflavin. Ribitol is linked to the isoalloxazine ring system. **(b)** Flavin mononucleotide (FMN, black) and flavin adenine dinucleotide (FAD, black and blue). The reactive center is shown in red.

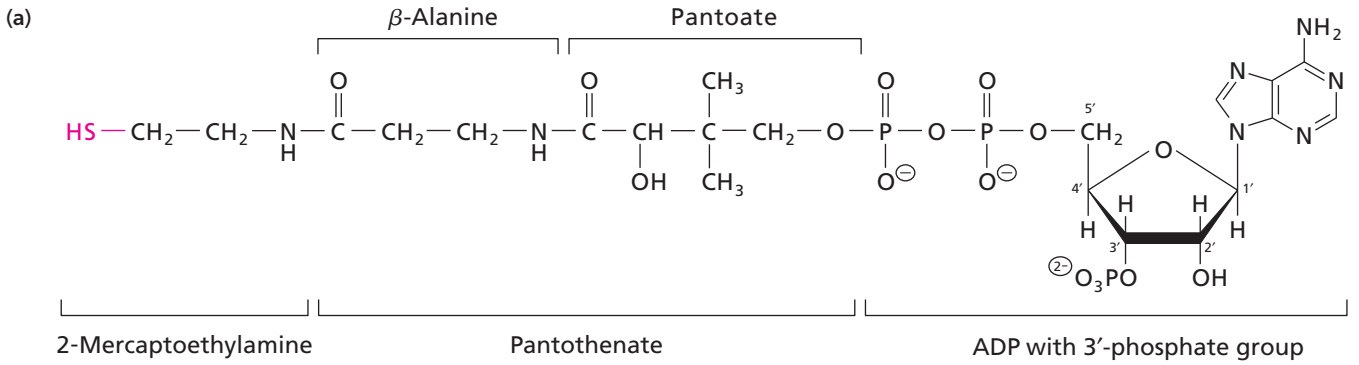
Downloaded From Vet-e-Books.com

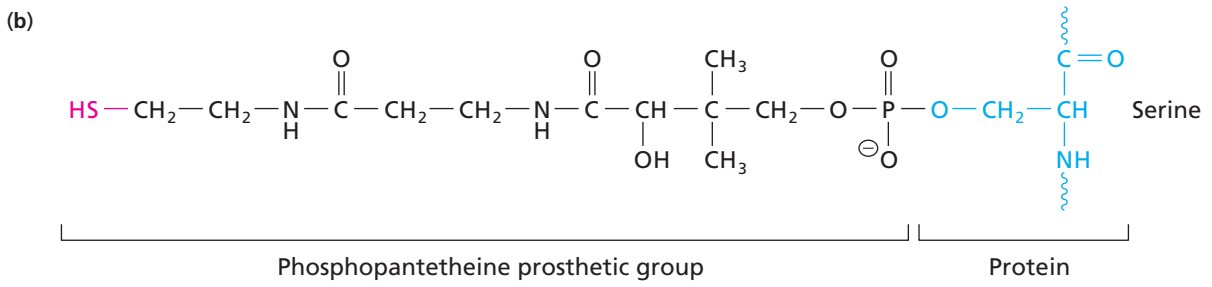


◀ **Figure 7.11**
Reduction and reoxidation of FMN or FAD. The conjugated double bonds between N-1 and N-5 are reduced by addition of a hydride ion and a proton to form $FMNH_2$ or $FADH_2$, respectively, the hydroquinone form of each coenzyme. Oxidation occurs in two steps. A single electron is removed by a one-electron oxidizing agent, with loss of a proton, to form a relatively stable free-radical intermediate. This semiquinone is then oxidized by removal of a proton and an electron to form fully oxidized FMN or FAD. These reactions are reversible.

whose 3′-hydroxyl group is esterified with a third phosphate group (Figure 7.12a). The reactive center of CoA is the —SH group. Acyl groups covalently attach to the —SH group to form thioesters. A common example is acetyl CoA (Figure 7.13), where the acyl group is an acetyl moiety. Acetyl CoA is a "high energy" compound due to the thioester linkage (Section 19.8). Coenzyme A was originally named for its role as the

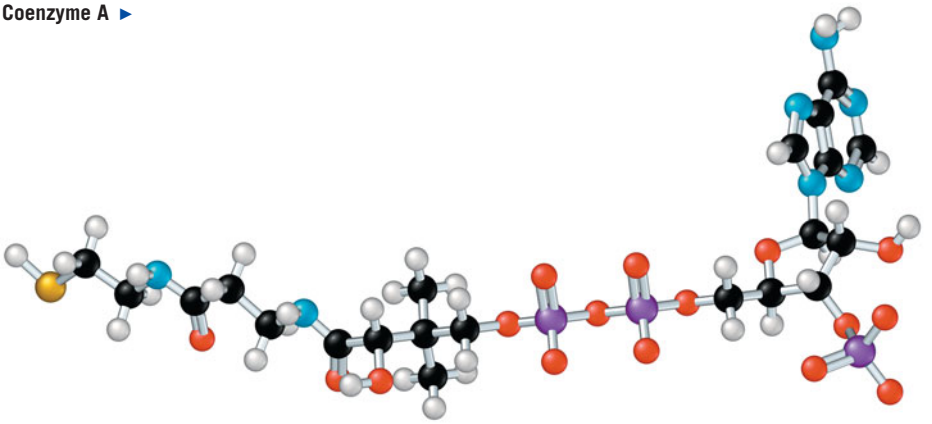
▼ **Figure 7.12**
Coenzyme A and acyl carrier protein (ACP). **(a)** In coenzyme A, 2-mercaptoethylamine is bound to the vitamin pantothenate, which in turn is bound via a phosphoester linkage to an ADP group that has an additional 3′-phosphate group. The reactive center is the thiol group (red). **(b)** In acyl carrier protein, the phosphopantetheine prosthetic group, which consists of the 2-mercaptoethylamine and pantothenate moieties of coenzyme A, is esterified to a serine residue of the protein.





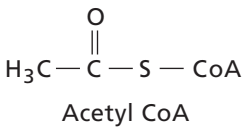
Downloaded From Vet-e-Books.com

Coenzyme A ▶



acetylation coenzyme. We will see acetyl CoA frequently when we discuss the metabolism of carbohydrates, fatty acids, and amino acids.

Phosphopantetheine, a phosphate ester containing the 2-mercaptoethylamine and pantothenate moieties of coenzyme A, is the prosthetic group of a small protein (77 amino acid residues) known as the acyl carrier protein (ACP). The prosthetic group is esterified to ACP via the side-chain oxygen of a serine residue (Figure 7.12b). The —SH of the prosthetic group of ACP is acylated by intermediates in the biosynthesis of fatty acids (Chapter 16).



▲ Figure 7.13
Acetyl CoA

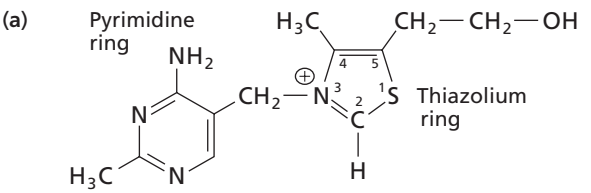
7.7 Thiamine Diphosphate

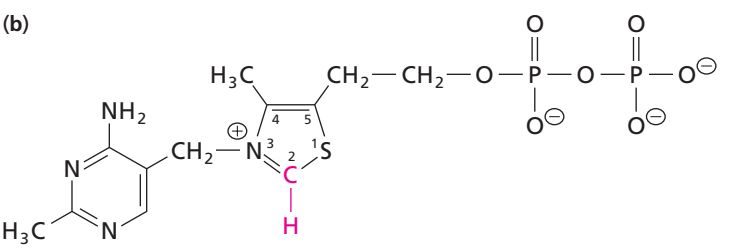
Thiamine (or vitamin B_1) contains a pyrimidine ring and a positively charged thiazolium ring (Figure 7.14a). The coenzyme is thiamine diphosphate (TDP), also called thiamine pyrophosphate (TPP) in the older literature (Figure 7.14b). TDP is synthesized from thiamine by enzymatic transfer of a pyrophosphoryl group from ATP.

About half a dozen decarboxylases (carboxylases) are known to require TDP as a coenzyme. For example, TDP is the prosthetic group of yeast pyruvate decarboxylase whose mechanism is shown in Figure 7.15. TDP is also a coenzyme involved in the oxidative decarboxylation of α-keto acids other than pyruvate. The first steps in those reactions proceed by the mechanism shown in Figure 7.15. In addition, TDP is a prosthetic group for enzymes known as transketolases that catalyze transfer between sugar molecules of two-carbon groups that contain a keto group.

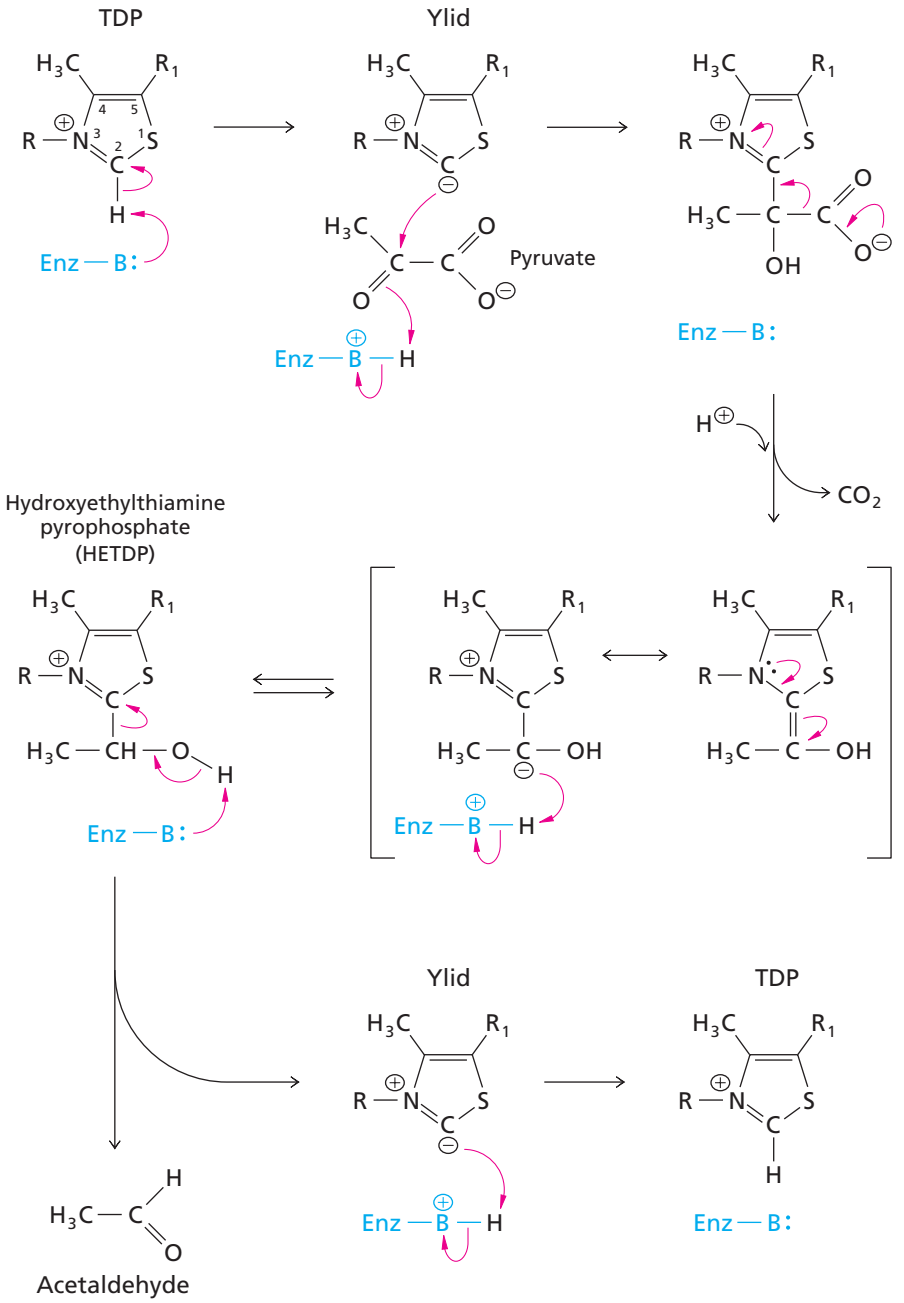
The metabolic role of pyruvate decarboxylase will be encountered in Section 11.3. Transketolases are discussed in Section 12.9. The role of TDP as a coenzyme in pyruvate dehydrogenase is described in Section 13.2.

Figure 7.14 ▶
Thiamine diphosphate (TDP). (a) Thiamine (vitamin B_1). **(b)** Thiamine diphosphate (TDP). The thiazolium ring of the coenzyme contains the reactive center (red).



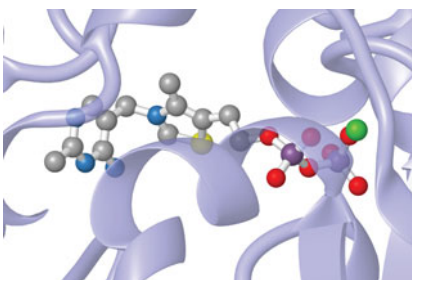


Downloaded From Vet-e-Books.com



◀ **Figure 7.15**
Mechanism of yeast pyruvate decarboxylase. The positive charge of the thiazolium ring of TDP attracts electrons, weakening the bond between C-2 and hydrogen. This proton is presumably removed by a basic residue of the enzyme. Ionization generates a dipolar carbanion known as an ylid (a molecule with opposite charges on adjacent atoms). The negatively charged C-2 attacks the electron-deficient carbonyl carbon of the substrate pyruvate and the first product (CO_2) is released. Two carbons of pyruvate are now attached to the thiazole ring as part of a resonance-stabilized carbanion. In the following step, protonation of the carbanion produces hydroxyethylthiamine diphosphate (HETDP). HETDP is cleaved, releasing acetaldehyde (the second product) and regenerating the ylid form of the enzyme-TDP complex. TDP re-forms when the ylid is protonated by the enzyme.

The thiazolium ring of the coenzyme contains the reactive center. C-2 of TDP has unusual reactivity; it is acidic despite its extremely high pK_a in aqueous solution. Similarly, recent experiments indicate that the pK_a value for the ionization of hydroxyethylthiamine diphosphate (HETDP) (i.e., formation of the dipolar carbanion) is changed from 15 in water to 6 at the active site of pyruvate decarboxylase. This increased acidity is attributed to low polarity of the active site, which also accounts for the reactivity of TDP.



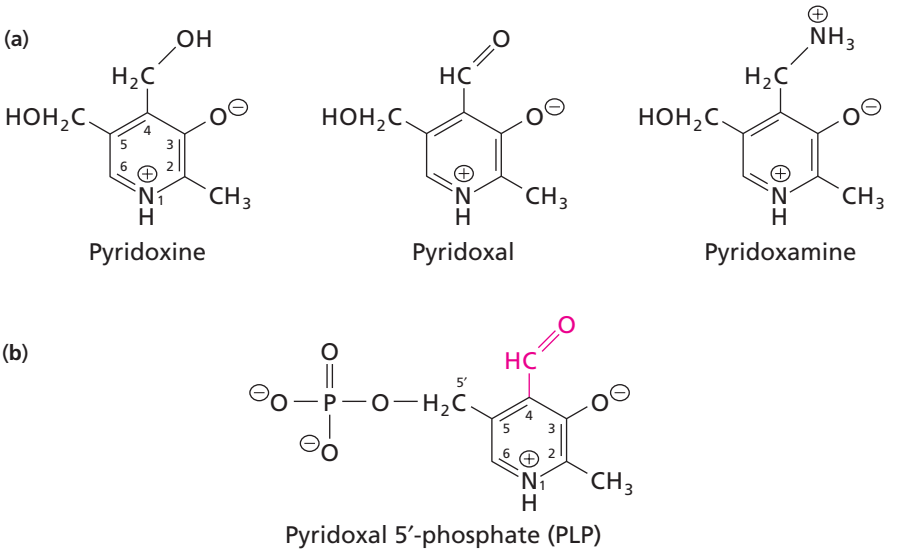
▲ **Thiamine diphosphate bound to pyruvate dehydrogenase.** The coenzyme is bound in an extended conformation and the diphosphate group is chelated to a magnesium ion (green). [PDB 1PYD]

7.8 Pyridoxal Phosphate

The B_6 family of water-soluble vitamins consists of three closely related molecules that differ only in the state of oxidation or amination of the carbon bound to position 4 of the pyridine ring (Figure 7.16a). Vitamin B_6—most often pyridoxal or pyridoxamine—is widely available from plant and animal sources. Induced B_6 deficiencies in rats result in dermatitis and various disorders related to protein metabolism but actual vitamin

Downloaded From Vet-e-Books.com

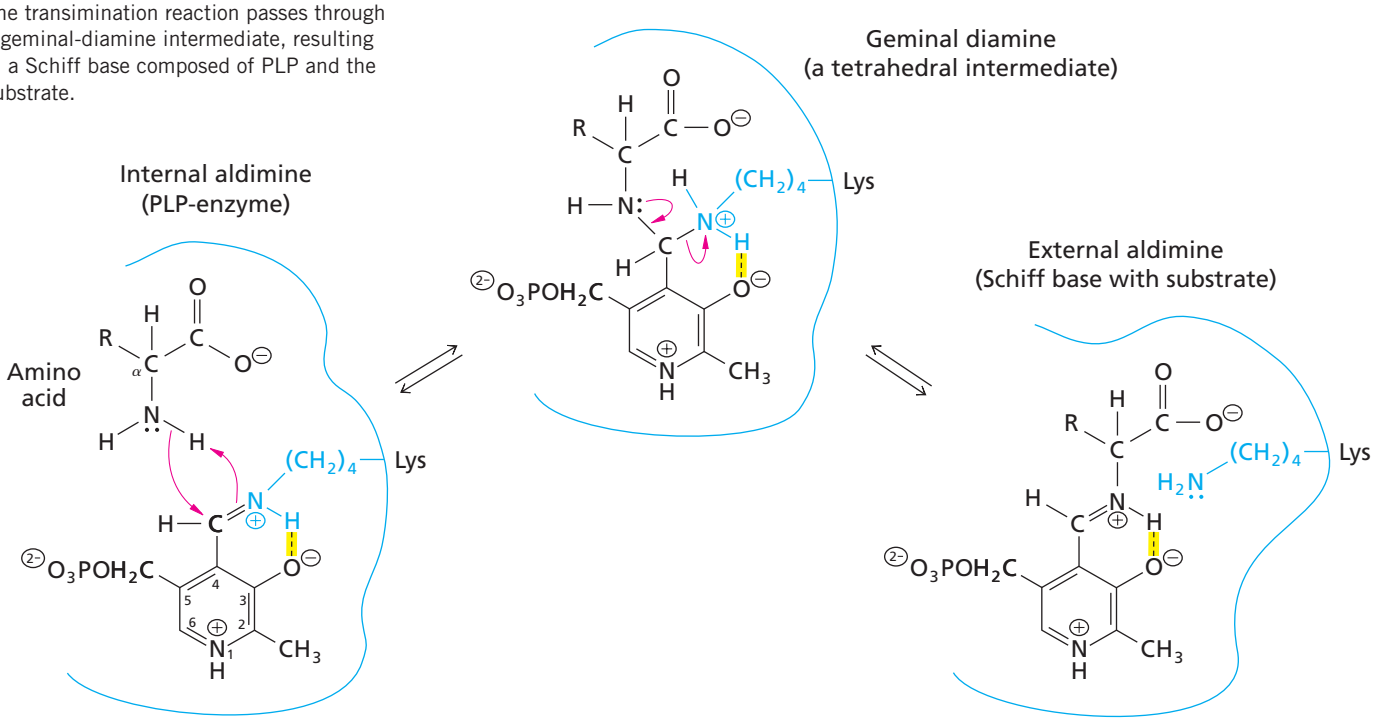
Figure 7.16 ►
B_6 vitamins and pyridoxal phosphate. (a) Vitamins of the B_6 family: pyridoxine, pyridoxal, and pyridoxamine. **(b)** Pyridoxal 5′-phosphate (PLP). The reactive center of PLP is the aldehyde group (red).



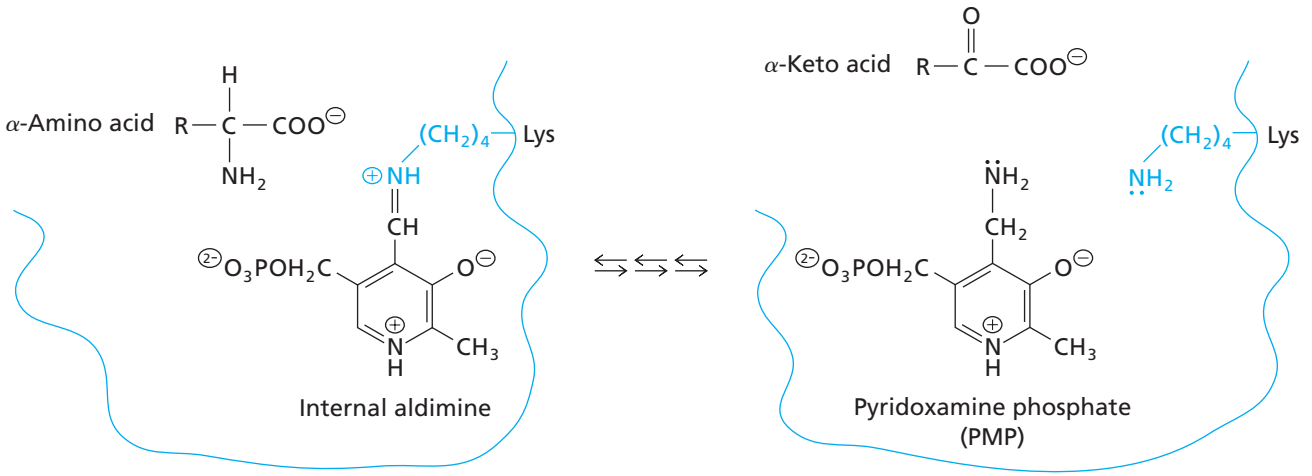
B_6 deficiencies in humans are rare. Enzymatic transfer of the γ-phosphoryl group from ATP forms the coenzyme pyridoxal 5′-phosphate (PLP) once vitamin B_6 enters a cell (Figure 7.16b).

Pyridoxal phosphate is the prosthetic group for many enzymes that catalyze a variety of reactions involving amino acids such as isomerizations, decarboxylations, and side-chain eliminations or replacements. In PLP-dependent enzymes, the carbonyl group of the prosthetic group is bound as a Schiff base (imine) to the ε-amino group of a lysine residue at the active site. (A Schiff base results from condensation of a primary amine with an aldehyde or ketone.) The enzyme-coenzyme Schiff base, shown on the left in Figure 7.17, is sometimes referred to as an internal aldimine. PLP is tightly bound to the enzyme by many weak noncovalent interactions; the additional covalent linkage of the internal aldimine helps prevent loss of the scarce coenzyme when the enzyme is not functioning.

Figure 7.17 ►
Binding of substrate to a PLP-dependent enzyme. The Schiff base linking PLP to a lysine residue of the enzyme is replaced by reaction of the substrate molecule with PLP. The transimination reaction passes through a geminal-diamine intermediate, resulting in a Schiff base composed of PLP and the substrate.



Downloaded From Vet-eBooks.com



▲ **Figure 7.18**
Mechanism of transaminases. An amino acid displaces lysine from the internal aldimine that links PLP to the enzyme, generating an external aldimine. Subsequent steps lead to the transfer of the amino group to PLP yielding an α-keto acid, which dissociates, and PMP, which remains bound to the enzyme. If another α-keto acid enters, each step proceeds in reverse. The amino group is transferred to the α-keto acid producing a new amino acid and regenerating the original PLP form of the enzyme.

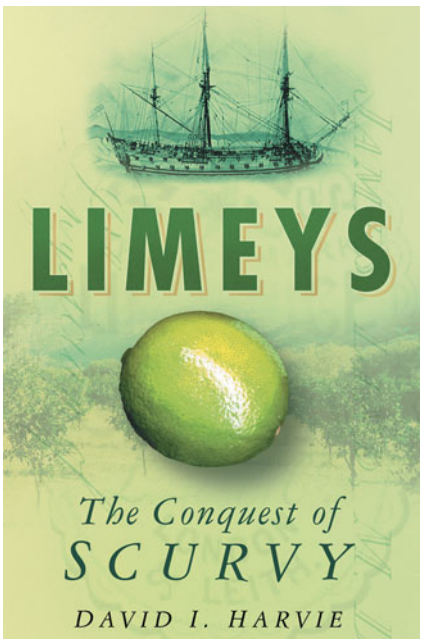
The initial step in all PLP-dependent enzymatic reactions with amino acids is the linkage of PLP to the α-amino group of the amino acid (formation of an external aldimine). When an amino acid binds to a PLP-enzyme, a transimination reaction takes place (Figure 7.17). This transfer reaction proceeds via a geminal-diamine intermediate rather than via formation of the free-aldehyde form of PLP. Note that the Schiff bases contain a system of conjugated double bonds in the pyridine ring ending with the positive charge on N-1. Similar ring structures with positively charged nitrogen atoms are present in $NAD^{\oplus}$. The prosthetic group serves as an electron sink during subsequent steps in the reactions catalyzed by PLP-enzymes. Once an α-amino acid forms a Schiff base with PLP, electron withdrawal toward N-1 weakens the three bonds to the α-carbon. In other words, the Schiff base with PLP stabilizes a carbanion formed when one of the three groups attached to the α-carbon of the amino acid is removed. Which group is lost depends on the chemical environment of the enzyme active site.

Removal of the α-amino group from amino acids is catalyzed by transaminases that participate in both the biosynthesis and degradation of amino acids (Chapter 17). Transamination is the most frequently encountered PLP-dependent reaction. The mechanism involves formation of an external aldimine (Figure 17.17) followed by release of the α-keto acid. The amino group remains bound to PLP forming pyridoxamine phosphate (PMP) (Figure 7.18). The next step in transaminase reactions is the reverse of the reaction shown in Figure 7.18 using a different α-keto acid as a substrate.

A specific transaminase is described in Section 17.2B.

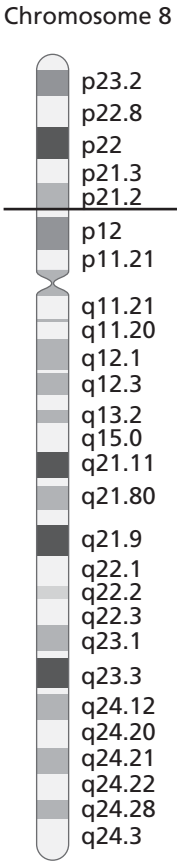
7.9 Vitamin C

The simplest vitamin is the antiscurvy agent ascorbic acid (vitamin C). Scurvy is a disease whose symptoms include skin lesions, fragile blood vessels, loose teeth, and bleeding gums.The link between scurvy and nutrition was recognized four centuries ago when British navy physicians discovered that citrus juice in limes and lemons were a remedy for scurvy in sailors whose diet lacked fresh fruits and vegetables. It was not until 1919, however, that ascorbic acid was isolated and shown to be the essential dietary component supplied by citrus juices.

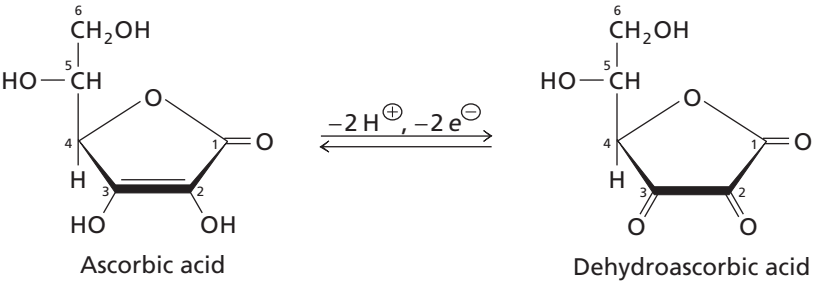


► *Limeys* is the story of Dr. James Lind and his attempt to promote citrus fruit as a cure for scurvy in the 1700s.

Downloaded From Vet-e-Books.com



▲ The human GULO pseudogene is located on the short arm of chromosome 8.



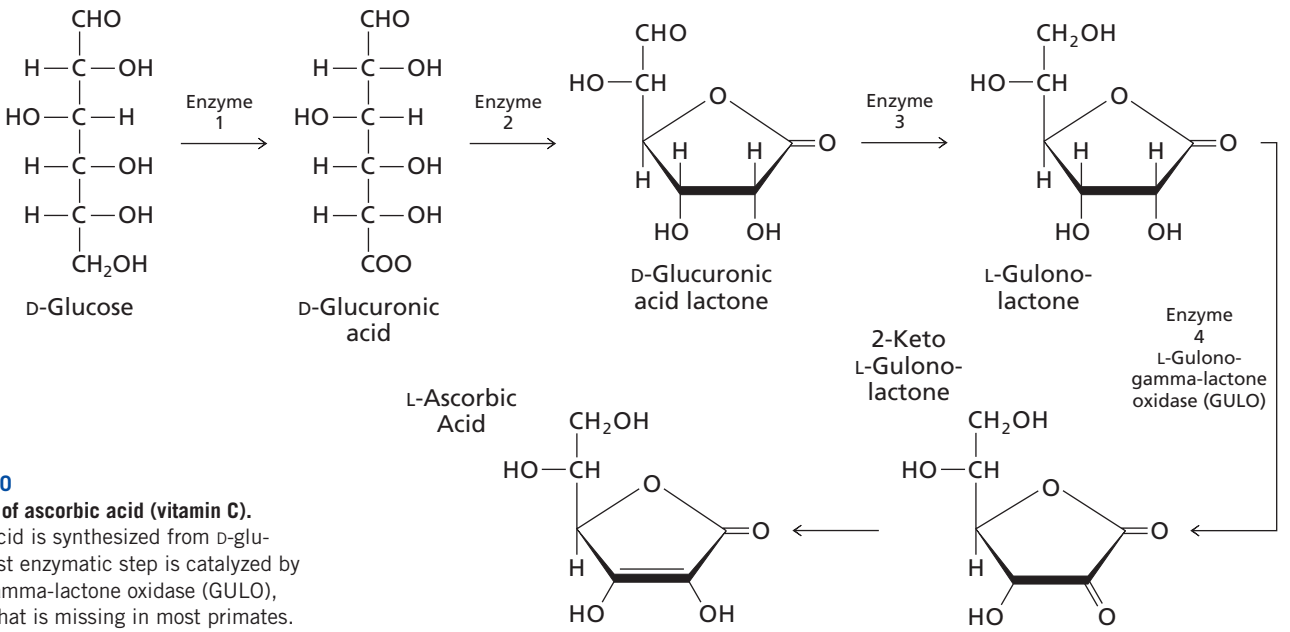
▲ **Figure 7.19**
Ascorbic acid (vitamin C) and its dehydro, oxidized form.

Back in the 18th century it was not easy to convince authorities that a simple solution like citrus fruit would solve the problem of scurvy because there were many competing theories. The story of Dr. James Lind and his efforts to convince the British navy is just one of many stories associated with vitamin C. It shows us that scientific evidence is not all that's required in order to make changes in the way we do things. Eventually, British sailors began to eat lemons and limes on a regular basis when they were at sea. Not only did this reduce the incidences of scurvy but it also gave rise to a famous nickname for British sailors. They were called "limeys" although lemons were much more effective than limes.

Ascorbic acid is a lactone, an internal ester in which the C-1 carboxylate group is condensed with the C-4 hydroxyl group, forming a ring structure. We now know that ascorbic acid is not a coenzyme but acts as a reducing agent in several different enzymatic reactions (Figure 7.19). The most important of these reactions is the hydroxylation of collagen (Section 4.12). Most mammals can synthesize ascorbic acid but guinea pigs, bats, and some primates (including humans) lack this ability and must therefore rely on dietary sources.

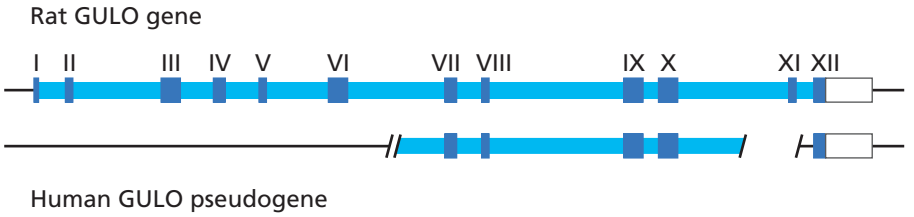
In most cases, we don't know very much about how certain enzymes disappeared from some species leading to a reliance on external sources for some essential metabolites. Most of the presumed gene disruption events happened so far in the distant past that few traces remain in modern genomes. The loss of ability to make vitamin C is an exception to that rule and serves as an instructive example of evolution.

Ascorbic acid is synthesized from D-glucose in a five-step pathway involving four enzymes (the last step is spontaneous). The last enzyme in the pathway is L-glucono-



▲ **Figure 7.20**
Biosynthesis of ascorbic acid (vitamin C). L-ascorbic acid is synthesized from D-glucose. The last enzymatic step is catalyzed by L-glucono-gamma-lactone oxidase (GULO), an enzyme that is missing in most primates.

Downloaded From Vet-eBooks.com



◀ **Figure 7.21**
Comparison of the intact rat GULO gene and the human pseudogene. The human pseudogene is missing the first six exons and exon 11. In addition, there are many mutations in the remaining exons that prevent them from producing protein product.

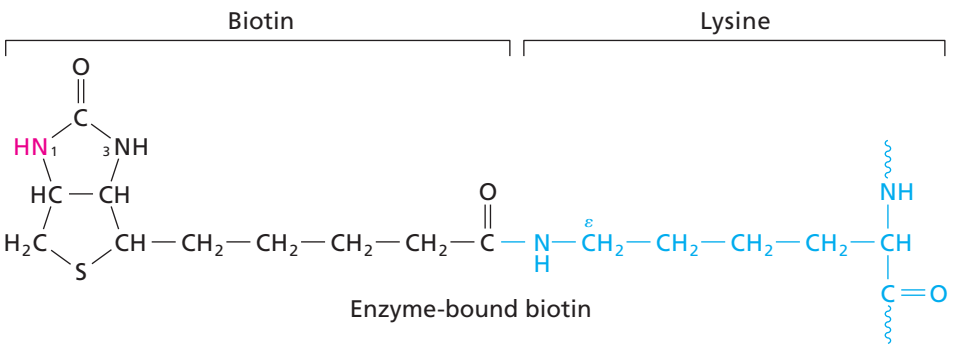
gamma-lactone oxidase (GULO) (Figure 7.20). GULO (the enzyme) is not present in primates of the haplorrhini family (monkeys and apes), but it is present in the strepsirrhini (lemurs, lorises etc.). These groups diverged about 80 million years ago. This led to the prediction that the GULO *gene* would be absent or defective in the monkeys and apes but intact in the other primates.

The prediction was confirmed with the discovery of a human GULO pseudogene on chromosome 8 in a block of genes that contains an active GULO gene in other animals. A comparison of the human pseudogene and a functional rat gene reveals many differences (Figure 7.21). The human pseudogene is missing the first six exons of the normal gene plus exon 11. The pseduogene in other apes is also missing these exons indicating that the ancestor of all apes had a similar defective GULO gene.

The original mutation that made the GULO gene inactive isn't known. Once inactivated, the pseudogene accumulated additional mutations that became fixed by random genetic drift. We can assume that lack of ability to synthesize vitamin C was not detrimental in these species because they obtained sufficient quantities in their normal diet.

7.10 Biotin

Biotin is a prosthetic group for enzymes that catalyze carboxyl group transfer reactions and ATP-dependent carboxylation reactions. Biotin is covalently linked to the active site of its host enzyme by an amide bond to the ε-amino group of a lysine residue (Figure 7.22).

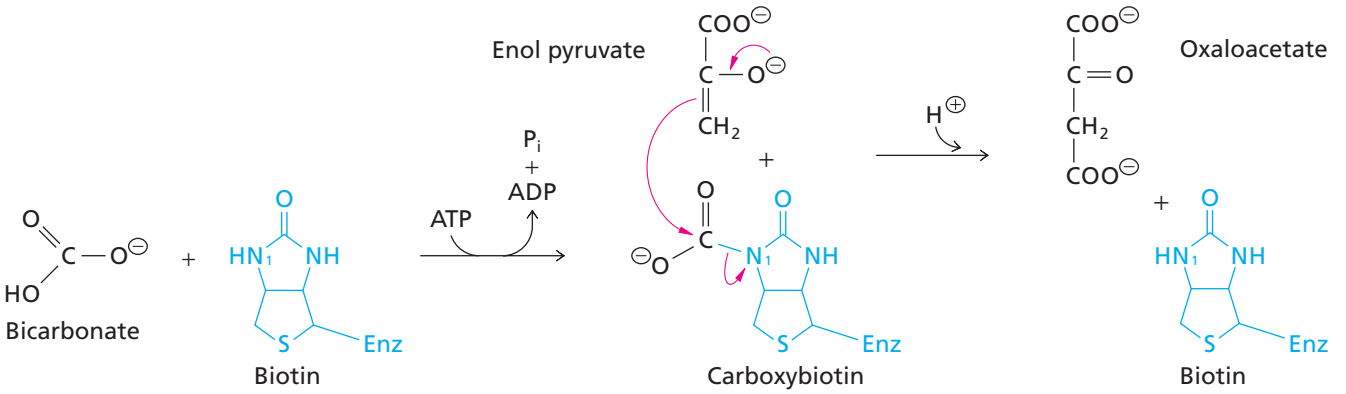


◀ **Figure 7.22**
Enzyme-bound biotin. The carboxylate group of biotin is covalently bound via amide linkage to the ε-amino group of a lysine residue (blue). The reactive center of the biotin moiety is N-1 (red).

The pyruvate carboxylase reaction demonstrates the role of biotin as a carrier of carbon dioxide (Figure 7.23). In this ATP-dependent reaction, pyruvate, a three-carbon acid, reacts with bicarbonate forming the four-carbon acid oxaloacetate. Enzyme-bound biotin is the intermediate carrier of the mobile carboxyl metabolic group. The pyruvate carboxylase reaction is an important CO_2 fixation reaction. It is required in the gluconeogenesis pathway (Chapter 11).

Biotin was first identified as an essential factor for the growth of yeast. Biotin deficiency is rare in humans or animals on normal diets because biotin is synthesized by intestinal bacteria and is required only in very small amounts (micrograms per day). A biotin deficiency can be induced, however, by ingesting raw egg whites that contain a protein called avidin. Avidin binds tightly to biotin making it unavailable for absorption

Downloaded From Vet-e-Books.com



▲ **Figure 7.23**
Reaction catalyzed by pyruvate carboxylase. First, biotin, bicarbonate, and ATP react to form carboxybiotin. The carboxybiotinyl-enzyme complex provides a stable, activated form of CO_2 that can be transferred to pyruvate. Next, the enolate form of pyruvate attacks the carboxyl group of carboxybiotin, forming oxaloacetate and regenerating biotin.

from the intestinal tract. Avidin is denatured when eggs are cooked and it loses its affinity for biotin.

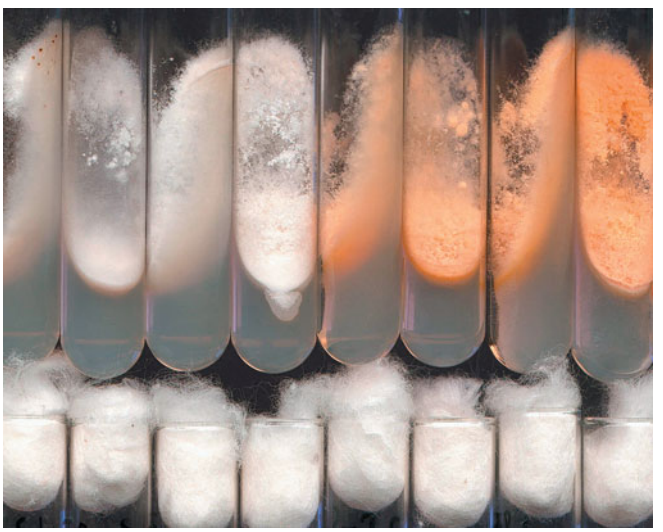
A variety of laboratory techniques take advantage of the high affinity of avidin for biotin. For example, a substance to which biotin is covalently attached can be extracted from a complex mixture by affinity chromatography (Section 3.6) on a column of immobilized avidin. The association constant for biotin and avidin is about 10^{15} M^{-1}—one of the tightest binding interactions known in biochemistry (see Section 4.9).

BOX 7.3 ONE GENE: ONE ENZYME

George Beadle and Edward Tatum wanted to test the idea that each gene encoded a single enzyme in a metabolic pathway. It was back in the late 1930s and this correspondence, which we now take for granted, was still a hypothesis. Remember, this was a time when it wasn't even clear whether genes were proteins or some other kind of chemical.

Beadle and Tatum chose the fungus *Neurospora crassa* for their experiments. *Neurospora* grows on a well-defined medium needing only sugar and biotin (vitamin B_7) as supplements. They reasoned that by irradiating *Neurospora* spores with X rays they could find mutants that would grow on rich supplemented medium but not on the simple defined medium. All they had to do next was identify the one supplement that needed to be added to the minimal medium to correct the defect. This would identify a gene for an enzyme that synthesized the now-essential supplement.

The 299th mutant required vitamin B_6 and the 1085th mutant required vitamin B_1. The B_6 and B_1 biosynthesis pathways were the first two pathways to be identified in this set of experiments. Later on, they worked out the genes/enzymes used in the tryptophan pathway. The results were published in 1941 and Beadle and Tatum received the Nobel Prize in Physiology or Medicine in 1958.



▲ ***Neurospora crassa* growing on defined medium in a test tube.** The strains on the right are producing orange carotenoid and the ones on the left are nonproducing strains.

(Source: Courtesy of Manchester University, United Kingdom).

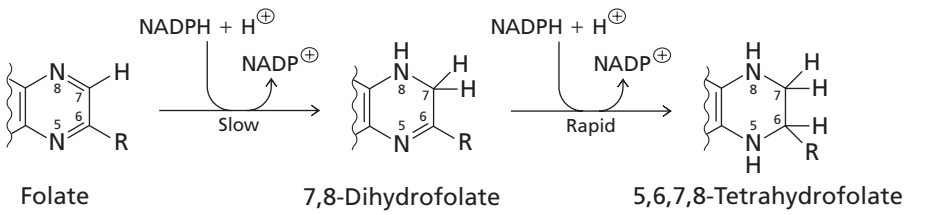
Downloaded From Vet-e-Books.com

7.11 Tetrahydrofolate

The vitamin folate was first isolated in the early 1940s from green leaves, liver, and yeast. Folate has three main components: pterin (2-amino-4-oxopteridine), a *p*-aminobenzoic acid moiety, and a glutamate residue. The structures of pterin and folate are shown in Figures 7.24a and 7.24b. Humans require folate in their diets because we cannot synthesize the pterin-*p*-aminobenzoic acid intermediate (PABA) and we cannot add glutamate to exogenous PABA.

The coenzyme forms of folate, known collectively as tetrahydrofolate, differ from the vitamin in two respects: they are reduced compounds (5,6,7,8-tetrahydropterins), and they are modified by the addition of glutamate residues bound to one another through γ-glutamyl amide linkages (Figure 7.24c). The anionic polyglutamyl moiety, usually five to six residues long, participates in the binding of the coenzymes to enzymes. When using the term *tetrahydrofolate*, keep in mind that it refers to compounds that have polyglutamate tails of varying lengths.

Tetrahydrofolate is formed from folate by adding hydrogen to positions 5, 6, 7, and 8 of the pterin ring system. Folate is reduced in two NADPH-dependent steps in a reaction catalyzed by dihydrofolate reductase (DHFR).

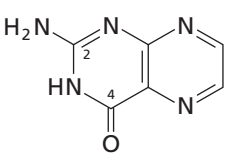


(7.4)

The primary metabolic function of dihydrofolate reductase is the reduction of dihydrofolate produced during the formation of the methyl group of thymidylate (dTMP) (Chapter 18). This reaction, which uses a derivative of tetrahydrofolate, is an essential step in the biosynthesis of DNA. Because cell division cannot occur when DNA synthesis is interrupted, dihydrofolate reductase has been extensively studied as a target for chemotherapy in the treatment of cancer (Box 18.4). In most species, dihydrofolate reductase is a relatively small monomeric enzyme that has evolved efficient binding sites for the two large substrates (folate and NADPH) (Figure 6.12).

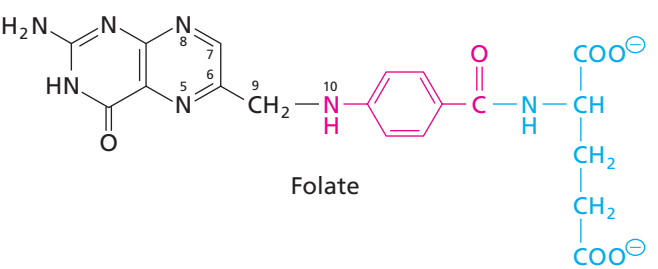
▼ **Figure 7.24**
Pterin, folate, and tetrahydrofolate. Pterin **(a)** is part of folate **(b)**, a molecule containing *p*-aminobenzoate (red) and glutamate (blue). **(c)** The polyglutamate forms of tetrahydrofolate usually contain five or six glutamate residues. The reactive centers of the coenzyme, N-5 and N-10, are shown in red.

(a)



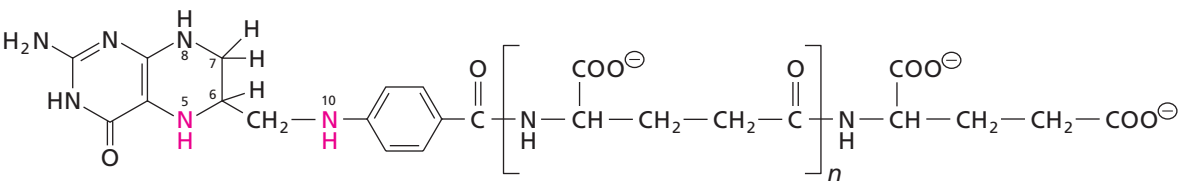
Pterin
(2-Amino-4-oxopteridine)

(b)



Folate

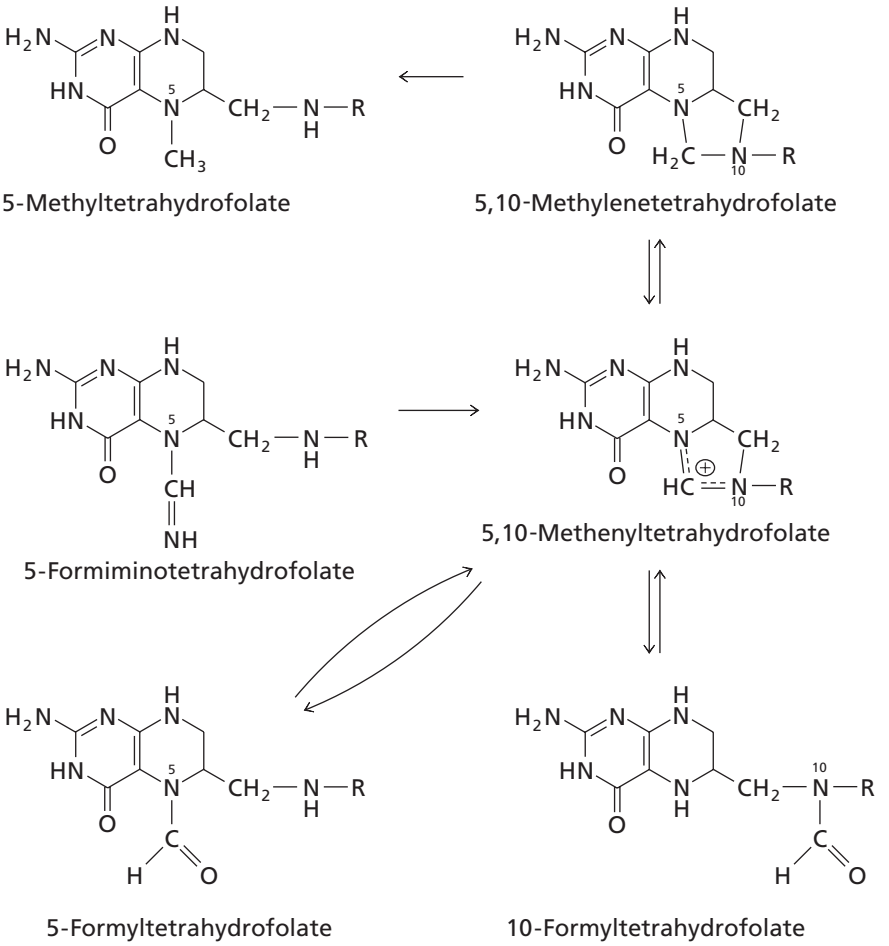
(c)



Tetrahydrofolate (Tetrahydrofolyl polyglutamate)

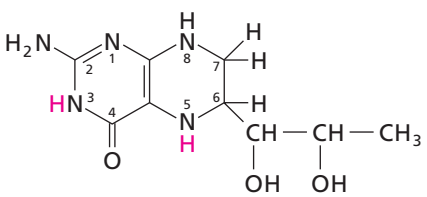
Downloaded From Vet-e-Books.com

Figure 7.25 ►
One-carbon derivatives of tetrahydrofolate. The derivatives can be interconverted enzymatically by the routes shown. (R represents the benzoyl polyglutamate portion of tetrahydrofolate.)





▲ **Many fruits and vegetables contain adequate supplies of folate.** Yeast and liver products are also excellent sources of folate.



▲ **Figure 7.26**
5,6,7,8-Tetrahydrobiopterin. The hydrogen atoms lost on oxidation are shown in red.

5,6,7,8-Tetrahydrofolate is required by enzymes that catalyze biochemical transfers of several one-carbon units. The groups bound to tetrahydrofolate are methyl, methylene, or formyl groups. Figure 7.25 shows the structures of several one-carbon derivatives of tetrahydrofolate and the enzymatic interconversions that occur among them. The one-carbon metabolic groups are covalently bound to the secondary amine N-5 or N-10 of tetrahydrofolate, or to both in a ring form. 10-Formyltetrahydrofolate is the donor of formyl groups and 5,10-methylenetetrahydrofolate is the donor of hydroxymethyl groups.

Another pterin coenzyme, 5,6,7,8-tetrahydrobiopterin, has a three-carbon side chain at C-6 of the pterin moiety in place of the large side chain found in tetrahydrofolate (Figure 7.26). This coenzyme is not derived from a vitamin but is synthesized by animals and other organisms. Tetrahydrobiopterin is the cofactor for several hydroxylases and will be encountered as a reducing agent in the conversion of phenylalanine to tyrosine (Chapter 17). It also is required by the enzyme that catalyzes the synthesis of nitric oxide from arginine (Section 17.12).

The sale of vitamins and supplements is big business in developed nations. It's often difficult to decide whether an extra supply of vitamins is necessary for good health because the scientific evidence is often missing or contradictory. Folate (vitamin B_9) deficiency is uncommon in normal, healthy adults and children in developed nations but there are documented cases of folate deficiency in pregnant women. A lack of tetrahydrofolate can lead to anemia and to severe defects in the developing fetus. While there are many fruits and vegetables that contain folate, it's a good idea for pregnant women to supplement their diet with folate in order to ensure their own health and that of the baby.

Downloaded From Vet-eBooks.com

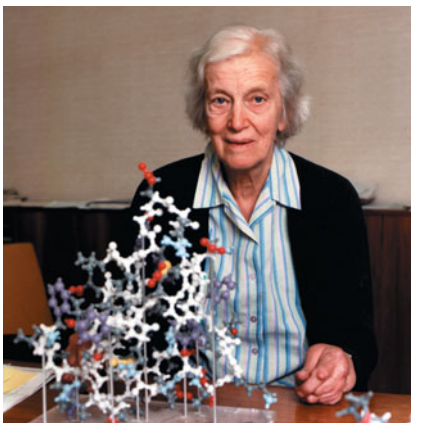
7.12 Cobalamin

Cobalamin (vitamin B_{12}) is the largest B vitamin and was the last to be isolated. The structure of cobalamin (Figure 7.27a) includes a corrin ring system that resembles the porphyrin ring system of heme (Figure 4.37). Note that cobalamin contains cobalt rather than the iron found in heme. The abbreviated structure shown in Figure 7.27b emphasizes the positions of two axial ligands bound to the cobalt, a benzimidazole ribonucleotide below the corrin ring and an R group above it. In the coenzyme forms of cobalamin, the R group is either a methyl group (in methylcobalamin) or a 5′-deoxyadenosyl group (in adenosylcobalamin).

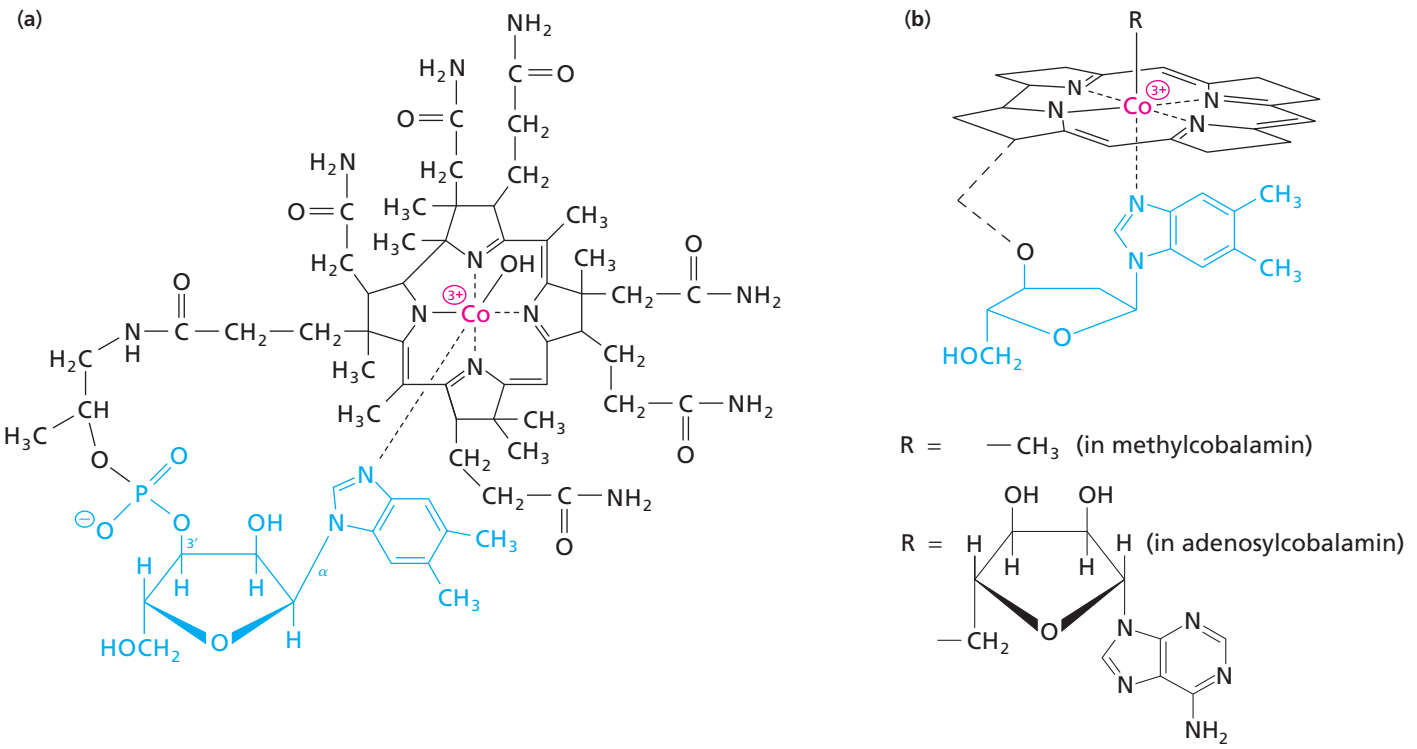
Cobalamin is synthesized by only a few microorganisms. It is required as a micronutrient by all animals and by some bacteria and algae. Humans obtain cobalamin from foods of animal origin. A deficiency of cobalamin can lead to pernicious anemia, a potentially fatal disease in which there is a decrease in the production of blood cells by bone marrow. Pernicious anemia can also cause neurological disorders. Most victims of pernicious anemia do not secrete a necessary glycoprotein (called intrinsic factor) from the stomach mucosa. This protein specifically binds cobalamin and the complex is absorbed by cells of the small intestine. Impaired absorption of cobalamin is now treated by regular injections of the vitamin.

The role of adenosylcobalamin reflects the reactivity of its C—Co bond. The coenzyme participates in several enzyme-catalyzed intramolecular rearrangements in which a hydrogen atom and a second group, bound to adjacent carbon atoms within a substrate, exchange places (Figure 7.28a). An example is the methylmalonyl–CoA mutase reaction (Figure 7.28b) that is important in the metabolism of odd-chain fatty acids (Chapter 16) and leads to the formation of succinyl CoA, an intermediate of the citric acid cycle.

Methylcobalamin participates in the transfer of methyl groups, as in the regeneration of methionine from homocysteine in mammals.



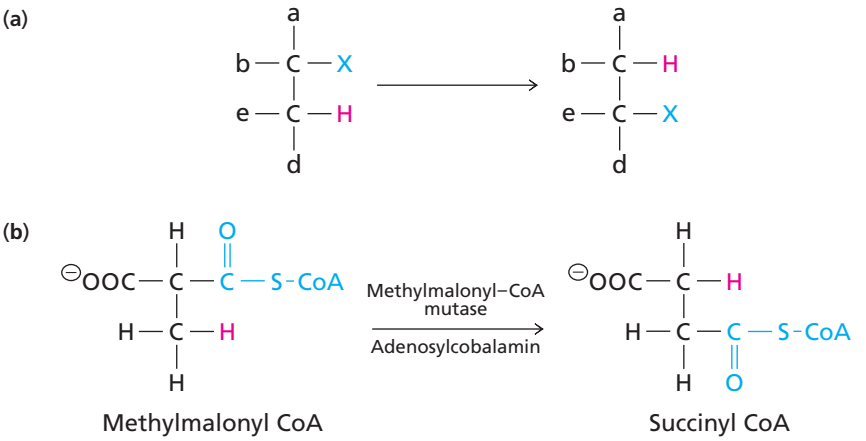
▲ **Dorothy Crowfoot Hodgkin (1910–1994).** Hodgkin received the Nobel Prize in 1964 for determining the structure of vitamin B_{12} (cobalamin). The structure of insulin, shown in the photograph, was published in 1969.



▲ **Figure 7.27**
Cobalamin (vitamin B_{12}) and its coenzymes. (a) Detailed structure of cobalamin showing the corrin ring system (black) and 5,6-dimethylbenzimidazole ribonucleotide (blue). The metal coordinated by corrin is cobalt (red). The benzimidazole ribonucleotide is coordinated with the cobalt of the corrin ring and is also bound via a phosphoester linkage to a side chain of the corrin ring system. **(b)** Abbreviated structure of cobalamin coenzymes. A benzimidazole ribonucleotide lies below the corrin ring, and an R group lies above the ring.

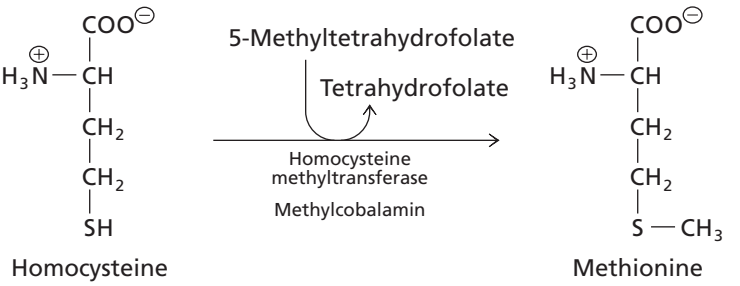
Downloaded From Vet-e-Books.com

Figure 7.28 ►
Intramolecular rearrangements catalyzed by adenosylcobalamin-dependent enzymes. **(a)** Rearrangement in which a hydrogen atom and a substituent on an adjacent carbon atom exchange places. **(b)** Rearrangement of methylmalonyl CoA to succinyl CoA, catalyzed by methylmalonyl–CoA mutase.





▲ **Intestinal bacteria.** Normal, healthy humans harbor billions of bacteria in their intestines. There are at least several dozen different species. The one shown here is *Helicobacter pylori*, which causes stomach ulcers when it invades the stomach. The bacteria are sitting on the surface of the intestine that has many projections for absorbing nutrients. Other common species are *Escherichia coli* and various species of *Actinomyces* and *Streptococcus*. These bacteria help break down ingested food and they supply many of the essential vitamins and amino acids that humans need, especially cobalamin.



(7.5)

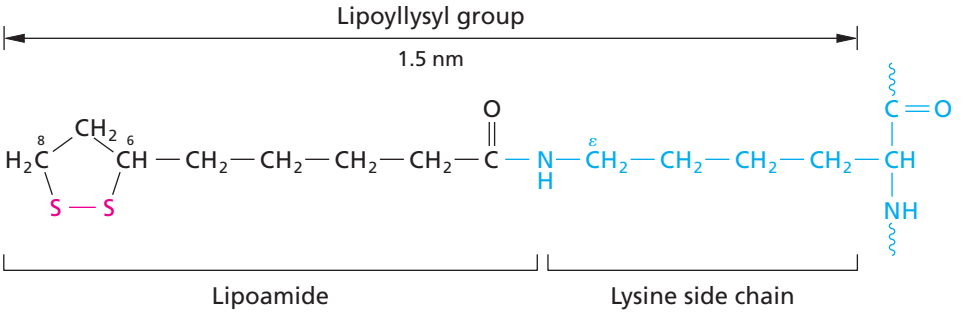
In this reaction, the methyl group of 5-methyltetrahydrofolate is passed to a reactive, reduced form of cobalamin to form methylcobalamin that can transfer the methyl group to the thiol side chain of homocysteine.

7.13 Lipoamide

The lipoamide coenzyme is the protein-bound form of lipoic acid. Lipoic acid is sometimes described as a vitamin but animals appear to be able to synthesize it. It is required by certain bacteria and protozoa for growth. Lipoic acid is an eight-carbon carboxylic acid (octanoic acid) in which two hydrogen atoms, on C-6 and C-8, have been replaced by sulfhydryl groups in disulfide linkage. Lipoic acid does not occur free—it is covalently attached via an amide linkage through its carboxyl group to the ε-amino group of a lysine residue of a protein (Figure 7.29). This structure is found in dihydrolipoamide acyltransferases that are components of the pyruvate dehydrogenase complex and related enzymes.

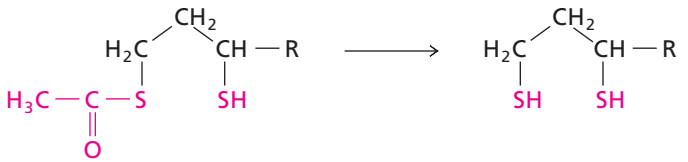
Lipoamide carries acyl groups between active sites in multienzyme complexes. For example, in the pyruvate dehydrogenase complex (Section 12.2), the disulfide ring of

Figure 7.29 ►
Lipoamide. Lipoic acid is bound in amide linkage to the ε-amino group of a lysine residue (blue) of dihydrolipoamide acyltransferases. The dithiolane ring of the lipoyllysyl groups is extended 1.5 nm from the polypeptide backbone. The reactive center of the coenzyme is shown in red.



Downloaded From Vet-e-Books.com

the lipoamide prosthetic group reacts with HETDP (Figure 7.15) binding its acetyl group to the sulfur atom attached to C-8 of lipoamide and forming a thioester. The acyl group is then transferred to the sulfur atom of a coenzyme A molecule generating the reduced (dihydrolipoamide) form of the prosthetic group.



(7.6)

The final step catalyzed by the pyruvate dehydrogenase complex is the oxidation of dihydrolipoamide. In this reaction, NADH is formed by the action of a flavoprotein component of the complex. The actions of the multiple coenzymes of the pyruvate dehydrogenase complex show how coenzymes, by supplying reactive groups that augment the catalytic versatility of proteins, are used to conserve both energy and carbon building blocks.

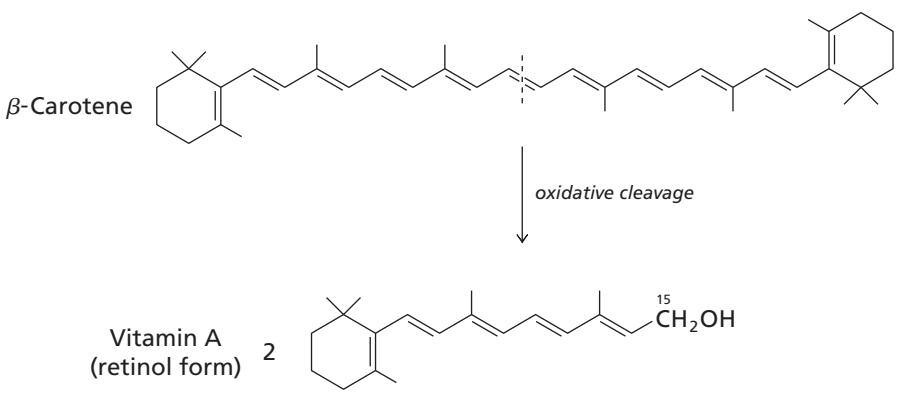
7.14 Lipid Vitamins

The structures of the four lipid vitamins (A, D, E, and K) contain rings and long aliphatic side chains. The lipid vitamins are highly hydrophobic although each possesses at least one polar group. In humans and other mammals, ingested lipid vitamins are absorbed in the intestine by a process similar to the absorption of other lipid nutrients (Section 16.1a). After digestion of any proteins that may bind them, they are carried to the cellular interface of the intestine as micelles formed with bile salts. The study of these hydrophobic molecules has presented several technical difficulties so research on their mechanisms has progressed more slowly than that on their water-soluble counterparts. Lipid vitamins differ widely in their functions, as we will see below.

A. Vitamin A

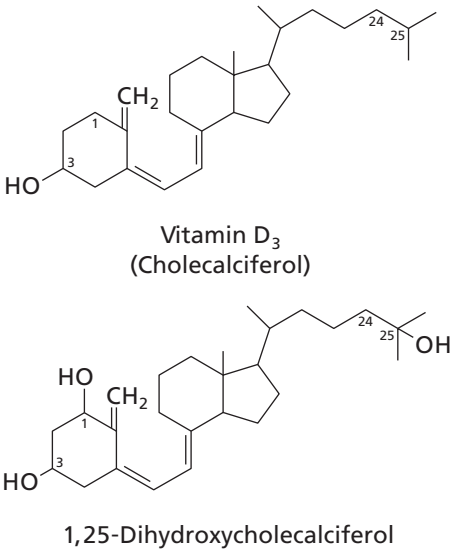
Vitamin A, or retinol, is a 20-carbon lipid molecule obtained in the diet either directly or indirectly from β-carotene. Carrots and other yellow vegetables are rich in β-carotene, a 40-carbon plant lipid whose enzymatic oxidative cleavage yields vitamin A (Figure 7.30). Vitamin A exists in three forms that differ in the oxidation state of the terminal functional group: the stable alcohol retinol, the aldehyde retinal, and retinoic acid. Their hydrophobic side chain is formed from repeated isoprene units (Section 9.6).

All three vitamin A derivatives have important biological functions. Retinoic acid is a signal compound that binds to receptor proteins inside cells; the ligand–receptor



◀ **Figure 7.30**
Formation of vitamin A from β-carotene.

Downloaded From Vet-e-Books.com



▲ **Figure 7.31**
Vitamin D_3 (cholecalciferol) and 1,25-dihydroxycholecalciferol. (Vitamin D_2 has an additional methyl group at C-24 and a *trans* double bond between C-22 and C-23.) 1,25-Dihydroxycholecalciferol is produced from vitamin D_3 by two separate hydroxylations.

complexes then bind to chromosomes and can regulate gene expression during cell differentiation. The aldehyde retinal is a light-sensitive compound with an important role in vision. Retinal is the prosthetic group of the protein rhodopsin; absorption of a photon of light by retinal triggers a neural impulse.

B. Vitamin D

Vitamin D is the collective name for a group of related lipids. Vitamin D_3 (cholecalciferol) is formed nonenzymatically in the skin from the steroid 7-dehydrocholesterol when humans are exposed to sufficient sunlight. Vitamin D_2, a compound related to vitamin D_3 (D_2 has an additional methyl group), is the additive in fortified milk. The active form of vitamin D_3, 1,25-dihydroxycholecalciferol, is formed from vitamin D_3 by two hydroxylation reactions (Figure 7.31); vitamin D_2 is similarly activated. The active compounds are hormones that help control Ca^{2+} utilization in humans—vitamin D regulates both intestinal absorption of calcium and its deposition in bones. In vitamin D–deficiency diseases, such as rickets in children and osteomalacia in adults, bones are weak because calcium phosphate does not properly crystallize on the collagen matrix of the bones.

C. Vitamin E

Vitamin E, or α-tocopherol (Figure 7.32), is one of several closely related tocopherols, compounds having a bicyclic oxygen-containing ring system with a hydrophobic side chain. The phenol group of vitamin E can undergo oxidation to a stable free radical. Vitamin E is believed to function as a reducing agent that scavenges oxygen and free radicals. This antioxidant action may prevent damage to fatty acids in biological membranes. A deficiency of vitamin E is rare but may lead to fragile red blood cells and neurological damage. The deficiency is almost always caused by genetic defects in absorption of fat molecules. There is currently no scientific evidence to support claims that vitamin E supplements in the diet of normal, healthy individuals will improve health.

D. Vitamin K

Phylloquinone (vitamin K) are important components of photosynthesis reaction centers in bacteria, algae, and plants.

Vitamin K (phylloquinone) (Figure 7.32) is a lipid vitamin from plants that is required for the synthesis of some of the proteins involved in blood coagulation. It is a coenzyme for a mammalian carboxylase that catalyzes the conversion of specific glutamate residues to γ-carboxyglutamate residues (Equation 7.7). The reduced (hydroquinone) form of vitamin K participates in the carboxylation as a reducing agent. Oxidized vitamin K has to be regenerated in order to support further modifications of clotting factors. This is accomplished by vitamin K reductase.

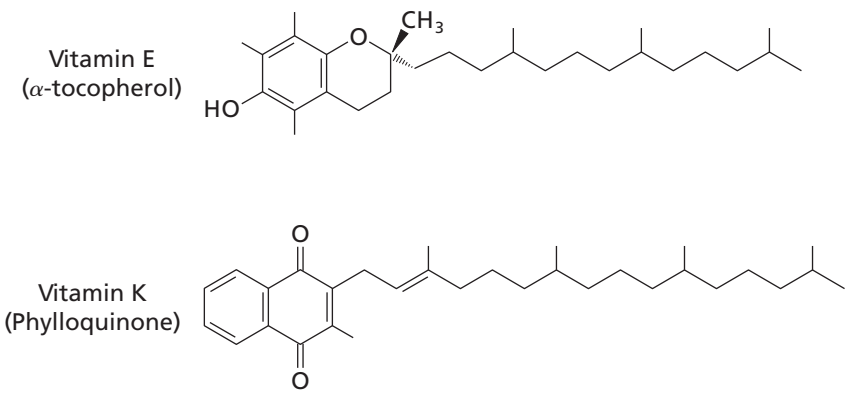


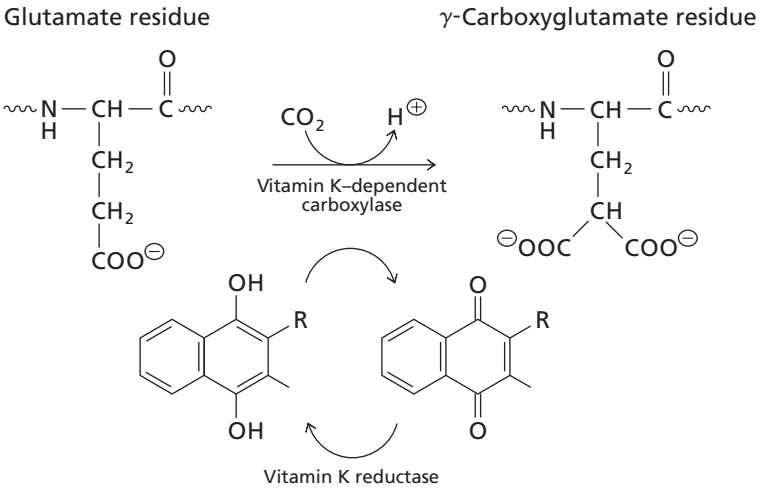
Figure 7.32 ►
Structures of vitamin E and vitamin K.

Downloaded From Vet-e-Books.com





▲ **Vitamin D and the evolution of skin color.** Black skin protects cells from damage by sunlight but it may inhibit formation of vitamin D. This isn't a problem in Nairobi, Kenya (left) but it might be in Stockholm, Sweden (right). One hypothesis for the evolution of skin color suggests that light-colored skin evolved in northern climates in order to increase vitamin D production.



(7.7)

When calcium binds to the γ-carboxyglutamate residues of the coagulation proteins, the proteins adhere to platelet surfaces where many steps of the coagulation process take place.

7.15 Ubiquinone

Ubiquinone—also called coenzyme Q and therefore abbreviated "Q"—is a lipid-soluble coenzyme synthesized by almost all species. Ubiquinone is a benzoquinone with four substituents, one of which is a long hydrophobic chain. This chain of 6 to 10 isoprenoid units allows ubiquinone to dissolve in lipid membranes. In the membrane, ubiquinone transports electrons between enzyme complexes. Some bacteria use menaquinone instead of ubiquinone (Figure 7.33 a). An analog of ubiquinone, plastoquinone (Figure 7.33b), serves a similar function in photosynthetic electron transport in chloroplasts (Chapter 15).

Ubiquinone is a stronger oxidizing agent than either $NAD^{\oplus}$ or the flavin coenzymes. Consequently, it can be reduced by NADH or $FADH_2$. Like FMN and FAD, ubiquinone can accept or donate two electrons one at a time because it has three oxidation states: oxidized Q, a partially reduced semiquinone free radical, and fully reduced QH_2, called ubiquinol (Figure 7.34). Coenzyme Q plays a major role in membrane-associated electron transport. It is responsible for moving protons from one side of the membrane to the other by a process known as the Q cycle. (Chapter 14). The resulting proton gradient contributes to ATP synthesis.

Downloaded From Vet-e-Books.com

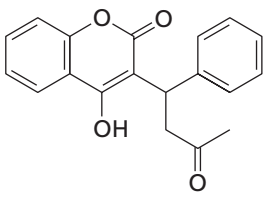
BOX 7.4 RAT POISON

Warfarin is an effective rat poison that has been used for many decades. It's a competitive inhibitor of vitamin K reductase, the enzyme that regenerates the reduced form of vitamin K (Equation 7.7). Blocking the formation of blood clotting factors leads to death in the rodents by internal bleeding. Rodents are very sensitive to inhibition of vitamin K reductase.

Later on it was discovered that low concentrations of warfarin were effective in individuals who suffer from excessive blood clotting. The drug was renamed (e.g., Coumadin®) for use in humans since its association with rat poison had a somewhat negative connotation.

Vitamin K analogs are widely used as anticoagulants in patients who are prone to thrombosis where they can prevent strokes and other embolisms. Like all medications, the dosage must be carefully regulated and controlled in order to prevent adverse effects, but in this case the dosage is even more critical.

Since the drugs only affect the synthesis of new clotting factors, they often take several days to have an effect.This is why patients will often be started at low dosages of these analogs and the amount of drug will be increased slowly over the course of many months.



▲ Warfarin.



▲ A rat (*Rattus norvegicus*).

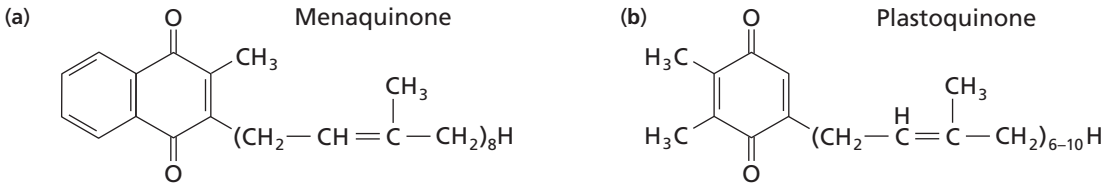


Figure 7.33 ►
Structures of (a) menaquinone and (b) plastoquinone. The hydrophobic tail of each molecule is composed of 6 to 10 five-carbon isoprenoid units.

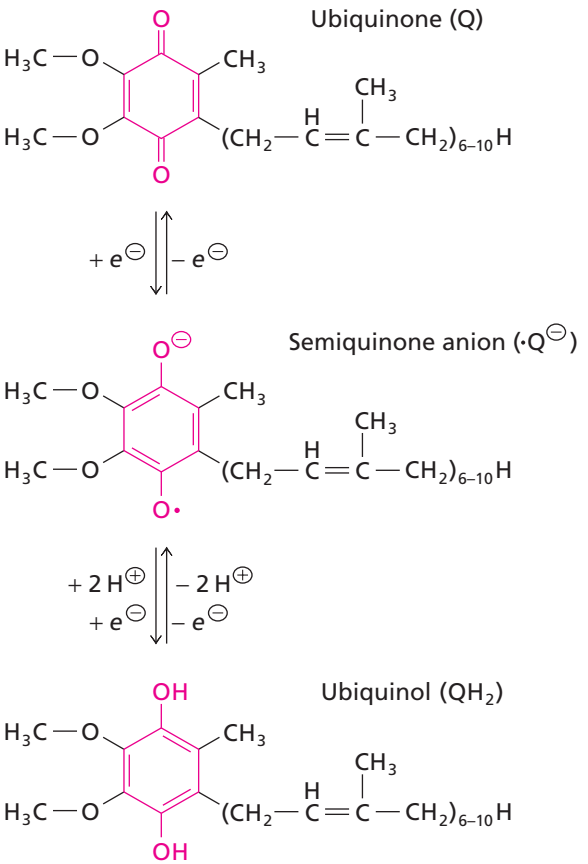


Figure 7.34 ►
Three oxidation states of ubiquinone. Ubiquinone is reduced in two one-electron steps via a semiquinone free-radical intermediate. The reactive center of ubiquinone is shown in red.

Downloaded From Vet-eBooks.com

Unlike FAD or FMN, ubiquinone and its derivatives cannot accept or donate a pair of electrons in a single step.

The strength of coenzyme oxidizing agents (standard reduction potential) is described in Section 10.9.

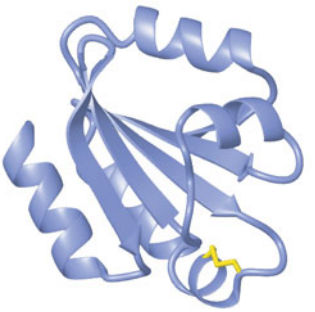
7.16 Protein Coenzymes

Some proteins act as coenzymes. They do not catalyze reactions by themselves but are required by certain other enzymes. These coenzymes are called either group transfer proteins or **protein coenzymes**. They contain a functional group either as part of their protein backbone or as a prosthetic group. Protein coenzymes are generally smaller and more heat-stable than most enzymes. They are called *coenzymes* because they participate in many different reactions and associate with a variety of different enzymes.

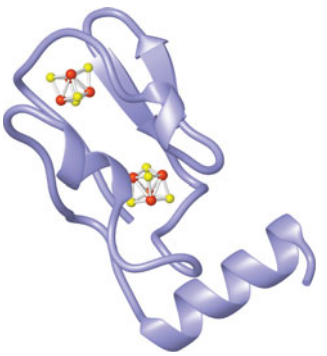
Some protein coenzymes participate in group transfer reactions or in oxidation–reduction reactions in which the transferred group is hydrogen or an electron. Metal ions, iron-sulfur clusters, and heme groups are reactive centers commonly found in these protein coenzymes. (Cytochromes are an important class of protein coenzymes that contain heme prosthetic groups. See Section 7.17.) Several protein coenzymes have two reactive thiol side chains that cycle between their dithiol and disulfide forms. For example, thioredoxins have cysteines three residues apart (—Cys—X—X—Cys—). The thiol side chains of these cysteine residues undergo reversible oxidation to form the disulfide bond of a cystine unit. We will encounter thioredoxins as reducing agents when we examine the citric acid cycle (Chapter 13), photosynthesis (Chapter 15), and deoxyribonucleotide synthesis (Chapter 18). The disulfide reactive center of thioredoxin is on the surface of the protein where it is accessible to the active sites of appropriate enzymes (Figure 7.35).

Ferredoxin is another common oxidation-reduction coenzyme. It contains two iron-sulfur clusters that can accept or donate electrons (Figure 7.36).

Some other protein coenzymes contain firmly bound coenzymes or portions of coenzymes. In *Escherichia coli*, a carboxyl carrier protein containing covalently bound biotin is one of three protein components of acetyl CoA carboxylase that catalyzes the first committed step of fatty acid synthesis. (In animal acetyl CoA carboxylases, the three protein components are fused into one protein chain.) ACP, introduced in Section 7.6, contains a phosphopantetheine moiety as its reactive center. The reactions of ACP therefore resemble those of coenzyme A. ACP is a component of all fatty acid synthases that have been tested. A protein coenzyme necessary for the degradation of glycine in mammals, plants, and bacteria (Chapter 17) contains a molecule of covalently bound lipoamide as a prosthetic group.



▲ **Figure 7.35**
Oxidized thioredoxin. Note that the cystine group is on the exposed surface of the protein. The sulfur atoms are shown in yellow. See Figure 4.24m for another view of thioredoxin. [PDB 1ERU].



▲ **Figure 7.36**
Ferredoxin. This ferredoxin from *Pseudomonas aeruginosa* contains two [4 Fe–4 S] iron-sulfur clusters that can be oxidized and reduced. Ferredoxin is a common cosubstrate in many oxidation–reduction reactions. [PDB 2FGO]

7.17 Cytochromes

Cytochromes are heme-containing protein coenzymes whose Fe(III) atoms undergo reversible one-electron reduction. Some structures of cytochromes were shown in Figures 4.21 and 4.24b. Cytochromes are classified as *a*, *b*, and *c* on the basis of their visible absorption spectra. The absorption spectra of reduced and oxidized cytochrome *c* are shown in Figure 7.37. Although the most strongly absorbing band is the Soret (or γ) band, the band labeled α is used to characterize cytochromes as either *a*, *b*, or *c*. Cytochromes in the same class may have slightly different spectra; therefore, a subscript number denoting the peak wavelength of the α absorption band of the reduced cytochrome often differentiates the cytochromes of a given class (e.g., cytochrome b_{560}). Wavelengths of maximum absorption for reduced cytochromes are given in Table 7.3.

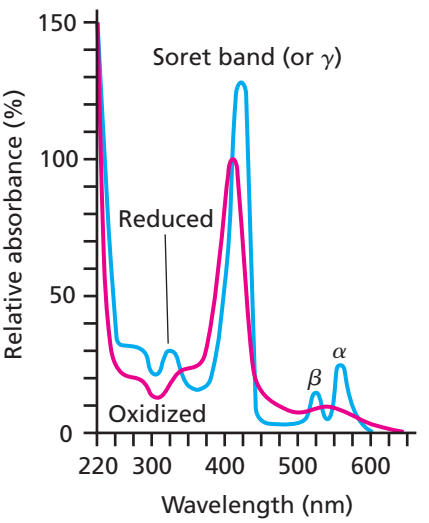


Figure 7.37 ►
Comparison of the absorption spectra of oxidized (red) and reduced (blue) horse cytochrome *c*. The reduced cytochrome has three absorbance peaks, designated α, β, and γ. On oxidation, the Soret (or γ) band decreases in intensity and shifts to a slightly shorter wavelength, whereas the α and β peaks disappear, leaving a single broad band of absorbance.

Downloaded From Vet-e-Books.com

Table 7.3 Absorption maxima (in nm) of major spectral bands in the visible absorption spectra of the reduced cytochromes

	Absorption band		
Heme protein	α	β	γ
Cytochrome *c*	550–558	521–527	415–423
Cytochrome *b*	555–567	526–546	408–449
Cytochrome *a*	592–604	Absent	439–443

The classes have slightly different heme prosthetic groups (Figure 7.38). The heme of *b*-type cytochromes is the same as that of hemoglobin and myoglobin (Figure 4.44). The heme of cytochrome *a* has a 17-carbon hydrophobic chain at C-2 of the porphyrin ring and a formyl group at C-8, whereas the *b*-type heme has a vinyl group attached to C-2 and a methyl group at C-8. In *c*-type cytochromes, the heme is covalently attached to the apoprotein by two thioether linkages formed by addition of the thiol groups of two cysteine residues to vinyl groups of the heme.

The tendency to transfer an electron to another substance, measured as a reduction potential, varies among individual cytochromes. The differences arise from the different environment each apoprotein provides for its heme prosthetic group. The reduction potentials of iron-sulfur clusters also vary widely depending on the chemical and physical environment provided by the apoprotein. The range of reduction potentials among prosthetic groups is an important feature of membrane-associated electron transport pathways (Chapter 14) and photosynthesis (Chapter 15).

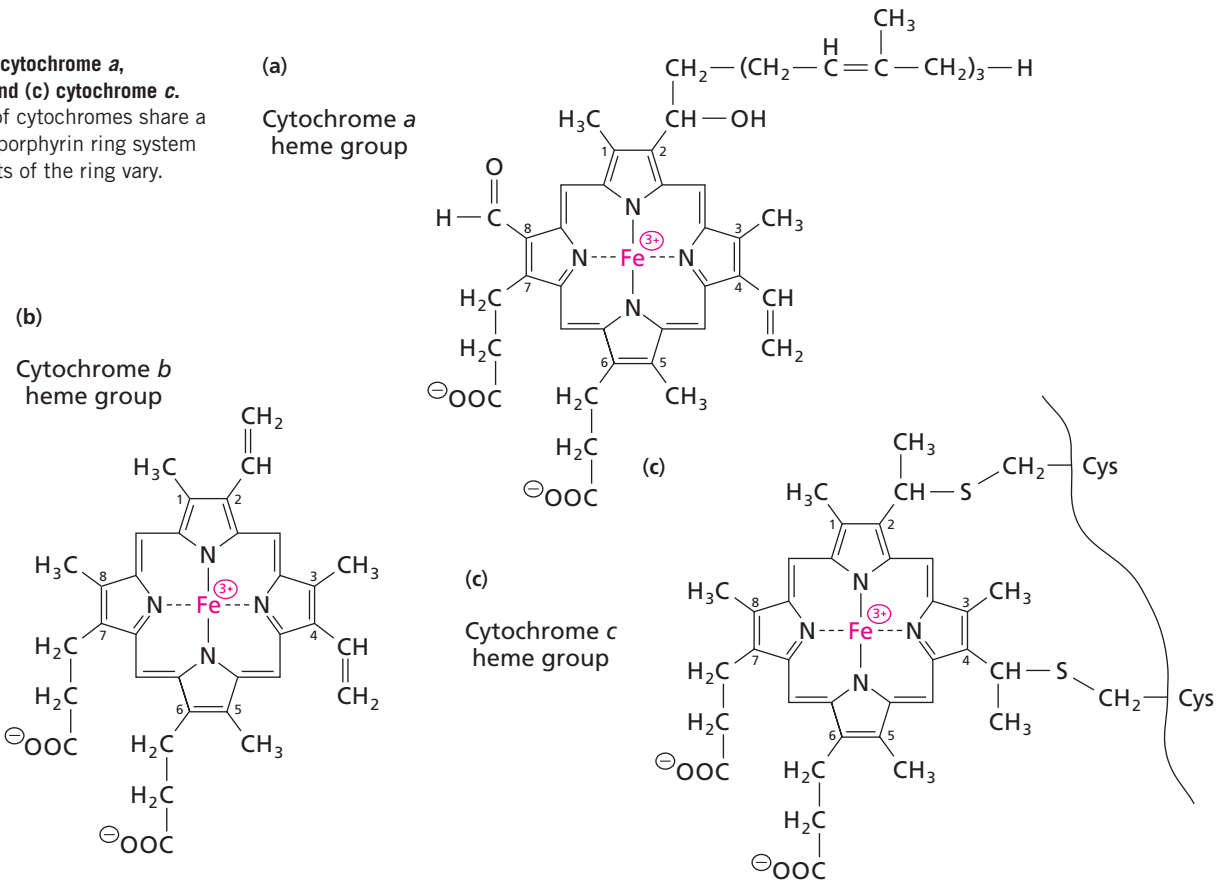


Figure 7.38 ▶
Heme groups of (a) cytochrome *a*, (b) cytochrome *b*, and (c) cytochrome *c*. The heme groups of cytochromes share a highly conjugated porphyrin ring system but the substituents of the ring vary.

Downloaded From Vet-e-Books.com

BOX 7.5 NOBEL PRIZES FOR VITAMINS AND COENZYMES

The discovery of vitamins in the first part of the 20th century stimulated an enormous amount of biochemistry research. What were these mysterious chemicals that seemed essential for life? Why were they essential?

We now take vitamins and coenzymes for granted but that doesn't do justice to the workers who discovered their role in metabolism. Here's a list of the scientists who received Nobel Prizes for their work on vitamins and coenzymes.

Chemistry 1928: **Adolf Otto Reinhold Windaus** "for the services rendered through his research into the constitution of the sterols and their connection with the vitamins."

Physiology or Medicine 1929: **Christiaan Eijkman** "for his discovery of the antineuritic vitamin." **Sir Frederick Gowland Hopkins** "for his discovery of the growth-stimulating vitamins."

Chemistry 1937: **Paul Karrer** "for his investigations on carotenoids, flavins and vitamins A and B_2." **Walter Norman Haworth** "for his investigations on carbohydrates and vitamin C."

Physiology or Medicine 1937: **Albert von Szent-Györgyi Nagyrapolt** "for his discoveries in connection with the biological combustion processes, with special reference to vitamin C and the catalysis of fumaric acid."

Chemistry 1938: **Richard Kuhn** "for his work on carotenoids and vitamins."

Physiology or Medicine 1943: **Henrik Carl Peter Dam** "for his discovery of vitamin K." **Edward Adelbert Doisy** "for his discovery of the chemical nature of vitamin K."

Physiology or Medicine 1953: **Fritz Albert Lipmann** "for his discovery of co-enzyme A and its importance for intermediary metabolism."

Chemistry 1964: **Dorothy Crowfoot Hodgkin** "for her determinations by X-ray techniques of the structures of important biochemical substances."

Chemistry 1970: **Luis F. Leloir** "for his discovery of sugar nucleotides and their role in the biosynthesis of carbohydrates."

Chemistry 1997: **Paul D. Boyer and John E. Walker** "for their elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP)."



▲ **Nobel Medals.** Chemistry (left), Physiology or Medicine (right).

Summary

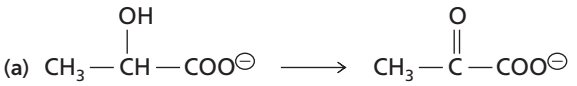
1. Many enzyme-catalyzed reactions require cofactors. Cofactors include essential inorganic ions and group-transfer reagents called coenzymes. Coenzymes can either function as cosubstrates or remain bound to enzymes as prosthetic groups.
2. Inorganic ions, such as $K^{\oplus}$, $Mg^{2\oplus}$, $Ca^{2\oplus}$, $Zn^{2\oplus}$, and $Fe^{3\oplus}$, may participate in substrate binding or in catalysis.
3. Some coenzymes are synthesized from common metabolites; others are derived from vitamins. Vitamins are organic compounds that must be supplied in small amounts in the diets of humans and other animals.
4. The pyridine nucleotides, $NAD^{\oplus}$ and $NADP^{\oplus}$, are coenzymes for dehydrogenases. Transfer of a hydride ion ($H^{\ominus}$) from a specific substrate reduces $NAD^{\oplus}$ or $NADP^{\oplus}$ to NADH or NADPH, respectively, and releases a proton.
5. The coenzyme forms of riboflavin—FAD and FMN—are tightly bound as prosthetic groups. FAD and FMN are reduced by hydride (two-electron) transfers to form $FADH_2$ and $FMNH_2$, respectively. The reduced flavin coenzymes donate electrons one or two at a time.
6. Coenzyme A, a derivative of pantothenate, participates in acyl-group–transfer reactions. Acyl carrier protein is required in the synthesis of fatty acids.
7. The coenzyme form of thiamine is thiamine diphosphate (TDP), whose thiazolium ring binds the aldehyde generated on decarboxylation of an α-keto acid substrate.
8. Pyridoxal 5′-phosphate is a prosthetic group for many enzymes in amino acid metabolism. The aldehyde group at C-4 of PLP forms a Schiff base with an amino acid substrate, through which it stabilizes a carbanion intermediate.
9. Vitamin C is a vitamin but not a coenzyme. It's a substrate in several reactions including those required in the synthesis of collagen. Vitamin C deficiency causes scurvy. Primates need an external source of vitamin C because they have lost one of the key enzymes required for its synthesis. The gene for this enzyme is a pseudogene in certain primate genomes.
10. Biotin, a prosthetic group for several carboxylases and carboxyltransferases, is covalently linked to a lysine residue at the enzyme active site.
11. Tetrahydrofolate is a reduced derivative of folate and participates in the transfer of one-carbon units at the oxidation levels of methanol, formaldehyde, and formic acid. Tetrahydrobiopterin is a reducing agent in some hydroxylation reactions.
12. The coenzyme forms of cobalamin—adenosylcobalamin and methylcobalamin—contain cobalt and a corrin ring system. These coenzymes participate in a few intramolecular rearrangements and methylation reactions.
13. Lipoamide, a prosthetic group for α-keto acid dehydrogenase multienzyme complexes, accepts an acyl group, forming a thioester.
14. The four fat-soluble, or lipid, vitamins are A, D, E, and K. These vitamins have diverse functions.

Downloaded From Vet-eBooks.com

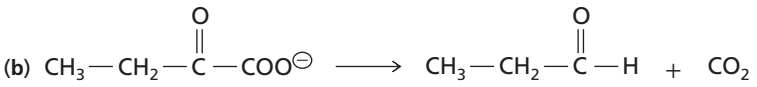
15. Ubiquinone is a lipid-soluble electron carrier that transfers electrons one or two at a time.
16. Some proteins, such as acyl carrier protein and thioredoxin, act as coenzymes in group-transfer reactions or in oxidation–reduction reactions in which the transferred group is hydrogen or an electron.
17. Cytochromes are small, heme-containing protein coenzymes that participate in electron transport. They are differentiated by their absorption spectra.

Problems

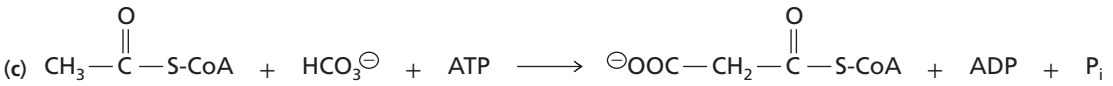
1. For each of the following enzyme-catalyzed reactions, determine the type of reaction and the coenzyme that is likely to participate.



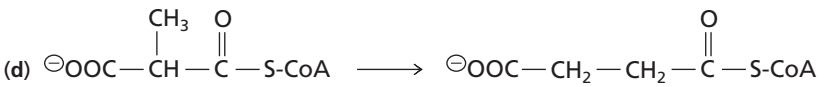
(a)



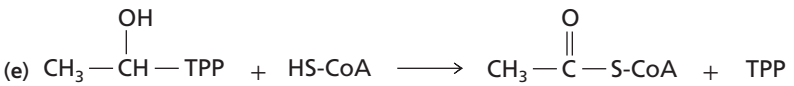
(b)



(c)

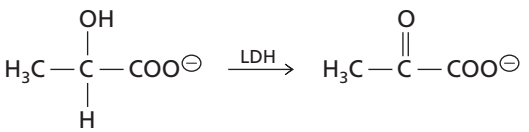


(d)



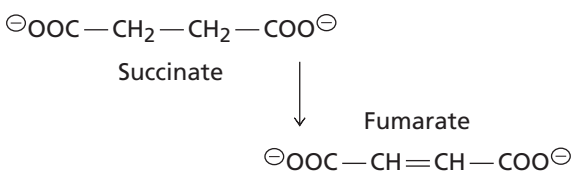
(e)

2. List the coenzymes that
   (a) participate as oxidation–reduction reagents.
   (b) act as acyl carriers.
   (c) transfer methyl groups.
   (d) transfer groups to and from amino acids.
   (e) are involved in carboxylation or decarboxylation reactions.
3. In the oxidation of lactate to pyruvate by lactate dehydrogenase (LDH), $NAD^{\oplus}$ is reduced in a two-electron transfer process from lactate. Since two protons are removed from lactate as well, is it correct to write the reduced form of the coenzyme as $NADH_2$? Explain.



L-Lactate Pyruvate

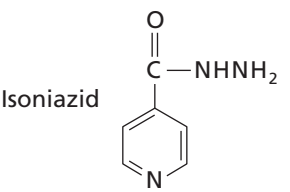
4. Succinate dehydrogenase requires FAD to catalyze the oxidation of succinate to fumarate in the citric acid cycle. Draw the isoalloxazine ring system of the cofactor resulting from the oxidation of succinate to fumarate and indicate which hydrogens in $FADH_2$ are lacking in FAD.



Succinate

Fumarate

5. What is the common structural feature of $NAD^{\oplus}$, FAD, and coenzyme A?
6. Certain nucleophiles can *add* to C-4 of the nicotinamide ring of $NAD^{\oplus}$, in a manner similar to the addition of a hydride in the reduction of $NAD^{\oplus}$ to NADH. Isoniazid is the most widely used drug for the treatment of tuberculosis. X-ray studies have shown that isoniazid inhibits a crucial enzyme in the tuberculosis bacterium where a covalent adduct is formed between the carbonyl of isoniazid and the 4′ position of the nicotinamide ring of a bound $NAD^{\oplus}$ molecule. Draw the structure of this NAD-isoniazid inhibitory adduct.

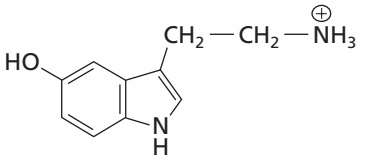


Isoniazid

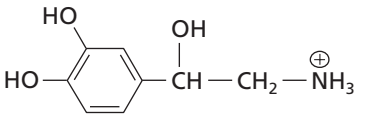
7. A vitamin B_6 deficiency in humans can result in irritability, nervousness, depression, and sometimes convulsions. These symptoms may result from decreased levels of the neurotransmitters serotonin and norepinephrine, which are metabolic derivatives of tryptophan and tyrosine, respectively. How could a deficiency of vitamin B_6 result in decreased levels of serotonin and norepinephrine?

Downloaded From Vet-e-Books.com

Serotonin



Norepinephrine



8. Macrocytic anemia is a disease in which red blood cells mature slowly due to a decreased rate of DNA synthesis. The red blood cells are abnormally large (macrocytic) and are more easily ruptured. How could the anemia be caused by a deficiency of folic acid?

9. A patient suffering from methylmalonic aciduria (high levels of methylmalonic acid) has high levels of homocysteine and low levels of methionine in the blood and tissues. Folic acid levels are normal.

(a) What vitamin is likely to be deficient?

(b) How could the deficiency produce the symptoms listed above?

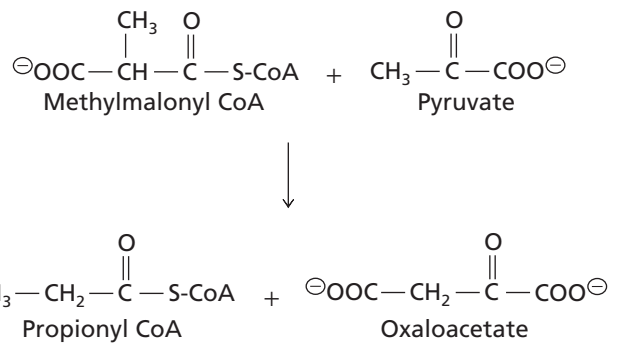
(c) Why is this vitamin deficiency more likely to occur in a person who follows a strict vegetarian diet?

10. Alcohol dehydrogenase (ADH) from yeast is a metalloenzyme that catalyzes the $NAD^{\oplus}$-dependent oxidation of ethanol to acetaldehyde. The mechanism of yeast ADH is similar to that of lactate dehydrogenase (LDH) (Figure 7.9) except that the zinc ion of ADH occupies the place of His-195 in LDH.

(a) Draw a mechanism for the oxidation of ethanol to acetaldehyde by yeast ADH.

(b) Does ADH require a residue analogous to Arg-171 in LDH?

11. In biotin-dependent transcarboxylase reactions, an enzyme transfers a carboxyl group between substrates in a two-step process without the need for ATP or bicarbonate. The reaction catalyzed by the enzyme methylmalonyl CoA-pyruvate transcarboxylase is shown below. Draw the structures of the products expected from the first step of the reaction.



12. (a) Histamine is produced from histidine by the action of a decarboxylase. Draw the external aldimine produced by the reaction of histidine and pyridoxal phosphate at the active site of histidine decarboxylase.

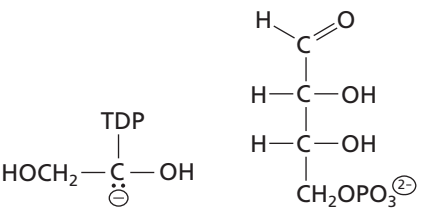
(b) Since racemization of amino acids by PLP-dependent enzymes proceeds via Schiff base formation, would racemization of L-histidine to D-histidine occur during the histidine decarboxylase reaction?

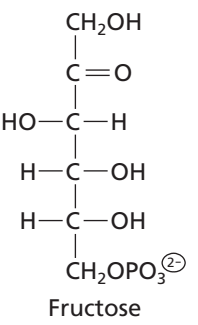
13. (a) Thiamine pyrophosphate is a coenzyme for oxidative decarboxylation reactions in which the keto carbonyl carbon is oxidized to an acid or an acid derivative. Oxidation occurs by removal of two electrons from a resonance-stabilized carbanion intermediate. What is the mechanism for the reaction pyruvate + HS-CoA → acetyl CoA + CO_2, beginning from the resonance-stabilized carbanion intermediate formed after decarboxylation (Figure 7.15) (such as a thioester in the case below)?

(b) Pyruvate dehydrogenase (PDH) is an enzyme complex that catalyzes the oxidative decarboxylation of pyruvate to acetyl CoA and CO_2 in a multistep reaction. The oxidation and acetyl-group transfer steps require TDP and lipoic acid in addition to other coenzymes. Draw the chemical structures for the molecules in the following two steps in the PDH reaction.

HETDP + lipoamide ⟶ acetyl-TDP + dihydrolipoamide ⟶
TDP + acetyl-dihydrolipoamide

(c) In a transketolase enzyme TDP-dependent reaction, the resonance-stabilized carbanion intermediate shown adjacent is generated as an intermediate. This intermediate is then involved in a condensation reaction (resulting in C—C bond formation) with the aldehyde group of erythrose 4-phosphate (E4P) to form fructose 6-phosphate (F6P). Starting from the carbanion intermediate, show a mechanism for this transketolase reaction. (Fischer projections of carbohydrate structures are sometimes drawn as shown here.)





Downloaded From Vet-e-Books.com

Selected Readings

Metal Ions

Berg, J. M. (1987). Metal ions in proteins: structural and functional roles. *Cold Spring Harbor Symp. Quant. Biol.* 52:579–585.

Rees, D. C. (2002). Great metalloclusters in enzymology. *Annu. Rev. Biochem.* 71: 221–246.

Specific Cofactors

Banerjee, R., and Ragsdale, S.W. (2003). The many faces of vitamin B_{12}: catalysis by cobalmin-dependent enzymes. *Annu. Rev. Biochem.* 72:209–247.

Bellamacina, C. R. (1996). The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J.* 10:1257–1268.

Blakley, R. L., and Benkovic, S. J., eds. (1985). *Folates and Pterins*, Vol. 1 and Vol. 2. (New York: John Wiley & Sons).

Chiang, P. K., Gordon, R. K., Tal, J., Zeng, G. C., Doctor, B. P., Pardhasaradhi, K., and McCann, P. P. (1996). *S*-Adenosylmethionine and methylation. *FASEB J.* 10:471–480.

Coleman, J. E. (1992). Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annu. Rev. Biochem.* 61:897–946.

Ghisla, S., and Massey, V. (1989). Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.* 181:1–17.

Hayashi, H., Wada, H., Yoshimura, T., Esaki, N., and Soda, K. (1990). Recent topics in pyridoxal 5′-phosphate enzyme studies. *Annu. Rev. Biochem.* 59:87–110.

Jordan, F. (1999). Interplay of organic and biological chemistry in understanding coenzyme mechanisms: example of thiamin diphosphate-dependent decarboxylations of 2-oxo acids. *FEBS Lett.* 457:298–301.

Jordan, F., Li, H., and Brown, A. (1999). Remarkable stabilization of zwitterionic intermediates may account for a billion-fold rate acceleration by thiamin diphosphate-dependent decarboxylases. *Biochem.* 38:6369–6373.

Jurgenson, C. T., Begley, T. P. and Ealick, S. E. (2009). The structural and biochemical foundations of thiamin biosynthesis. *Ann. Rev. Biochem.* 78:569–603.

Knowles, J. R. (1989). The mechanism of biotin-dependent enzymes. *Annu. Rev. Biochem.* 58:195–221.

Ludwig, M. L., and Matthews, R. G. (1997). Structure-based perspectives on B_{12}-dependent enzymes. *Annu. Rev. Biochem.* 66:269–313.

Palfey, B. A., Moran, G. R., Entsch, B., Ballou, D. P., and Massey, V. (1999). Substrate recognition by "password" in *p*-hydroxybenzoate hydroxylase. *Biochem.* 38:1153–1158.

NAD-Binding Motifs

Bellamacina, C. R. (1996). The nictotinamide d inucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J.* 10:1257–1269.

Rossman, M. G., Liljas, A., Brändén, C.-I., and Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes.* Vol. 11, Part A, 3rd ed., P. D., Boyer, ed. (New York: Academic Press), pp. 61–102.

Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R., and Holbrook, J. J. (1988). A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. *Science* 242:1541–1544.

Downloaded From Vet-eBooks.com





Carbohydrates

Molecular biology has dealt largely on the triad of DNA, RNA and protein. Biochemistry is concerned with all the molecules of the cell. Excluded from the province of molecular biology have been most of the structures and functions essential for growth and maintenance: carbohydrates, coenzymes, lipids, and membranes.

—Arthur Kornberg
"For the love of enzymes: the odyssey of a biochemist" (1989)

Carbohydrates (also called saccharides) are—on the basis of mass—the most abundant class of biological molecules on Earth. Although all organisms can synthesize carbohydrate, much of it is produced by photosynthetic organisms, including bacteria, algae, and plants. These organisms convert solar energy to chemical energy that is then used to make carbohydrate from carbon dioxide. Carbohydrates play several crucial roles in living organisms. In animals and plants, carbohydrate polymers act as energy storage molecules. Animals can ingest carbohydrates that can then be oxidized to yield energy for metabolic processes. Polymeric carbohydrates are also found in cell walls and in the protective coatings of many organisms. Other carbohydrate polymers are marker molecules that allow one type of cell to recognize and interact with another type. Carbohydrate derivatives are found in a number of biological molecules, including some coenzymes (Chapter 7) and the nucleic acids (Chapter 19).

The name *carbohydrate*, "hydrate of carbon," refers to their empirical formula $(CH_2O)_n$, where n is 3 or greater (n is usually 5 or 6 but can be up to 9). Carbohydrates can be described by the number of monomeric units they contain. **Monosaccharides** are the smallest units of carbohydrate structure. **Oligosaccharides** are polymers of two to about 20 monosaccharide residues. The most common oligosaccharides are disaccharides, which consist of two linked monosaccharide residues. **Polysaccharides** are polymers that contain many (usually more than 20) monosaccharide residues. Oligosaccharides and polysaccharides do not have the empirical formula $(CH_2O)_n$ because water is eliminated during polymer formation. The term *glycan* is a more general term for carbohydrate polymers. It can refer to a polymer of identical sugars (homoglycan) or of different sugars (heteroglycan).

Glycoconjugates are carbohydrate derivatives in which one or more carbohydrate chains are linked covalently to a peptide, protein, or lipid. These derivatives include proteoglycans, peptidoglycans, glycoproteins, and glycolipids.

In this chapter, we discuss nomenclature, structure, and function of monosaccharides, disaccharides, and the major homoglycans—starch, glycogen, cellulose, and

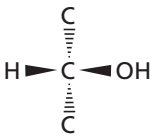
Photosynthesis is described in detail in Chapter 15.

Top: Darkling beetle. The exoskeletons of insects contain chitin, a homoglycan.

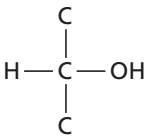
Downloaded From Vet-e-Books.com

KEY CONCEPT

A Fischer projection is a convention designed to convey information about the stereochemistry of a molecule. It does not resemble the actual conformation of the molecule in solution.



Stereo view



Fischer projection

For *each* chiral carbon atom in a Fischer projection the vertical bonds project into the plane of the page and the horizontal bonds project upward toward the viewer.

chitin. We then consider proteoglycans, peptidoglycans, and glycoproteins, all of which contain heteroglycan chains.

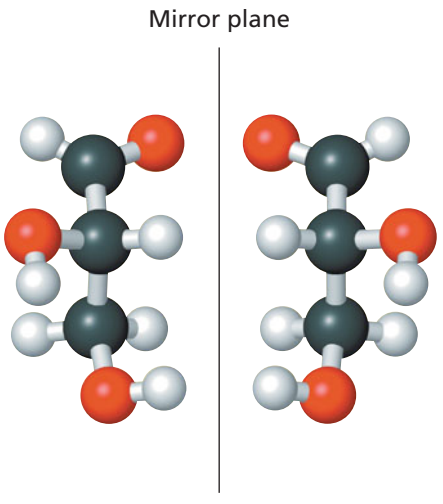
8.1 Most Monosaccharides Are Chiral Compounds

Monosaccharides are water-soluble, white, crystalline solids that have a sweet taste. Examples include glucose and fructose. Chemically, monosaccharides are polyhydroxy aldehydes, or **aldoses**, or polyhydroxy ketones, or **ketoses**. They are classified by their type of carbonyl group and their number of carbon atoms. As a rule, the suffix *-ose* is used in naming carbohydrates, although there are a number of exceptions. All monosaccharides contain at least three carbon atoms. One of these is the carbonyl carbon, and each of the remaining carbon atoms bears a hydroxyl group. In aldoses, the most oxidized carbon atom is designated C-1 and is drawn at the top of a Fischer projection. In ketoses, the most oxidized carbon atom is usually C-2.

We've encountered Fischer projections before but now it's time to present the convention in more detail. A Fischer projection is a two-dimensional representation of a three-dimensional molecule. It is designed to preserve information about the stereochemistry of a molecule. In a Fischer projection of sugars, the C-1 atom is always at the top of the figure. For each separate chiral carbon atom, the two horizontal bonds project upward from the page toward you. The two vertical bonds project downward into the page. Remember, this applies to each chiral carbon atom, so in a carbohydrate with multiple carbon atoms the Fischer projection represents a molecule that curls back into the page. For longer molecules, the top and bottom groups may even come in virtual contact, forming a loop. The Fischer projection is a convention for preserving stereochemical information; it does not represent a realistic model of how a molecule might look in solution.

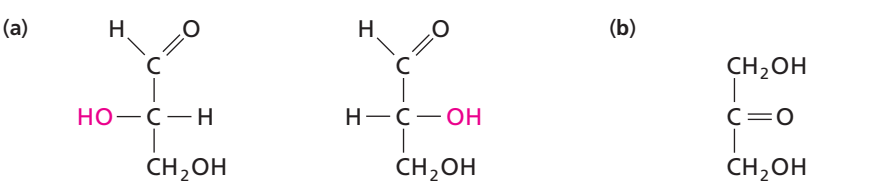
The smallest monosaccharides are **trioses**, or three-carbon sugars. One- or two-carbon compounds having the general formula $(CH_2O)_n$ do not have properties typical of carbohydrates (such as sweet taste and the ability to crystallize). The aldehydic triose, or aldotriose, is glyceraldehyde (Figure 8.1a). Glyceraldehyde is chiral because its central carbon, C-2, has four different groups attached to it, (Section 3.1). The ketonic triose, or ketotriose, is dihydroxyacetone (Figure 8.1b). It is achiral because it has no asymmetric carbon atom. All other monosaccharides, longer-chain versions of these two sugars, are chiral.

The stereoisomers D- and L-glyceraldehyde are shown as ball-and-stick models in Figure 8.2. Chiral molecules are optically active; that is, they rotate the plane of polarized light. The convention for designating D and L isomers was originally based on the optical properties of glyceraldehyde. The form of glyceraldehyde that caused rotation to the right (dextrorotatory) was designated D and the form that caused rotation to the left (levorotatory) was designated L. Structural knowledge was limited when this convention was established in the late 19th century so the configurations for the enantiomers of glyceraldehyde were assigned arbitrarily, with a 50% probability of error. X-ray crystallographic experiments later proved that the original structural assignments were correct.



L-Glyceraldehyde D-Glyceraldehyde

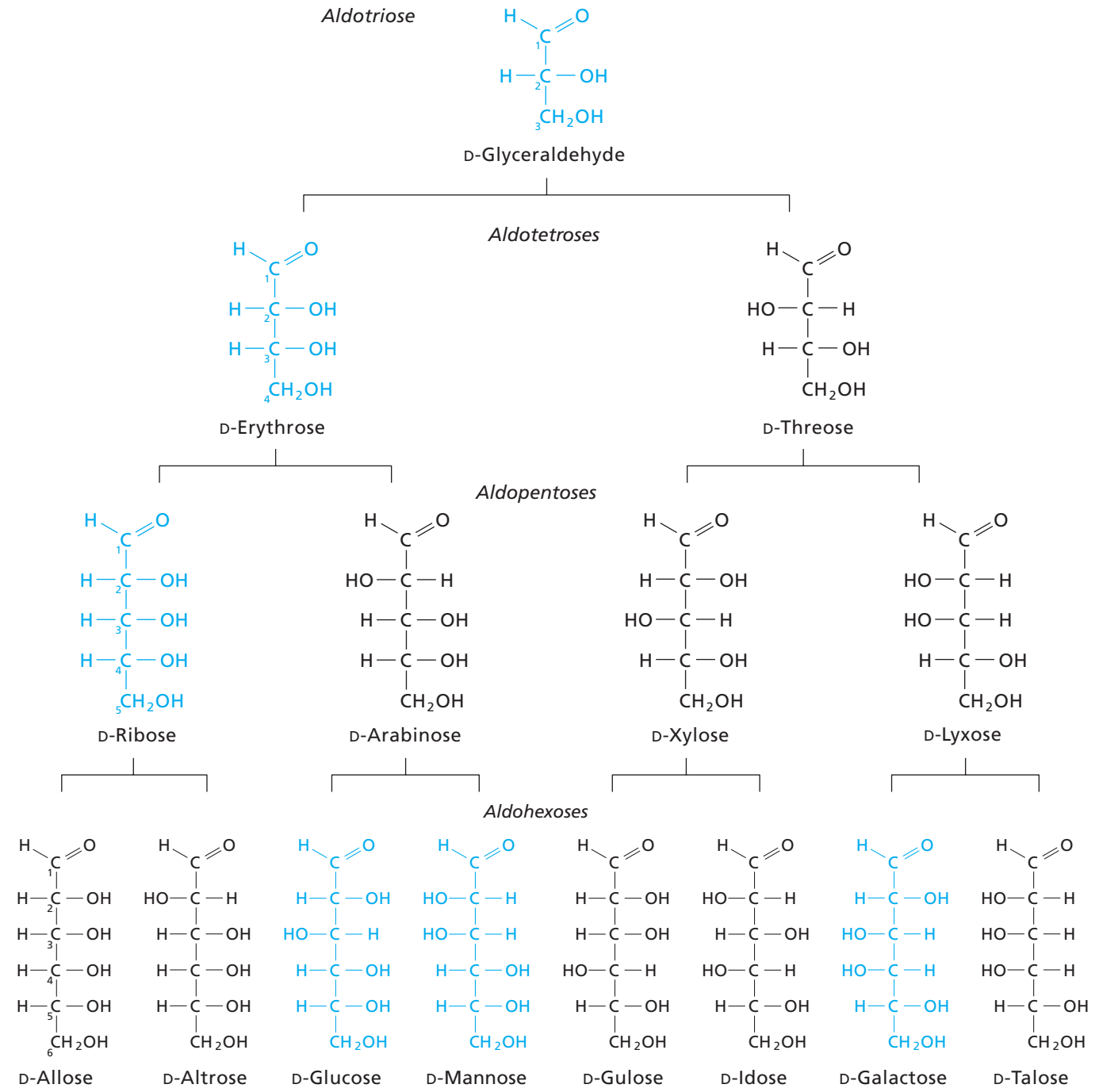
▲ **Figure 8.2**
View of L-glyceraldehyde (left) and D-glyceraldehyde (right). These molecules are drawn in a conformation that corresponds to the Fischer projections in Figure 8.1.



L-Glyceraldehyde D-Glyceraldehyde Dihydroxyacetone

▲ **Figure 8.1**
Fischer projections of (a) glyceraldehyde and (b) dihydroxyacetone. The designations L (for left) and D (for right) for glyceraldehyde refer to the configuration of the hydroxyl group of the chiral carbon (C-2). Dihydroxyacetone is achiral.

Downloaded From Vet-e-Books.com



▲ **Figure 8.3**
Fischer projections of the three- to six-carbon D-aldoses. The aldoses shown in blue are the most important in our study of biochemistry.

Longer aldoses and ketoses can be regarded as extensions of glyceraldehyde and dihydroxyacetone, respectively, with chiral H—C—OH groups inserted between the carbonyl carbon and the primary alcohol group. Figure 8.3 shows the complete list of the names and structures of the tetroses (four-carbon aldoses), pentoses (five-carbon aldoses), and hexoses (six-carbon aldoses) related to D-glyceraldehyde. Many of these monosaccharides are not synthesized by most organisms and we will not encounter them again in this book.

Note that the carbon atoms are numbered from the carbon of the aldehyde group that is assigned the number 1. By convention, sugars are said to have the D configuration when the configuration of the chiral carbon with the highest number—the chiral carbon most distant from the carbonyl carbon—is the same as that of C-2 of D-glyceraldehyde

Downloaded From Vet-e-Books.com

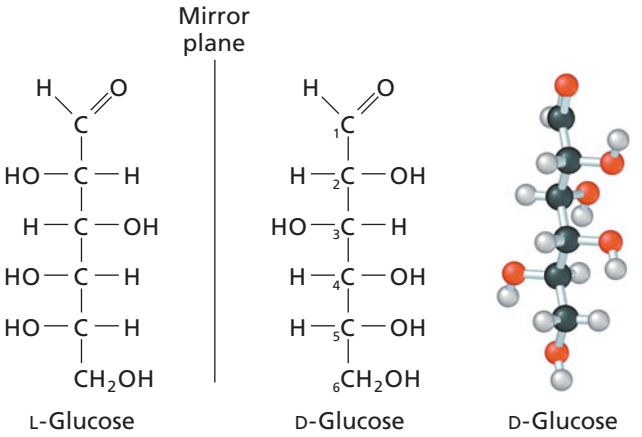


Figure 8.4 ►
L- and D-glucose. Fischer projections (left) showing that L- and D-glucose are mirror images. Conformation of the extended form of D-glucose in solution.

(i.e., the —OH group attached to this carbon atom is on the right side in a Fischer projection). The arrangement of asymmetric carbon atoms is unique for each monosaccharide, giving each its distinctive properties. Except for glyceraldehyde (which was used as the standard), there is no predictable association between the absolute configuration of a sugar and whether it is dextrorotatory or levorotatory.

It is mostly the D enantiomers that are synthesized in living cells—just as the L enantiomers of amino acids are more common. The L enantiomers of the 15 aldoses in Figure 8.3 are not shown. Recall that pairs of enantiomers are mirror images; in other words, the configuration at each chiral carbon is opposite. For example, the hydroxyl groups bound to carbon atoms 2, 3, 4, and 5 of D-glucose point right, left, right, and right, respectively, in the Fischer projection; those of L-glucose point left, right, left, and left (Figure 8.4).

The three-carbon aldose, glyceraldehyde, has only a single chiral atom (C-2) and therefore only two stereoisomers. There are four stereoisomers for aldotetroses (D- and L-erythrose and D- and L-threose) because erythrose and threose each possess two chiral carbon atoms. In general, there are 2^n possible stereoisomers for a compound with n chiral carbons. Aldohexoses, which possess four chiral carbons, have a total of 2^4, or 16, stereoisomers (the eight D aldohexoses in Figure 8.3 and their L enantiomers).

Sugar molecules that differ in configuration at only one of several chiral centers are called **epimers**. For example, D-mannose and D-galactose are epimers of D-glucose (at C-2 and C-4, respectively), although they are not epimers of each other (Figure 8.3).

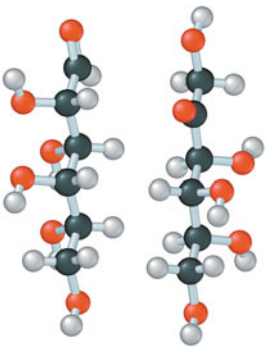
Longer-chain ketoses (Figure 8.5) are related to dihydroxyacetone in the same way that longer-chain aldoses are related to glyceraldehyde. Note that a ketose has one fewer chiral carbon atom than the aldose of the same empirical formula. For example, there are only two stereoisomers for the one ketotetrose (D- and L-erythrulose), and four stereoisomers for ketopentoses (D- and L-xylulose and D- and L-ribulose). Ketotetrose and ketopentoses are named by inserting *-ul-* in the name of the corresponding aldose. For example, the ketose xylulose corresponds to the aldose xylose. This nomenclature does not apply to the ketohexoses (tagatose, sorbose, psicose, and fructose) because they have traditional (trivial) names.

8.2 Cyclization of Aldoses and Ketoses

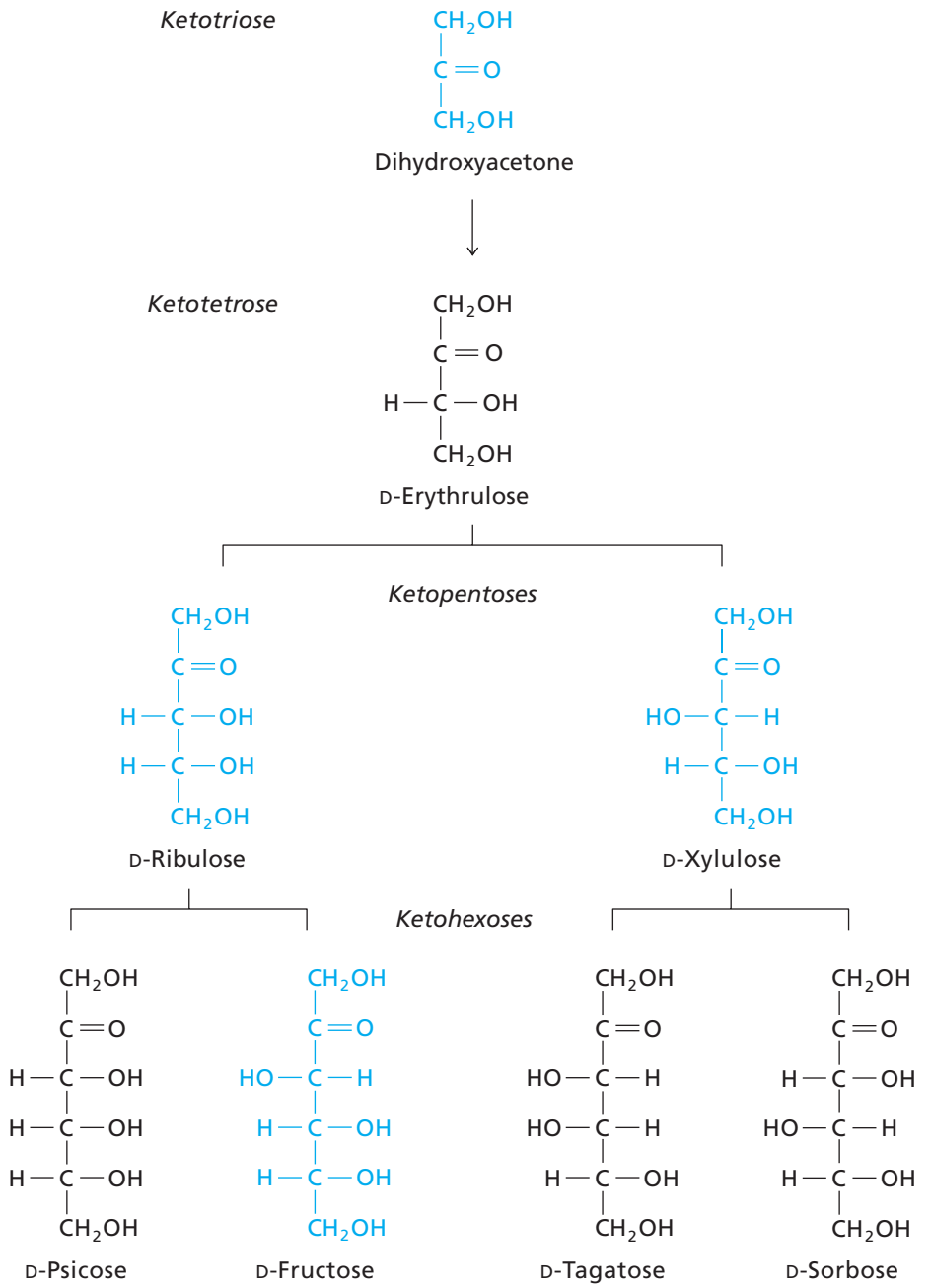
The optical behavior of some monosaccharides suggests they have one more chiral carbon atom than is evident from the structures shown in Figures 8.3 and 8.5. D-Glucose, for example, exists in two forms that contain five (not four) asymmetric carbons. The source of this additional asymmetry is an intramolecular cyclization reaction that produces a new chiral center at the carbon atom of the carbonyl group. This cyclization resembles the reaction of an alcohol with an aldehyde to form a hemiacetal or with a ketone to form a hemiketal (Figure 8.7).

The carbonyl carbon of an aldose containing at least five carbon atoms or of a ketose containing at least six carbon atoms can react with an intramolecular hydroxyl

Downloaded From Vet-eBooks.com



▲ **Who am I?** The structures of the D sugars are shown in Figures 8.3 and 8.5. You can deduce the structures of the L configurations. Knowing the convention for Fischer projections, you should have no trouble identifying these molecules.



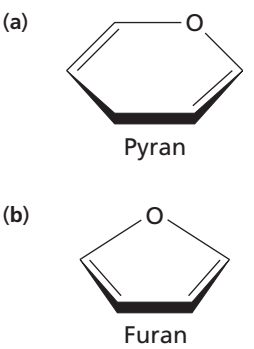
◄ **Figure 8.5**
Fischer projections of the three- to six-carbon D-ketoses. The ketoses shown in blue are the most important in our study of biochemistry.

group to form a cyclic hemiacetal or cyclic hemiketal, respectively. The oxygen atom from the reacting hydroxyl group becomes a member of the five- or six-membered ring structures (Figure 8.8).

Because it resembles the six-membered heterocyclic compound pyran (Figure 8.6a), the six-membered ring of a monosaccharide is called a **pyranose**. Similarly, because the five-membered ring of a monosaccharide resembles furan (Figure 8.6b), it is called a **furanose**. Note that, unlike pyran and furan, the rings of carbohydrates do not contain double bonds.

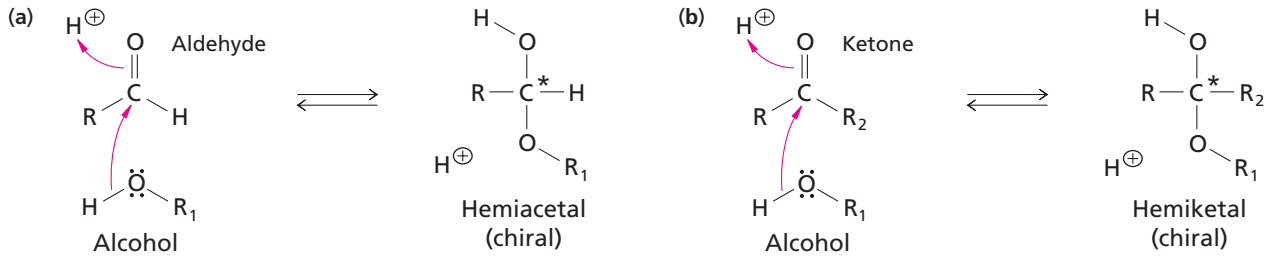
The most oxidized carbon of a cyclized monosaccharide, the one attached to two oxygen atoms, is referred to as the **anomeric carbon**. In ring structures, the anomeric carbon is chiral. Thus, the cyclized aldose or ketose can adopt either of two configurations (designated α or β), as illustrated for D-glucose in Figure 8.8. The α and β isomers are called **anomers**.

In solution, aldoses and ketoses that form ring structures equilibrate among their various cyclic and open-chain forms. At 31°C, for example, D-glucose exists in an equilibrium



▲ **Figure 8.6**
(a) Pyran and (b) furan.

Downloaded From Vet-e-Books.com

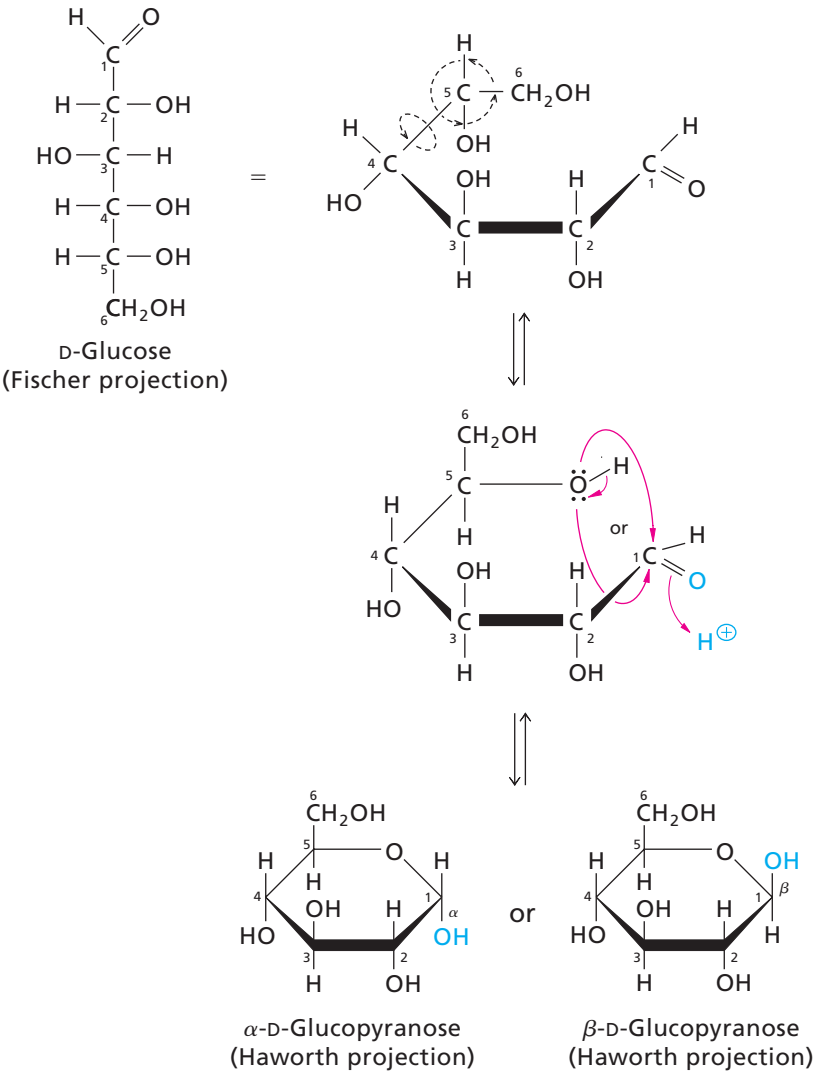


▲ **Figure 8.7**
Hemiacetal and hemiketal. (a) Reaction of an alcohol with an aldehyde to form a hemiacetal. **(b)** Reaction of an alcohol with a ketone to form a hemiketal. The asterisks indicate the newly formed chiral centers.

mixture of approximately 64% β-D-glucopyranose and 36% α-D-glucopyranose, with very small amounts of the furanose (Figure 8.9) and open-chain (Figure 8.4) forms. Similarly, D-ribose exists as a mixture of approximately 58.5% β-D-ribopyranose, 21.5% α-D-ribopyranose, 13.5% β-D-ribofuranose, and 6.5% α-D-ribofuranose, with a tiny fraction in the open-chain form (Figure 8.10). The relative abundance of the various forms of monosaccharides at equilibrium reflects the relative stabilities of each form. Although unsubstituted D-ribose is most stable as the β-pyranose, its structure in nucleotides (Section 8.5c) is the β-furanose form.

The ring drawings shown in these figures are called Haworth projections, after Norman Haworth who worked on the cyclization reactions of carbohydrates and first

Figure 8.8 ►
Cyclization of D-glucose to form glucopyranose. The Fischer projection (top left) is rearranged into a three-dimensional representation (top right). Rotation of the bond between C-4 and C-5 brings the C-5 hydroxyl group close to the C-1 aldehyde group. Reaction of the C-5 hydroxyl group with one side of C-1 gives α-D-glucopyranose; reaction of the hydroxyl group with the other side gives β-D-glucopyranose. The glucopyranose products are shown as Haworth projections in which the lower edges of the ring (thick lines) project in front of the plane of the paper and the upper edges project behind the plane of the paper. In the α-D-anomer of glucose, the hydroxyl group at C-1 points down; in the β-D-anomer, it points up.



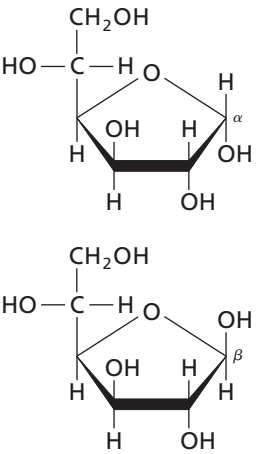
Downloaded From Vet-e-Books.com

proposed these representations. He received the Nobel Prize in Chemistry in 1937 for his work on carbohydrate structure and the synthesis of vitamin C.

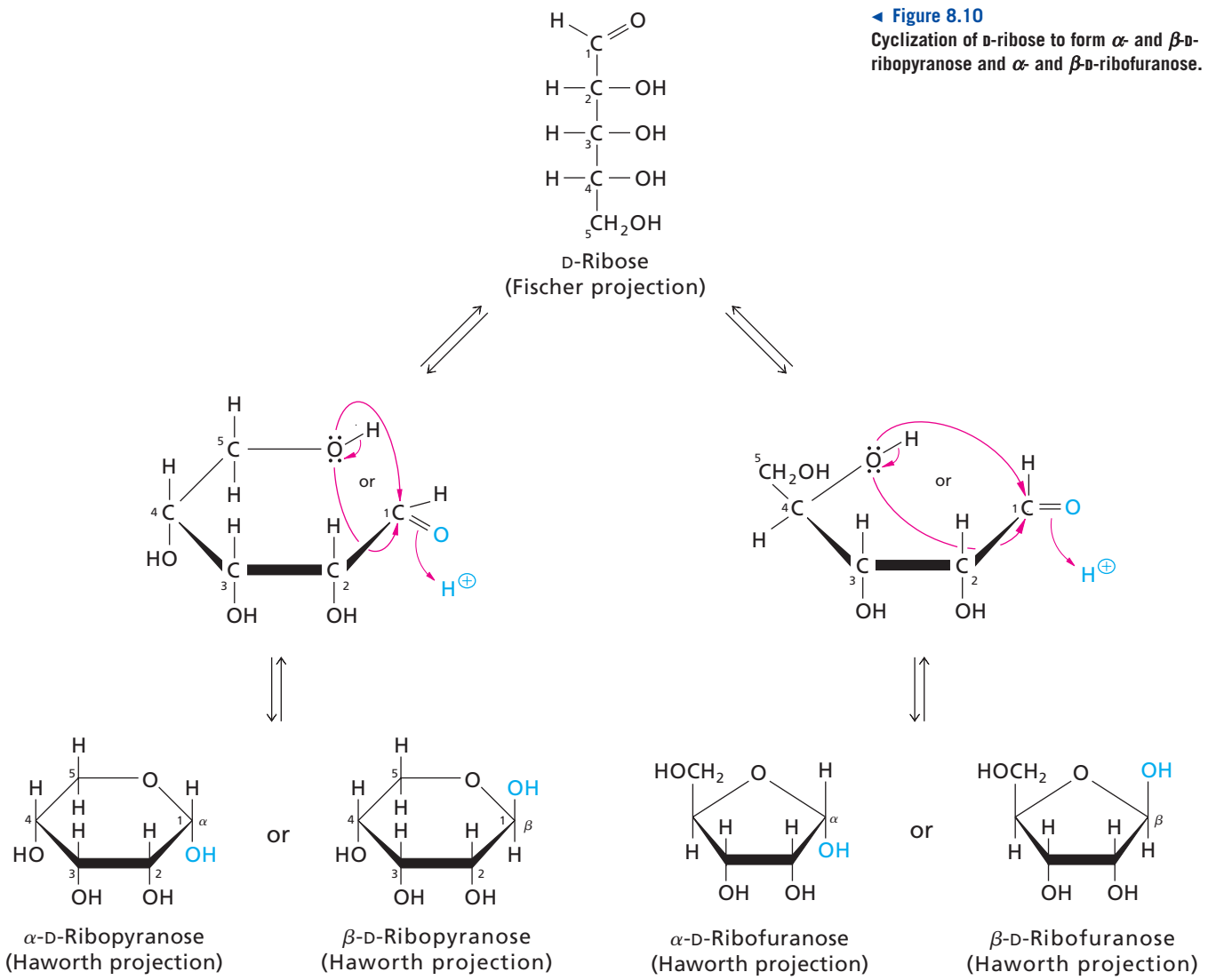
A Haworth projection adequately indicates stereochemistry and can be easily related to a Fischer projection: groups on the right in a Fischer projection point downwards in a Haworth projection. Because rotation around carbon–carbon bonds is constrained in the ring structure, the Haworth projection is a much more faithful representation of the actual conformation of sugars.

By convention, a cyclic monosaccharide is drawn so the anomeric carbon is on the right and the other carbons are numbered in a clockwise direction. In a Haworth projection, the configuration of the anomeric carbon atom is designated α if its hydroxyl group is *cis* to (on the same side of the ring as) the oxygen atom of the highest-numbered chiral carbon atom. It is β if its hydroxyl group is *trans* to (on the opposite side of the ring from) the oxygen attached to the highest-numbered chiral carbon. With α-D-glucopyranose, the hydroxyl group at the anomeric carbon points down; with β-D-glucopyranose, it points up.

Monosaccharides are often drawn in either the α- or β-D-furanose or the α- or β-D-pyranose form. However, you should remember that the anomeric forms of five- and six-carbon sugars are in rapid equilibrium. Throughout this chapter and the rest of the book, we draw sugars in the correct anomeric form if it is known. We refer to sugars in a nonspecific way (e.g., glucose) when we are discussing an equilibrium

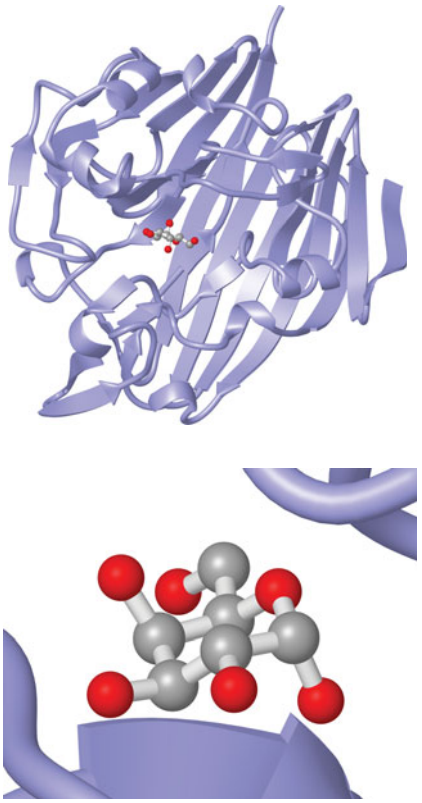


▲ **Figure 8.9**
α-D-glucofuranose (top) and β-D-glucofuranose (bottom).



◀ **Figure 8.10**
Cyclization of D-ribose to form α- and β-D-ribopyranose and α- and β-D-ribofuranose.

Downloaded From Vet-e-Books.com



▲ **Galactose mutarotase.** Mutarotases are enzymes that catalyze the interconversion of α and β configurations. This interconversion involves the breaking and remaking of covalent bonds, which is why they are different configurations. The enzyme shown here is galactose mutarotase from *Lactococcus lactis* with a molecule of α-D-galactose in the acitve site. The bottom figure shows the *conformation* of this molecule. Can you identify this conformation? [PDB 1L7K]

mixture of the various anomeric forms as well as the open-chain forms. When we are discussing a specific form of a sugar, however, we will refer to it precisely (e.g., β-D-glucopyranose). Also, since the D enantiomers of carbohydrates predominate in nature, we always assume that a carbohydrate has the D configuration unless specified otherwise.

8.3 Conformations of Monosaccharides

Haworth projections are commonly used in biochemistry because they accurately depict the *configuration* of the atoms and groups at each carbon atom of the sugar's backbone. However, the geometry of the carbon atoms of a monosaccharide ring is tetrahedral (bond angles near 110°), so monosaccharide rings are not actually planar. Cyclic monosaccharides can exist in a variety of *conformations* (three-dimensional shapes having the same configuration). Furanose rings adopt envelope conformations in which one of the five ring atoms (either C-2 or C-3) is out-of-plane and the remaining four are approximately coplanar (Figure 8.11). Furanoses can also form twist conformations where two of the five ring atoms are out-of-plane—one on either side of the plane formed by the other three atoms. The relative stability of each conformer depends on the degree of steric interference between the hydroxyl groups. The various conformers of unsubstituted monosaccharides can rapidly interconvert.

Pyranose rings tend to assume one of two conformations, the chair conformation or the boat conformation (Figure 8.12). There are two distinct chair conformers and six distinct boat conformers for each pyranose. The chair conformations minimize steric repulsion among the ring substituents and are generally more stable than boat conformations. The —H, —OH, and —CH_2OH substituents of a pyranose ring in the chair conformation may occupy two different positions. In the axial position the substituent is above or below the plane of the ring, while in the equatorial position the substituent lies in the plane of the ring. In pyranoses, five substituents are axial and five are equatorial. Whether a group is axial or equatorial depends on which carbon atom (C-1 or C-4) extends above the plane of the ring when the ring is in the chair conformation. Figure 8.13 shows the two different chair conformers of β-D-glucopyranose. The more stable conformation is the one in which the bulkiest ring substituents are equatorial (top structure). In fact, this conformation of β-D-glucose has the least steric strain of any aldohexose. Pyranose rings are occasionally forced to adopt slightly different conformations, such as the unstable half-chair adopted by a polysaccharide residue in the active site of lysozyme (Section 6.6).

KEY CONCEPT

Different *configurations* can only be formed by breaking and reforming covalent bonds. Molecules can adopt different *conformations* without breaking covalent bonds.

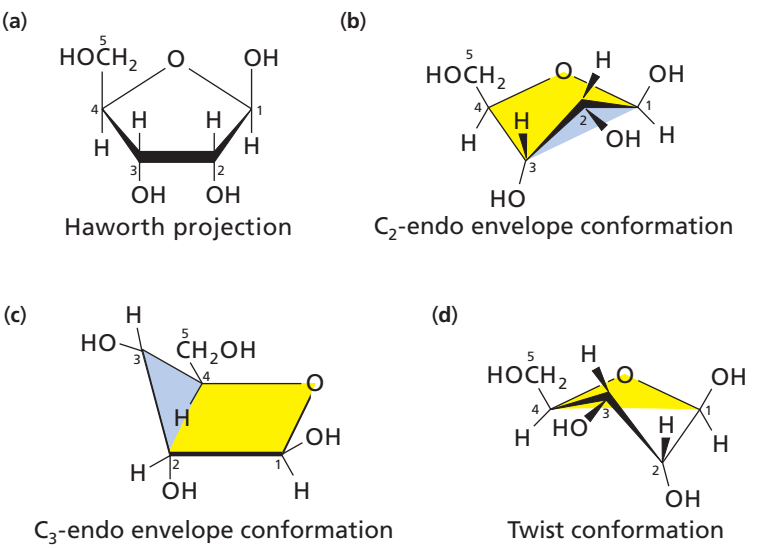
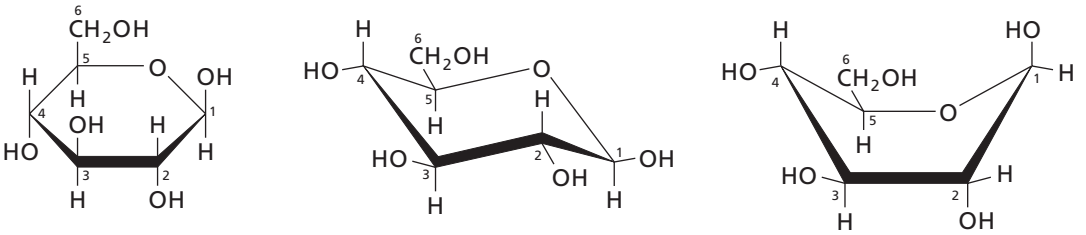


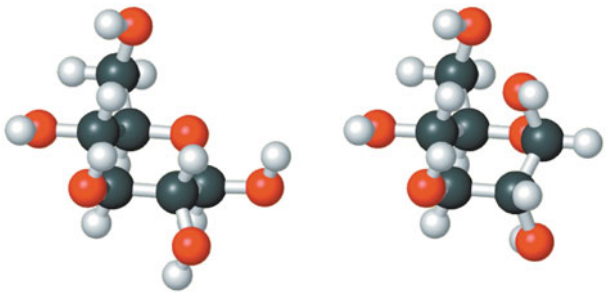
Figure 8.11 ►
Conformations of β-D-ribofuranose. (a) Haworth projection. **(b)** C_2-endo envelope conformation. **(c)** C_3-endo envelope conformation. **(d)** Twist conformation. In the C_2-endo conformation, C-2 lies above the plane defined by C-1, C-3, C-4, and the ring oxygen. In the C_3-endo conformation, C-3 lies above the plane defined by C-1, C-2, C-4, and the ring oxygen. In the twist conformation shown, C-3 lies above and C-2 lies below the plane defined by C-1, C-4, and the ring oxygen. The planes are shown in yellow.

Downloaded From Vet-e-Books.com

(a)



(b)

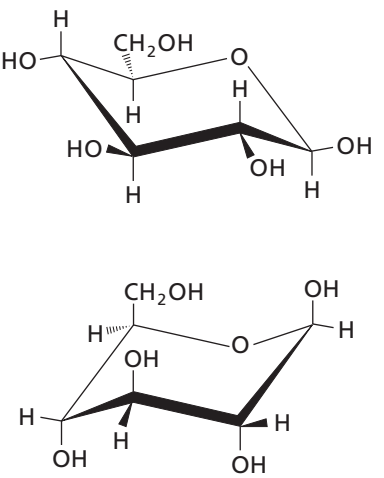


◀ **Figure 8.12**
Conformations of β-D-glucopyranose. **(a)** Haworth projection, a chair conformation, and a boat conformation. **(b)** Ball-and-stick model of a chair (left) and a boat (right) conformation.

8.4 Derivatives of Monosaccharides

There are many known derivatives of the basic monosaccharides. They include polymerized monosaccharides, such as oligosaccharides and polysaccharides, as well as several classes of nonpolymerized compounds. In this section, we introduce a few monosaccharide derivatives, including sugar phosphates, deoxy and amino sugars, sugar alcohols, and sugar acids.

Like other polymer-forming biomolecules, monosaccharides and their derivatives have abbreviations used in describing more complex polysaccharides. The accepted abbreviations contain three letters, with suffixes added in some cases. The abbreviations for some pentoses and hexoses and their major derivatives are listed in Table 8.1. We use these abbreviations later in this chapter.



▲ **Figure 8.13**
The two chair conformers of β-D-glucopyranose. The top conformer is more stable.

A. Sugar Phosphates

Monosaccharides are often converted to phosphate esters. Figure 8.14 shows the structures of several of the sugar phosphates we will encounter in our study of carbohydrate metabolism. The triose phosphates, ribose 5-phosphate, and glucose 6-phosphate are simple alcohol-phosphate esters. Glucose 1-phosphate is a hemiacetal phosphate, which is more reactive than an alcohol phosphate. The ability of UDP-glucose to act as a glucosyl donor (Section 7.3) is evidence of this reactivity.

B. Deoxy Sugars

The structures of two deoxy sugars are shown in Figure 8.15. In these derivatives, a hydrogen atom replaces one of the hydroxyl groups in the parent monosaccharide. 2-Deoxy-D-ribose is an important building block for DNA. L-Fucose (6-deoxy-L-galactose) is widely distributed in plants, animals, and microorganisms. Despite its unusual L configuration, fucose is derived metabolically from D-mannose.

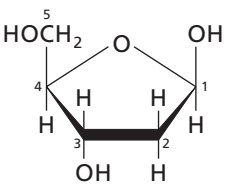
C. Amino Sugars

In a number of sugars, an amino group replaces one of the hydroxyl groups in the parent monosaccharide. Sometimes the amino group is acetylated. Three examples of amino

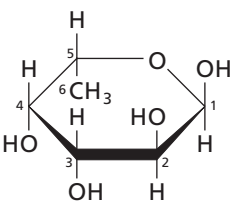
Downloaded From Vet-e-Books.com

Table 8.1 Abbreviations for some monosaccharides and their derivatives

Monosaccharide or derivative	Abbreviation
Pentoses	
Ribose	Rib
Xylose	Xyl
Hexoses	
Fructose	Fru
Galactose	Gal
Glucose	Glc
Mannose	Man
Deoxy sugars	
Abequose	Abe
Fucose	Fuc
Amino sugars	
Glucosamine	GlcN
Galactosamine	GalN
N-Acetylglucosamine	GlcNAc
N-Acetylgalactosamine	GalNAc
N-Acetylneuraminic acid	NeuNAc
N-Acetylmuramic acid	MurNAc
Sugar acids	
Glucuronic acid	GlcUA
Iduronic acid	IdoA

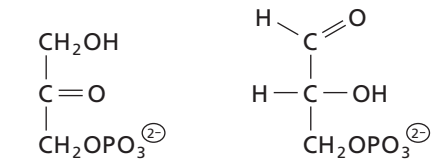


β-2-Deoxy-D-ribose



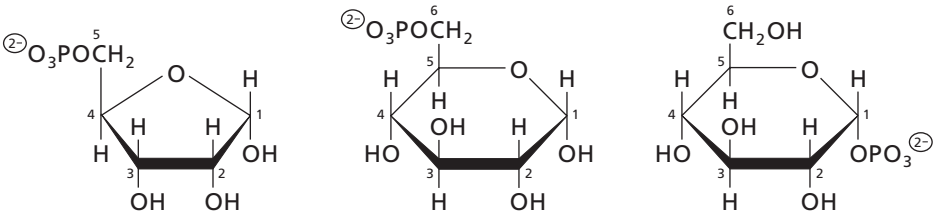
α-L-Fucose
(6-Deoxy-L-galactose)

▲ **Figure 8.15**
Structures of the deoxy sugars 2-deoxy-D-ribose and L-fucose.



Dihydroxyacetone phosphate

D-Glyceraldehyde 3-phosphate



α-D-Ribose 5-phosphate

α-D-Glucose 6-phosphate

α-D-Glucose 1-phosphate

▲ **Figure 8.14**
Structures of several metabolically important sugar phosphates.

sugars are shown in Figure 8.16. Amino sugars formed from glucose and galactose commonly occur in glycoconjugates. *N*-Acetylneuraminic acid (NeuNAc) is an acid formed from *N*-acetylmannosamine and pyruvate. When this compound cyclizes to form a pyranose, the carbonyl group at C-2 (from the pyruvate moiety) reacts with the hydroxyl group of C-6. NeuNAc is an important constituent of many glycoproteins and of a family of lipids called gangliosides (Section 9.5). Neuraminic acid and its derivatives, including NeuNAc, are collectively known as sialic acids.

D. Sugar Alcohols

In a sugar alcohol, the carbonyl oxygen of the parent monosaccharide has been reduced, producing a polyhydroxy alcohol. Figure 8.17 shows three examples of sugar alcohols. Glycerol and *myo*-inositol are important components of lipids (Section 10.4). Ribitol is a component of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) (Section 7.4). In general, sugar alcohols are named by replacing the suffix *-ose* of the parent monosaccharides with *-itol*.

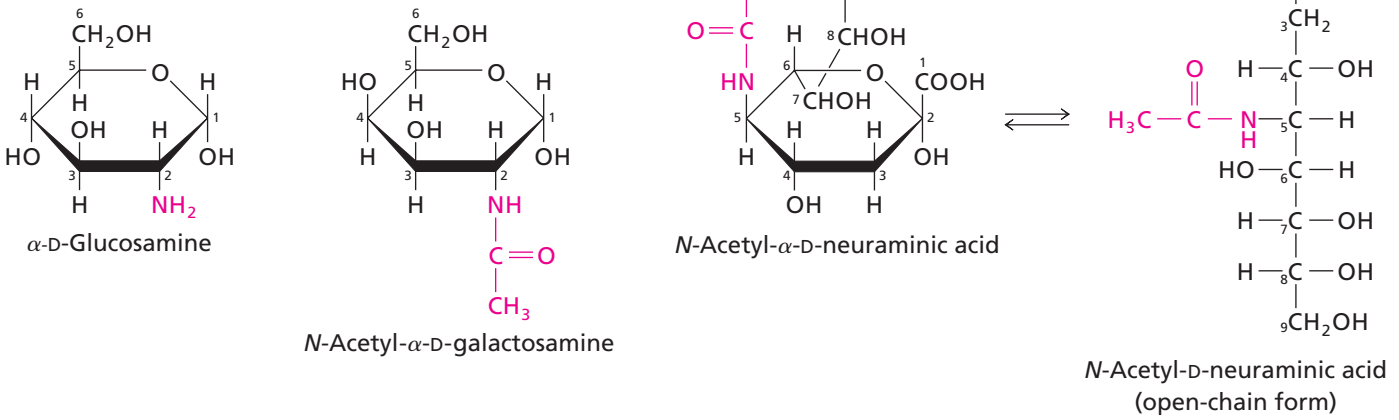
E. Sugar Acids

Sugar acids are carboxylic acids derived from aldoses, either by oxidation of C-1 (the aldehydic carbon) to yield an aldonic acid or by oxidation of the highest-numbered carbon (the carbon bearing the primary alcohol) to yield an alduronic acid. The structures of the aldonic and alduronic derivatives of glucose—gluconate and glucuronate—are shown in Figure 8.18. Aldonic acids exist in the open-chain form in alkaline solution and form lactones (intramolecular esters) on acidification. Alduronic acids can exist as pyranoses and therefore possess an anomeric carbon. Note that *N*-acetylneuraminic acid (Figure 8.16) is a sugar acid as well as an amino sugar. Sugar acids are important components of many polysaccharides. L-Ascorbic acid or vitamin C, is an enediol of a lactone derived from D-glucuronate (Section 7.9).

8.5 Disaccharides and Other Glycosides

The **glycosidic bond** is the primary structural linkage in all polymers of monosaccharides. A glycosidic bond is an acetal linkage in which the anomeric carbon of a sugar is condensed with an alcohol, an amine, or a thiol. As a simple example, glucopyranose

Downloaded From Vet-e-Books.com



▲ **Figure 8.16**
Structures of several amino sugars. The amino and acetylamino groups are shown in red.

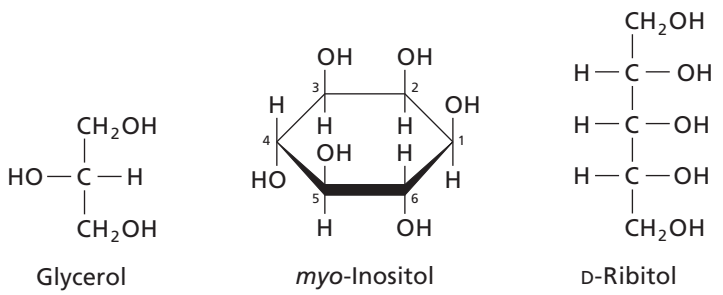
can react with methanol in an acidic solution to form an acetal (Figure 8.19). Compounds containing glycosidic bonds are called **glycosides**; if glucose supplies the anomeric carbon, they are specifically termed **glucosides**. The glycosides include disaccharides, polysaccharides, and some carbohydrate derivatives.

A. Structures of Disaccharides

Disaccharides are formed when the anomeric carbon of one sugar molecule interacts with one of several hydroxyl groups in the other sugar molecule. For disaccharides and other carbohydrate polymers, we must note both the types of monosaccharide residues that are present and the atoms that form the glycosidic bonds. In the systematic description of a disaccharide we must specify the linking atoms, the configuration of the glycosidic bond, and the name of each monosaccharide residue (including its designation as a pyranose or furanose). Figure 8.20 presents the structures and nomenclature for four common disaccharides.

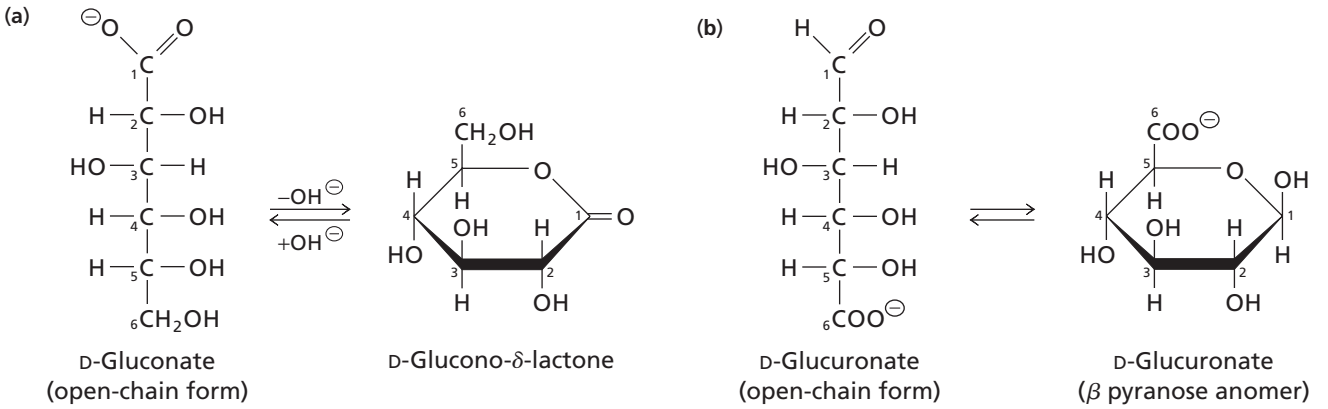
Maltose (Figure 8.20a) is a disaccharide released during the hydrolysis of starch, which is a polymer of glucose residues. It is present in malt, a mixture obtained from corn or grain that is used in malted milk and in brewing. Maltose is composed of two D-glucose residues joined by an α-glycosidic bond. The glycosidic bond links C-1 of one residue (on the left in Figure 8.20a) to the oxygen atom attached to C-4 of the second residue (on the right). Maltose is therefore α-D-glucopyranosyl-$(1 \rightarrow 4)$-D-glucose. Note that the glucose residue on the left, whose anomeric carbon is involved in the glycosidic bond, is fixed in the α configuration, whereas the glucose residue on the right (the reducing end, as explained in Section 8.5B) freely equilibrates among the α, β, and open-chain structures. (The open-chain form is present in very small amounts). The structure shown in Figure 8.20a is the β-pyranose anomer of maltose (the anomer whose reducing end is in the β configuration, the predominant anomeric form).

Cellobiose [β-D-glucopyranosyl-$(1 \rightarrow 4)$-D-glucose] is another glucose dimer (Figure 8.20b). Cellobiose is the repeating disaccharide in the structure of cellulose, a



◀ **Figure 8.17**
Structures of several sugar alcohols. Glycerol (a reduced form of glyceraldehyde) and myo-inositol (metabolically derived from glucose) are important constituents of many lipids. Ribitol (a reduced form of ribose) is a constituent of the vitamin riboflavin and its coenzymes.

Downloaded From Vet-e-Books.com



▲ **Figure 8.18**
Structures of sugar acids derived from D-glucose. (a) Gluconate and its δ-lactone. **(b)** The open-chain and pyranose forms of glucuronate.

plant polysaccharide, and is released during cellulose degradation. The only difference between cellobiose and maltose is that the glycosidic linkage in cellobiose is β (it is α in maltose). The glucose residue on the right in Figure 8.20b, like the residue on the right in Figure 8.20a, equilibrates among the α, β, and open-chain structures.

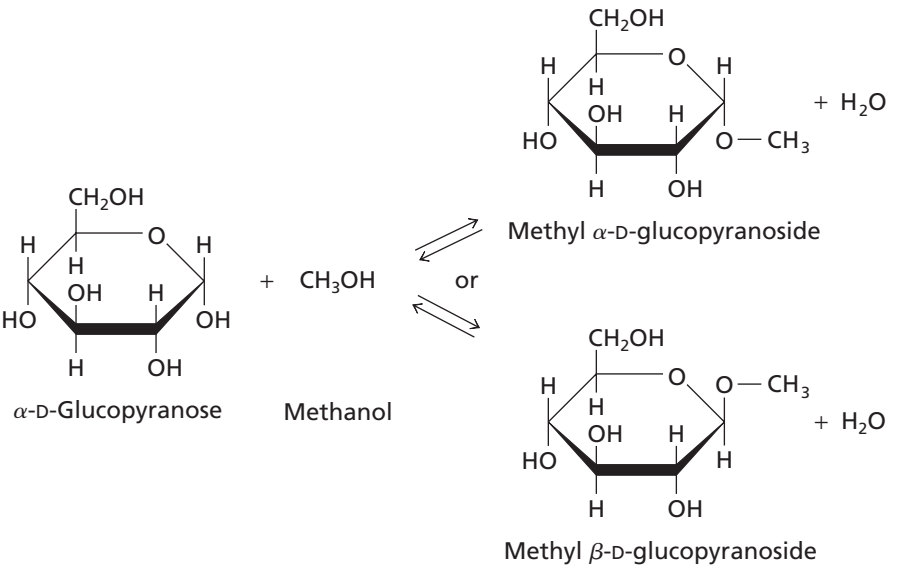
Lactose [β-D-galactopyranosyl-(1 → 4)-D-glucose], a major carbohydrate in milk, is a disaccharide synthesized only in lactating mammary glands (Figure 8.20c). Note that lactose is an epimer of cellobiose. The naturally occurring α anomer of lactose is sweeter and more soluble than the β anomer. The β anomer can be found in stale ice cream, where it has crystallized during storage and given a gritty texture to the ice cream.

Sucrose [α-D-glucopyranosyl-(1 → 2)-β-D-fructofuranoside], or table sugar, is the most abundant disaccharide found in nature (Figure 8.20d). Sucrose is synthesized only in plants. Sucrose is distinguished from the other three disaccharides in Figure 8.20 because its glycosidic bond links the anomeric carbon atoms of two monosaccharide residues. Therefore, the configurations of both the glucopyranose and fructofuranose residues in sucrose are fixed, and neither residue is free to equilibrate between α and β anomers.

B. Reducing and Nonreducing Sugars

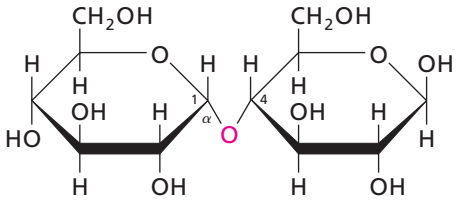
Monosaccharides, and most disaccharides, are hemiacetals with a reactive carbonyl group. They are readily oxidized to diverse products, a property often used in their analysis. Such carbohydrates, including glucose, maltose, cellobiose, and lactose, are sometimes called reducing sugars. Historically, reducing sugars were detected by their ability

Figure 8.19 ►
Reaction of glucopyranose with methanol produces a glycoside. In this acid-catalyzed condensation reaction, the anomeric —OH group of the hemiacetal is replaced by an —OCH_3 group, forming methyl glucoside, an acetal. The product is a mixture of the α and β anomers of methyl glucopyranoside.



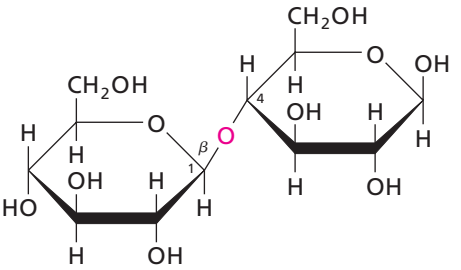
Downloaded From Vet-e-Books.com

(a)



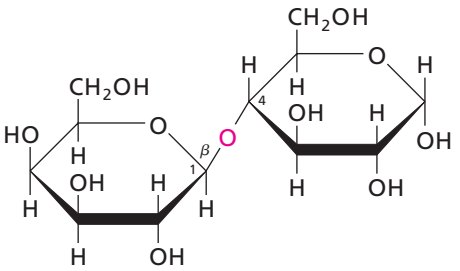
β anomer of maltose
(α-D-Glucopyranosyl-(1→4)-β-D-glucopyranose)

(b)



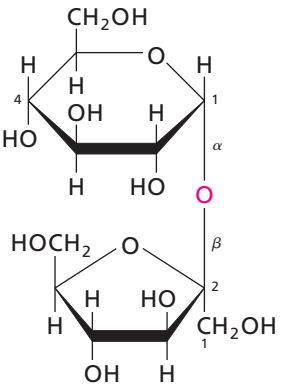
β anomer of cellobiose
(β-D-Glucopyranosyl-(1→4)-β-D-glucopyranose)

(c)



α anomer of lactose
(β-D-Galactopyranosyl-(1→4)-α-D-glucopyranose)

(d)



Sucrose
(α-D-Glucopyranosyl-(1→2)-β-D-fructofuranoside)

to reduce metal ions such as Cu^{2+} or Ag^{+} to insoluble products. Carbohydrates that are not hemiacetals, such as sucrose, are not readily oxidized because both anomeric carbon atoms are fixed in a glycosidic linkage. These are classified as nonreducing sugars.

The reducing ability of a sugar polymer is of more than analytical interest. The polymeric chains of oligosaccharides and polysaccharides show directionality based on their reducing and nonreducing ends. There is usually one reducing end (the residue containing the free anomeric carbon) and one nonreducing end in a linear polymer. All the internal glycosidic bonds of a polysaccharide involve acetals. The internal residues are not in equilibrium with open-chain forms and thus cannot reduce metal ions. A branched polysaccharide has a number of nonreducing ends but only one reducing end.

▲ **Figure 8.20**
Structures of (a) maltose, (b) cellobiose, (c) lactose, and (d) sucrose. The oxygen atom of each glycosidic bond is shown in red.

C. Nucleosides and Other Glycosides

The anomeric carbons of sugars form glycosidic linkages not only with other sugars but also with a variety of alcohols, amines, and thiols. The most commonly encountered glycosides, other than oligosaccharides and polysaccharides, are the nucleosides, in which a purine or pyrimidine is attached by its secondary amino group to a β-D-ribofuranose or β-D-deoxyribofuranose moiety. Nucleosides are called *N*-glycosides because a nitrogen atom participates in the glycosidic linkage. Guanosine (β-D-ribofuranosylguanine) is a typical nucleoside (Figure 8.21). We have already discussed ATP and other nucleotides that are metabolite coenzymes (Section 7.3). NAD and FAD also are nucleotides.

Two other examples of naturally occurring glycosides are shown in Figure 8.21. Vanillin glucoside (Figure 8.21b) is the flavored compound in natural vanilla extract. β-Galactosides constitute an abundant class of glycosides. In these compounds, a variety of nonsugar molecules are joined in β linkage to galactose. For example, galactocerebrosides (see Section 9.5) are glycolipids common in eukaryotic cell membranes and can be hydrolyzed readily by the action of enzymes called β-galactosidases.



▲ **Sugar cane is a major source of commercial sucrose.**

There is a more complete discussion of nucleosides and nucleotides in Chapter 19.

Downloaded From Vet-e-Books.com

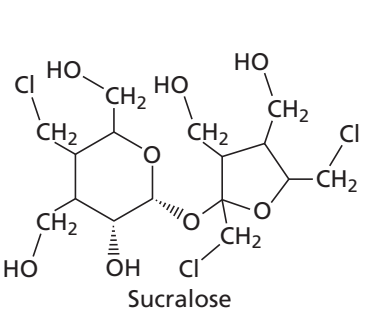
BOX 8.1 THE PROBLEM WITH CATS

One of the characteristics of sugars is that they taste sweet. You certainly know the taste of sucrose and you probably know that fructose and lactose also taste sweet. So do many of the other sugars and their derivatives, although we don't recommend that you go into a biochemistry lab and start tasting all the carbohydrates in those white plastic bottles on the shelves.

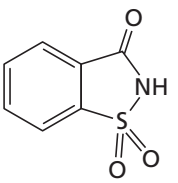
Sweetness is not a physical property of molecules. It's a subjective interaction between a chemical and taste receptors in your mouth. There are five different kinds of taste receptors: sweet, sour, salty, bitter, and umami (umami is like the taste of glutamate in monosodium glutamate). In order to trigger the sweet taste, a molecule like sucrose has to bind to the receptor and initiate a response that eventually makes it to your brain. Sucrose elicits a moderately strong response that serves as the standard for sweetness. The response to fructose is almost twice as strong and the response to lactose is only about one-fifth as strong as that of sucrose. Artificial sweeteners such as saccharin (Sweet'N Low®), sucralose (Splenda®), and aspartame (NutraSweet®) bind to the sweetness receptor and cause the sensation of sweetness. They are hundreds of times more sweet than sucrose.

The sweetness receptor is encoded by two genes called *Tas1r2* and *Tas1r3*. We don't know how sucrose and the other ligands bind to this receptor even though this is a very active area of research. In the case of sucrose and the artifical sweeteners, how can such different molecules elicit the taste of sweet?

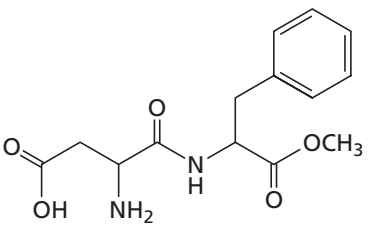
Cats, including lions, tigers and cheetahs, do not have a functional *Tas1r2* gene. It has been converted to a pseudogene because of a 247 bp deletion in exon 3. It's very likely that your pet cat has never experienced the taste of sweetness. That explains a lot about cats.



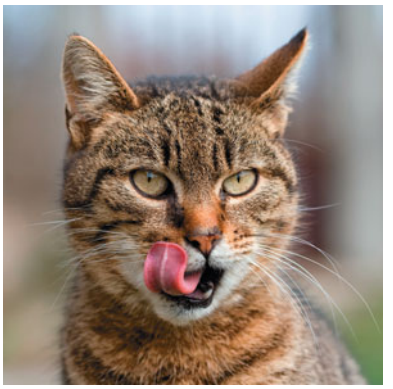
Sucralose



Saccharin



Aspartame



▲ **Cats are carnivores.** They probably can't taste sweetness.

8.6 Polysaccharides

Polysaccharides are frequently divided into two broad classes. **Homoglycans,** or homopolysaccharides, are polymers containing residues of only one type of monosaccharide. **Heteroglycans**, or heteropolysaccharides, are polymers containing residues of more than one type of monosaccharide. Polysaccharides are created without a template by the addition of particular monosaccharide and oligosaccharide residues. As a result, the lengths and compositions of polysaccharide molecules may vary within a population of these molecules. Some common polysaccharides and their structures are listed in Table 8.2.

Most polysaccharides can also be classified according to their biological roles. For example, starch and glycogen are storage polysaccharides while cellulose and chitin are structural polysaccharides. We will see additional examples of the variety and versatility of carbohydrates when we discuss the heteroglycans in the next section."

A. Starch and Glycogen

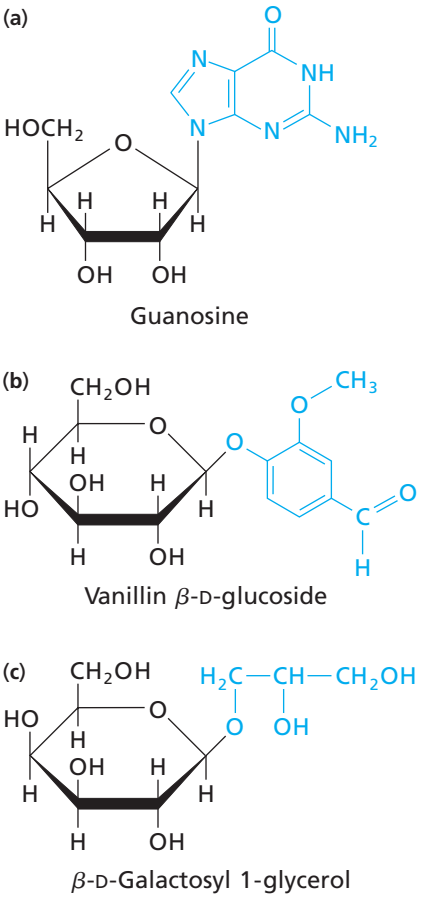
D-Glucose is synthesized in all species. Excess glucose can be broken down to produce metabolic energy. Glucose residues are stored as polysaccharides until they are needed for energy production. The most common storage homoglycan of glucose in plants and fungi is starch and in animals it is glycogen. Both types of polysaccharides occur in bacteria.

Downloaded From Vet-eBooks.com

Table 8.2 Structures of some common polysaccharides

Polysaccharide[a]	Component(s)[b]	Linkage(s)
Storage homoglycans		
Starch		
Amylose	Glc	α-(1 → 4)
Amylopectin	Glc	α-(1 → 4), α-(1 → 6) (branches)
Glycogen	Glc	α-(1 → 4), α-(1 → 6) (branches)
Structural homoglycans		
Cellulose	Glc	β(1 → 4)
Chitin	GlcNAc	β(1 → 4)
Heteroglycans		
Glycosaminoglycans	Disaccharides (amino sugars, sugar acids)	Various
Hyaluronic acid	GlcUA and GlcNAc	β(1 → 3), β(1 → 4)

[a]Polysaccharides are unbranched unless otherwise indicated.
[b]Glc, Glucose; GlcNAc, *N*-acetylglucosamine; GlcUA, D-glucuronate.



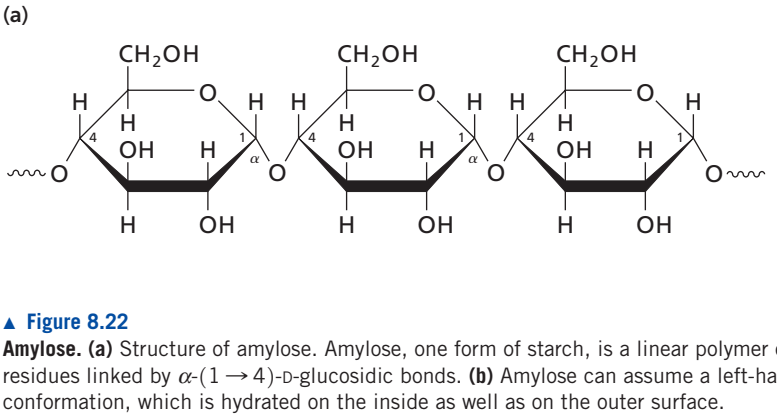
▲ **Figure 8.21**
Structures of three glycosides. The nonsugar components are shown in blue. **(a)** Guanosine. **(b)** Vanillin glucoside, the flavored compound in vanilla extract. **(c)** β-D-Galactosyl 1-glycerol, derivatives of which are common in eukaryotic cell membranes.

Starch is present in plant cells as a mixture of amylose and amylopectin and is stored in granules whose diameters range from 3 to 100 μm. Amylose is an unbranched polymer of about 100 to 1000 D-glucose residues connected by α-(1 → 4) glycosidic linkages, specifically termed α-(1 → 4) *glucosidic bonds* because the anomeric carbons belong to glucose residues (Figure 8.22a). The same type of linkage connects glucose monomers in the disaccharide maltose (Figure 8.20a). Although it is not truly soluble in water, amylose forms hydrated micelles in water and can assume a helical structure under some conditions (Figure 8.22b).

Amylopectin is a branched version of amylose (Figure 8.23). Branches, or polymeric side chains, are attached via α-(1 → 6) glucosidic bonds to linear chains of residues linked by α-(1 → 4) glucosidic bonds. Branching occurs, on average, once every 25 residues and the side chains contain about 15 to 25 glucose residues. Some side chains themselves are branched. Amylopectin molecules isolated from living cells may contain 300 to 6000 glucose residues.

An adult human consumes about 300 g of carbohydrate daily, much of which is in the form of starch. Raw starch granules resist enzymatic hydrolysis but cooking causes them to absorb water and swell. The swollen starch is a substrate for two different glycosidases. Dietary starch is degraded in the gastrointestinal tract by the actions of α-amylase and a debranching enzyme. α-Amylase, which is present in both animals and

Starch metabolism is described in Chapter 15.



▲ **Figure 8.22**
Amylose. (a) Structure of amylose. Amylose, one form of starch, is a linear polymer of glucose residues linked by α-(1 → 4)-D-glucosidic bonds. **(b)** Amylose can assume a left-handed helical conformation, which is hydrated on the inside as well as on the outer surface.

Downloaded From Vet-e-Books.com

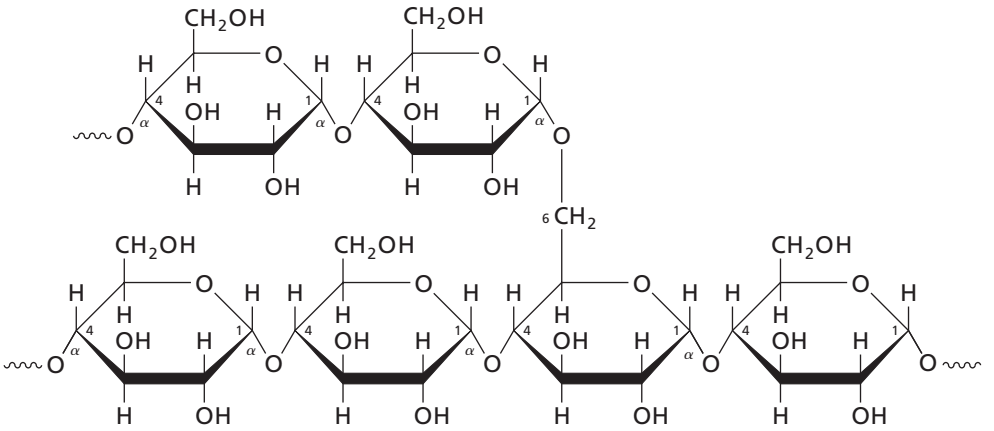


Figure 8.23 ►
Structure of amylopectin. Amylopectin, a second form of starch, is a branched polymer. The linear glucose residues of the main chain and the side chains of amylopectin are linked by α-(1 → 4)-D-glucosidic bonds, and the side chains are linked to the main chain by α-(1 → 6)-D-glucosidic bonds.

plants, is an endoglycosidase (it acts on internal glycosidic bonds). The enzyme catalyzes random hydrolysis of the α-(1 → 4) glucosidic bonds of amylose and amylopectin.

Another hydrolase, β-amylase, is found in the seeds and tubers of some plants. β-Amylase is an exoglycosidase (it acts on terminal glycosidic bonds). It catalyzes sequential hydrolytic release of maltose from the free, nonreducing ends of amylopectin.

Despite their α and β designations, both types of amylases act only on α-(1 → 4)-D-glycosidic bonds. Figure 8.24 shows the action of α-amylase and β-amylase on amylopectin. The α-(1 → 6) linkages at branch points are not substrates for either α- or β-amylase. After amylase-catalyzed hydrolysis of amylopectin, highly branched cores resistant to further hydrolysis, called **limit dextrins**, remain. Limit dextrins can be further degraded only after debranching enzymes have catalyzed hydrolysis of the α-(1 → 6) linkages at branch points.

Glycogen is also a branched polymer of glucose residues. Glycogen contains the same types of linkages found in amylopectin but the branches in glycogen are smaller and more frequent, occurring every 8–12 residues. In general, glycogen molecules are larger than starch molecules, Glycogen up to contains 50,000 glucose residues. In mammals,

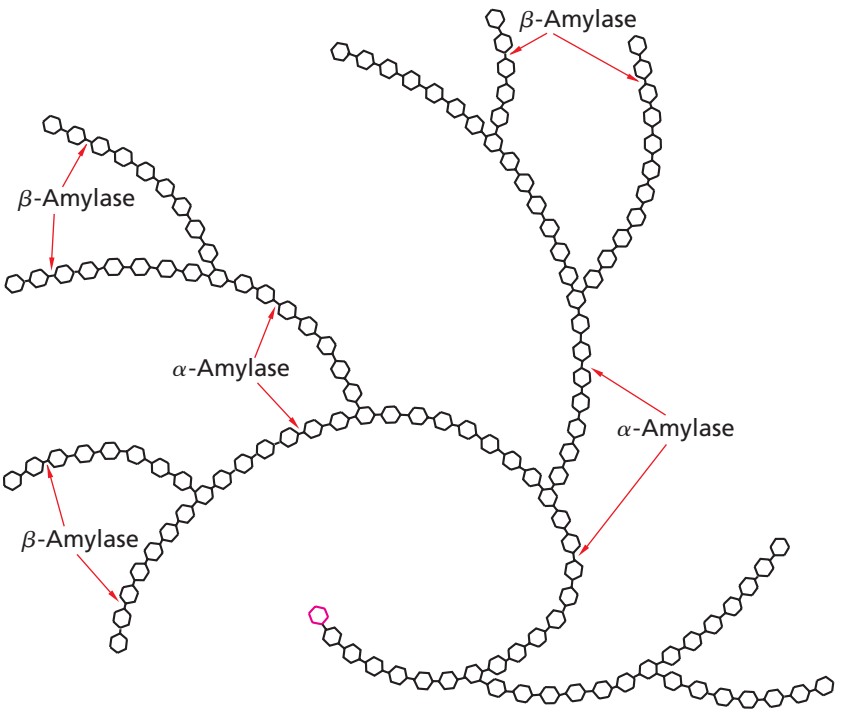


Figure 8.24 ►
Action of α-amylase and β-amylase on amylopectin. α-Amylase catalyzes random hydrolysis of internal α-(1 → 4) glucosidic bonds; β-amylase acts on the nonreducing ends. Each hexagon represents a glucose residue; the single reducing end of the branched polymer is red. (An actual amylopectin molecule contains many more glucose residues than shown here.)

Downloaded From Vet-e-Books.com

depending on the nutritional state, glycogen can account for up to 10% of the mass of the liver and 2% of the mass of muscle.

The branched structures of amylopectin and glycogen possess only one reducing end but many nonreducing ends. The reducing end of glycogen is covalently attached to a protein called glycogenin (Section 12.5A). Enzymatic lengthening and degradation of polysaccharide chains occurs at the nonreducing ends.

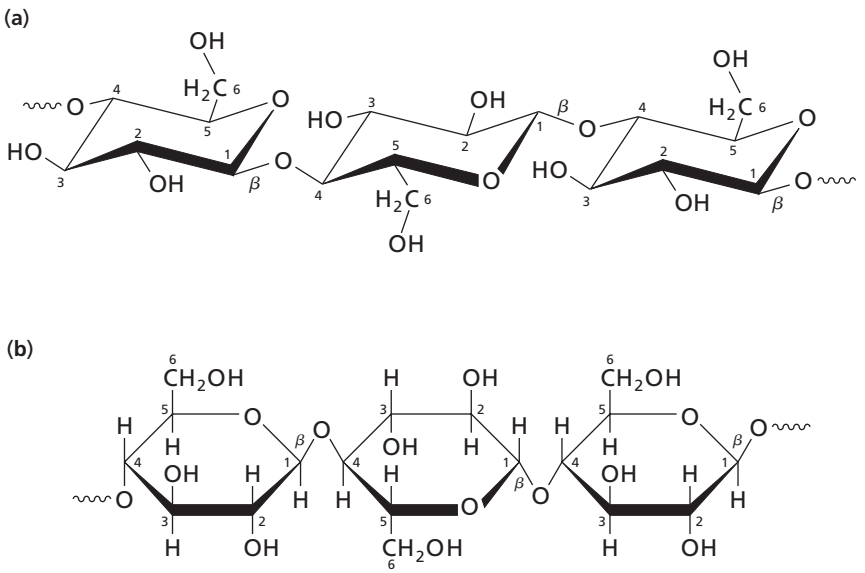
Enzymes that catalyze the intracellular synthesis and breakdown of glycogen are described in Chapter 12.

B. Cellulose

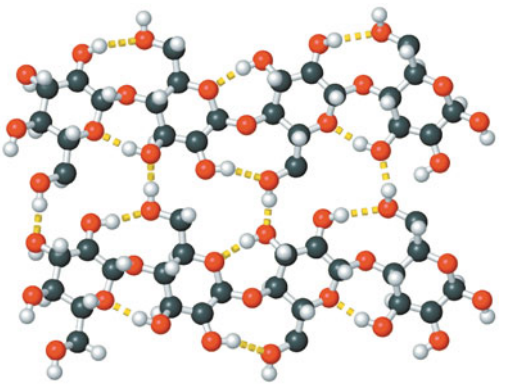
Cellulose is a structural polysaccharide. It is a major component of the rigid cell walls that surround many plant cells. The stems and branches of many plants consist largely of cellulose. This single polysaccharide accounts for a significant percentage of all organic matter on Earth. Like amylose, cellulose is a linear polymer of glucose residues, but in cellulose the glucose residues are joined by β-(1 → 4) linkages rather than α-(1 → 4) linkages. The two glucose residues of the disaccharide cellobiose also are connected by a β-(1 → 4) linkage (Figure 8.20b). Cellulose molecules vary greatly in size, ranging from about 300 to more than 15,000 glucose residues.

The β linkages of cellulose result in a rigid extended conformation in which each glucose residue is rotated 180° relative to its neighbors (Figure 8.25). Extensive hydrogen bonding within and between cellulose chains leads to the formation of bundles, or fibrils (Figure 8.26). Cellulose fibrils are insoluble in water and are quite strong and rigid. Cotton fibers are almost entirely cellulose and wood is about half cellulose. Because of its strength, cellulose is used for a variety of purposes and is a component of a number of synthetic materials including cellophane and the fabric rayon. We are most familiar with cellulose as the main component of paper.

Enzymes that catalyze the hydrolysis of α-D-glucosidic bonds (α-glucosidases, such as α- and β-amylase) do not catalyze the hydrolysis of β-D-glucosidic bonds. Similarly, β-glucosidases (such as cellulase) do not catalyze the hydrolysis of α-D-glucosidic bonds. Humans and other mammals can metabolize starch, glycogen, lactose, and sucrose and use the monosaccharide products in a variety of metabolic pathways. Mammals cannot metabolize cellulose because they lack enzymes capable of catalyzing the hydrolysis of β-glucosidic linkages. Ruminants such as cows and sheep have microorganisms in their rumen (a compartment in their multichambered stomachs) that produce β-glucosidases. Thus, ruminants can obtain glucose from grass and other plants that are rich in cellulose. Because they have cellulase-producing bacteria in their digestive tracts, termites also can obtain glucose from dietary cellulose.



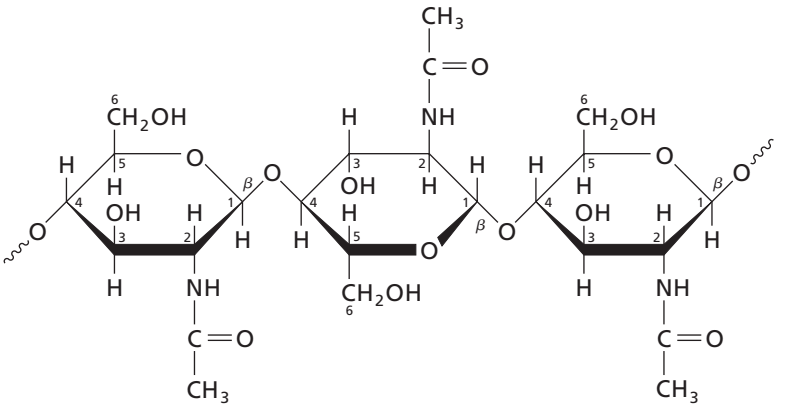
▲ **Figure 8.25**
Structure of cellulose. Note the alternating orientation of successive glucose residues in the cellulose chain. **(a)** Chair conformation. **(b)** Modified Haworth projection.



▲ **Figure 8.26**
Cellulose fibrils. Intra- and interchain hydrogen bonding gives cellulose its strength and rigidity.

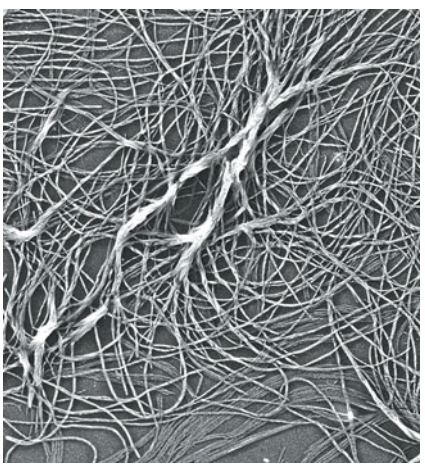
Downloaded From Vet-e-Books.com

Figure 8.27 ►
Structure of chitin. The linear homoglycan chitin consists of repeating units of β-(1 → 4)-linked GlcNAc residues. Each residue is rotated 180° relative to its neighbors.





▲ The giant redwood trees of California contains tons of cellulose.



▲ **Cellulose fibers.** Plants make large cellulose fibers that serve as structural support. A scanning electron micrograph of these fibers shows how they overlap to form a large net-like sheet. These cellulose fibers are about 253 million years old. They were recovered from deep within a salt mine in New Mexico.

C. Chitin

Chitin, probably the second most abundant organic compound on Earth, is a structural homoglycan found in the exoskeletons of insects and crustaceans and also in the cell walls of most fungi and red algae. Chitin is a linear polymer similar to cellulose. It is made up of β-(1 → 4)-linked GlcNAc residues rather than glucose residues (Figure 8.27). Each GlcNAc residue is rotated 180° relative to its neighbors. The GlcNAc residues in adjacent strands of chitin form hydrogen bonds with each other resulting in linear fibrils of great strength. Chitin is often closely associated with nonpolysaccharide compounds, such as proteins and inorganic material.

8.7 Glycoconjugates

Glycoconjugates consist of polysaccharides linked to (conjugated with) proteins or peptides. In most cases, the polysaccharides are composed of several different monosaccharide units. Thus, they are heteroglycans. (Starch, glycogen, cellulose, and chitin are homoglycans.) Heteroglycans appear in three types of glycoconjugates—proteoglycans, peptidoglycans, and glycoproteins. In this section, we see how the chemical and physical properties of the heteroglycans in glycoconjugates are suited to various biological functions.

A. Proteoglycans

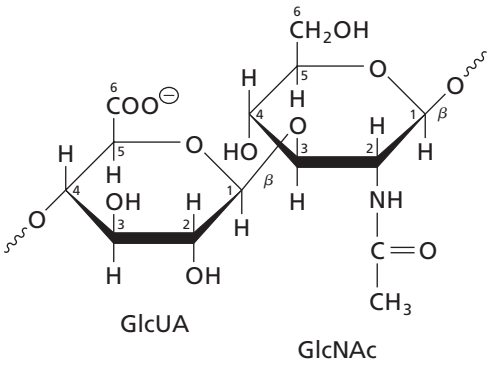
Proteoglycans are complexes of proteins and a class of polysaccharides called *glycosaminoglycans*. These glycoconjugates occur predominately in the extracellular matrix (connective tissue) of multicellular animals.

Glycosaminoglycans are unbranched heteroglycans of repeating disaccharide units. As the name *glycosaminoglycan* indicates, one component of the disaccharide is an amino sugar, either D-galactosamine (GalN) or D-glucosamine (GlcN). The amino group of the amino-sugar component can be acetylated forming *N*-acetylgalactosamine (GalNAc) or GlcNAc. The other component of the repeating disaccharide is usually an alduronic acid. Specific hydroxyl and amino groups of many glycosaminoglycans are sulfated. These sulfate groups and the carboxylate groups of alduronic acids make glycosaminoglycans polyanionic.

Several types of glycosaminoglycans have been isolated and characterized. Each type has its own sugar composition, linkages, tissue distribution, and function and each is attached to a characteristic protein. Hyaluronic acid is an example of a glycosaminoglycan composed of the repeating disaccharide unit shown in Figure 8.28. It is found in the fluid of joints where it forms a viscous solution that is an excellent lubricant. Hyaluronic acid is also a major component of cartilage.

Up to100 glycosaminoglycan chains can be attached to the protein of a proteoglycan. These heteroglycan chains are usually covalently bound by a glycosidic linkage to

Downloaded From Vet-e-Books.com



◀ **Figure 8.28**
Structure of the repeating disaccharide of hyaluronic acid. The repeating disaccharide of this glycosaminoglycan contains D-glucuronate (GlcUA) and GlcNAc. Each GlcUA residue is linked to a GlcNAc residue through β-(1 → 3) linkage; each GlcNAc residue is in turn linked to the next GlcUA residue through a β-(1 → 4) linkage.

the hydroxyl oxygens of serine residues. (Not all glycosaminoglycans are covalently linked to proteins.) Glycosaminoglycans can account for up to 95% of the mass of a proteoglycan.

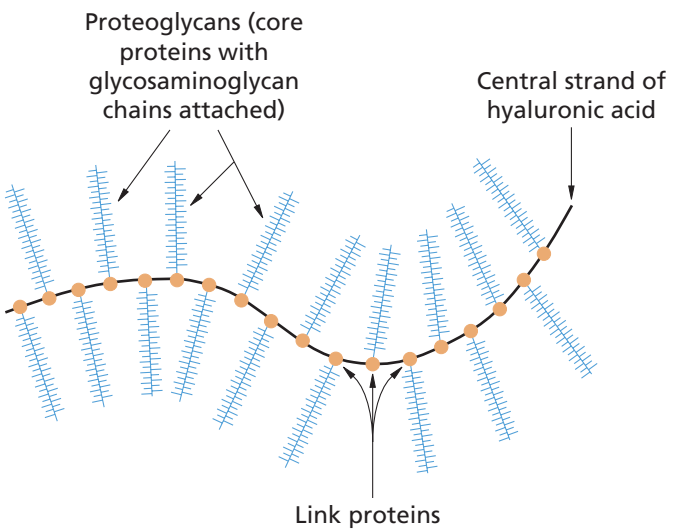
Proteoglycans are highly hydrated and occupy a large volume because their glycosaminoglycan component contains polar and ionic groups. These features confer elasticity and resistance to compression—important properties of connective tissue. For example, the flexibility of cartilage allows it to absorb shocks. Some of the water can be pressed out when cartilage is compressed but relief from pressure allows cartilage to rehydrate. In addition to maintaining the shapes of tissues, proteoglycans can also act as extracellular sieves and help direct cell growth and migration.

Examination of the structure of cartilage shows how proteoglycans are organized in this tissue. Cartilage is a mesh of collagen fibers (Section 4.11) interspersed with large proteoglycan aggregates ($M_r \sim 2 \times 10^8$). Each aggregate assumes a characteristic shape that resembles a bottle brush (Figure 8.29). These aggregates contain hyaluronic acid and several other glycosaminoglycans, as well as two types of proteins—core proteins and link proteins. A central strand of hyaluronic acid runs through the aggregate and many proteoglycans—core proteins with glycosaminoglycan chains attached—branch from its sides. The core proteins interact noncovalently with the hyaluronic acid strand, mostly by electrostatic interactions. Link proteins stabilize the core protein–hyaluronic acid interactions.

The major proteoglycan of cartilage is called *aggrecan*. The protein core of aggrecan ($M_r \sim 220{,}000$) carries approximately 30 molecules of keratan sulfate (a glycosaminoglycan composed chiefly of alternating *N*-acetylglucosamine 6-sulfate and galactose residues) and approximately 100 molecules of chondroitin sulfate (a glycosaminoglycan



▲ **Lobsters have an exoskeleton made of chitin.** The color of the exoskeleton is determined by the foods that the lobster eats. When it ingests β-carotene derivatives they are converted to a complex mixture of protein-bound carotenes called crustacayanin that has a greenish-brown color. When lobsters are cooked, the crustacyanin breaks down, releasing free β-carotene derivatives that are red in color, like the red color of maple leaves in autumn (see Section 15.1).



◀ **Figure 8.29**
Proteoglycan aggregate of cartilage. Core proteins carrying glycosaminoglycan chains are associated with a central strand of a single hyaluronic acid molecule. These proteins have many covalently attached glycosaminoglycan chains (keratan sulfate and chondroitin sulfate molecules). The interactions of the core proteins with hyaluronic acid are stabilized by link proteins, which interact noncovalently with both types of molecules. The aggregate has the appearance of a bottle brush.

Downloaded From Vet-e-Books.com

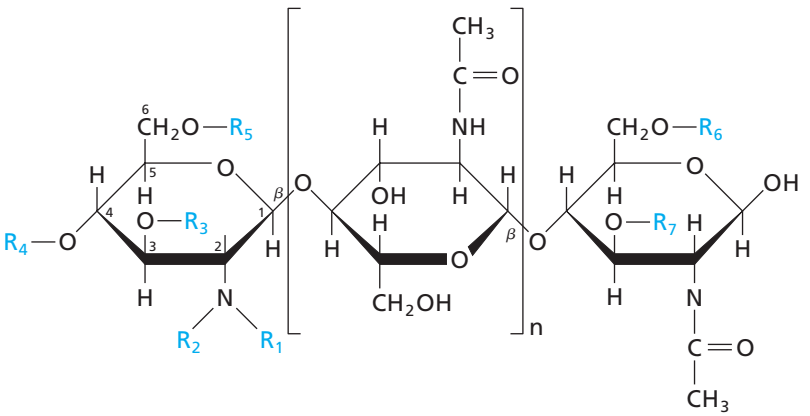
BOX 8.2 NODULATION FACTORS ARE LIPO-OLIGOSACCHARIDES

Legumes such as alfalfa, peas, and soybeans develop organs called *nodules* on their roots. Certain soil bacteria (rhizobia) infect the nodules and, in a symbiosis with the plants, carry out nitrogen fixation (reduction of atmospheric nitrogen to ammonia). The symbiosis is highly species-specific: only certain combinations of legumes and bacteria can cooperate and therefore these organisms must recognize each other. Rhizobia produce extracellular signal molecules that are oligosaccharides called nodulation factors. Extremely low concentrations of these compounds can induce their plant hosts to develop the nodules that the rhizobia can infect. A host plant responds only to a nodulation factor of a characteristic composition.

Infection begins when the plant root hair recognizes the nodulation factor via surface Nod-factor receptors. This results in a response that allows the bacteria to enter the root hair and migrate down to the cells in the root where the nodule forms.

All the nodulation factors studied to date are oligosaccharides that have a linear chain of β-(1 → 4) *N*-acetylglucosamine (GlcNAc)—the same repeating structure as in chitin (Section 8.6b). Most nodulation factors are sugar pentamers although the number of residues can vary between three and six (see figure below). Species specificity is provided by variation in polymer length and potential substitution on five sites at the nonreducing end (R1 to R5) and two sites at the reducing end (R6 and R7). R1, an acyl group substituting the nitrogen atom at C-2 of the nonreducing end, is a fatty acid, usually 18 carbons long. Thus, the nodulation factors are lipo-oligosaccharides. R6, bound to the alcohol at C-6 of the reducing end, can have a wide variety of structures, including sulfate or methyl fucose. Research on these growth regulators for legumes has stimulated the search for biological activities of other oligosaccharides.

See Section 17.1 for details about nitrogen fixation.



▲ General structure of nodulation factors, lipo-oligosaccharides with an *N*-acetylglucosamine (GlcNAc) backbone. The number of internal residues of *N*-acetylglucosamine is shown by *n*, which is usually 3 but can sometimes be 1, 2, or 4. R1 is a fatty acyl substituent, usually 18 carbons long.



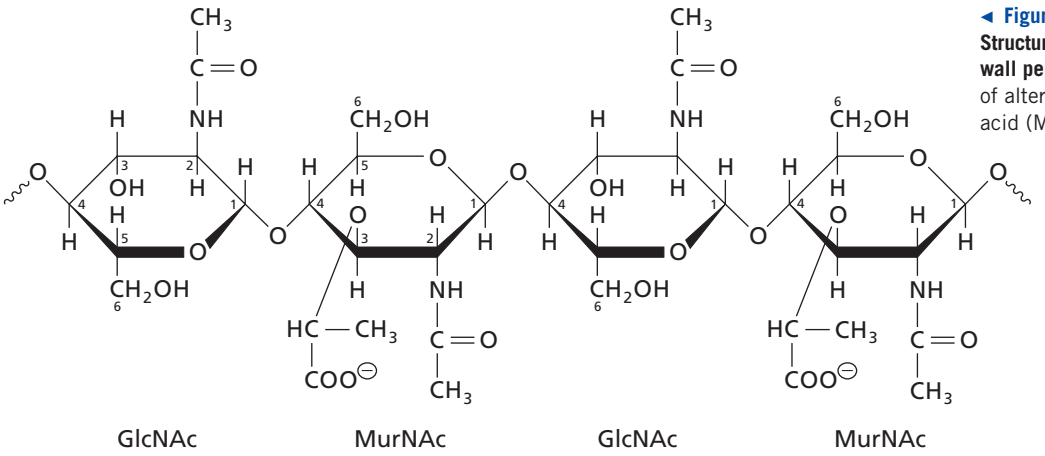
▲ **Formation of nodules in the legume *Lotus japonicus*.** Rhizobia (blue) have secreted nodulation factor leading to endocytosis by root hair cells and formation of an infection thread connecting the point of uptake (top) to the root nodule cells (below).

composed of alternating *N*-acetylgalactosamine sulfate and glucuronate residues). Aggrecan is a member of a small family of hyalectans, proteoglycans that bind to hyaluronic acid. Other hyalectans provide elasticity to blood vessel walls and modulate cell-cell interactions in the brain.

B. Peptidoglycans

Peptidoglycans are polysaccharides linked to small peptides. The cell walls of many bacteria contain a special class of peptidoglycan with a heteroglycan component attached to a four or five residue peptide. The heteroglycan component is composed of alternating residues of GlcNAc and *N*-acetylmuramic acid (MurNAc) joined by β-(1 → 4) linkages (Figure 8.30). MurNAc is a nine-carbon sugar found only in bacteria. MurNAc consists of the three-carbon acid D-lactate joined by an ether linkage to C-3 of GlcNAc.

Downloaded From Vet-eBooks.com



◀ **Figure 8.30**
Structure of the polysaccharide in bacterial cell wall peptidoglycan. The glycan is a polymer of alternating GlcNAc and *N*-acetylmuramic acid (MurNAc) residues.

The polysaccharide moiety of peptidoglycans resembles chitin except that every second GlcNAc residue is modified by addition of lactate to form MurNAc. The antibacterial action of lysozyme (Section 6.6) results from its ability to catalyze hydrolysis of the polysaccharide chains of peptidoglycans.

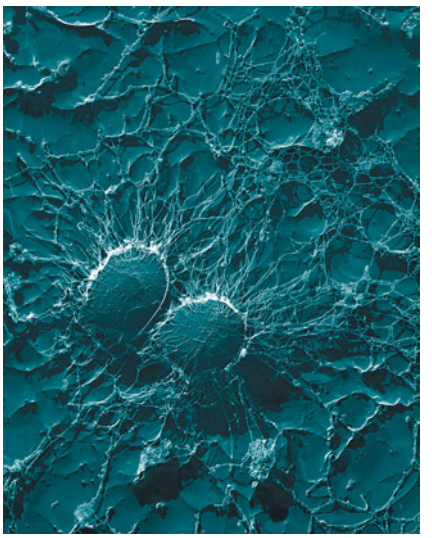
The peptide component of peptidoglycans varies among bacteria. The peptide component in *Staphylococcus aureus* is a tetrapeptide with alternating L and D amino acids: L-Ala–D-Isoglu–L-Lys–D-Ala. (Isoglu represents isoglutamate, a form of glutamate in which the γ-carboxyl group—not the α-carboxyl group—is linked to the next residue.) Other species have a different amino acid at the third position. An amide bond links the amino group of the L-alanine residue to the lactyl carboxylate group of a MurNAc residue of the glycan polymer (Figure 8.31). The tetrapeptide is cross-linked to another tetrapeptide on a neighboring peptidoglycan molecule by a chain of five glycine residues (pentaglycine). Pentaglycine joins the L-lysine residue of one tetrapeptide to the carboxyl group of the D-alanine residue of the other tetrapeptide. Extensive cross-linking essentially converts the peptidoglycan to one huge, rigid, macromolecule that defines the shape of the bacterium by covering its plasma membrane and protecting the cell from fluctuations in osmotic pressure.

Most bacteria have an additional exterior layer of dense polysaccharide called the *capsule*. The capsule is made up of chains of polysaccharide composed mainly of *N*-acetylglucosamine (GlcNAc) residues but various other amino sugars are present. The capsule protects the bacterial cell from injury. The capsule in pathogenic bacteria help cells avoid destruction by the immune system.

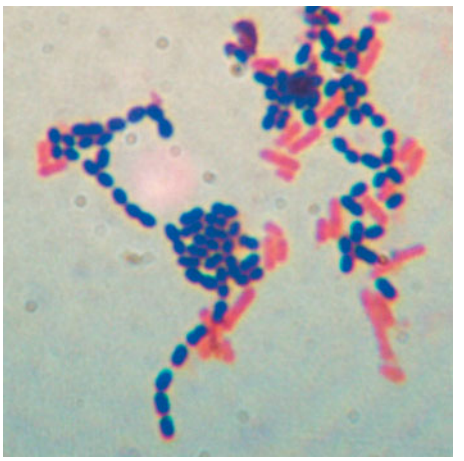
In gram-negative bacteria, the peptidoglycan cell wall lies between the inner plasma membrane and the outer membrane. In gram-positive bacteria, there is no outer membrane and the cell wall is much thicker. This is one of the reasons why the Gram stain (named after Christian Gram) will color the surfaces of some bacteria (gram positive) and not others (gram negative).

During peptidoglycan biosynthesis, a five-residue peptide—L-Ala–D-Isoglu–L-Lys–D-Ala–D-Ala—is attached to a MurNAc residue. In subsequent steps, five glycine residues are added sequentially to the ε-amino group of the lysine residue forming the pentaglycine bridge. In the final step of synthesis, a transpeptidase catalyzes formation of a peptide linkage between the penultimate alanine residue and a terminal glycine residue of a pentaglycine bridge of a neighboring peptidoglycan strand. This reaction is driven by release of the terminal D-alanine residue.

The structure of the antibiotic penicillin (Figure 8.32) resembles the terminal D-Ala–D-Ala residues of the immature peptidoglycan. Penicillin binds, probably irreversibly, to the transpeptidase active site inhibiting the activity of the enzyme and thereby blocking further peptidoglycan synthesis. The antibiotic prevents growth and proliferation of bacteria. Penicillin is selectively toxic to bacteria because the reaction it affects occurs only in certain bacteria, not in eukaryotic cells.

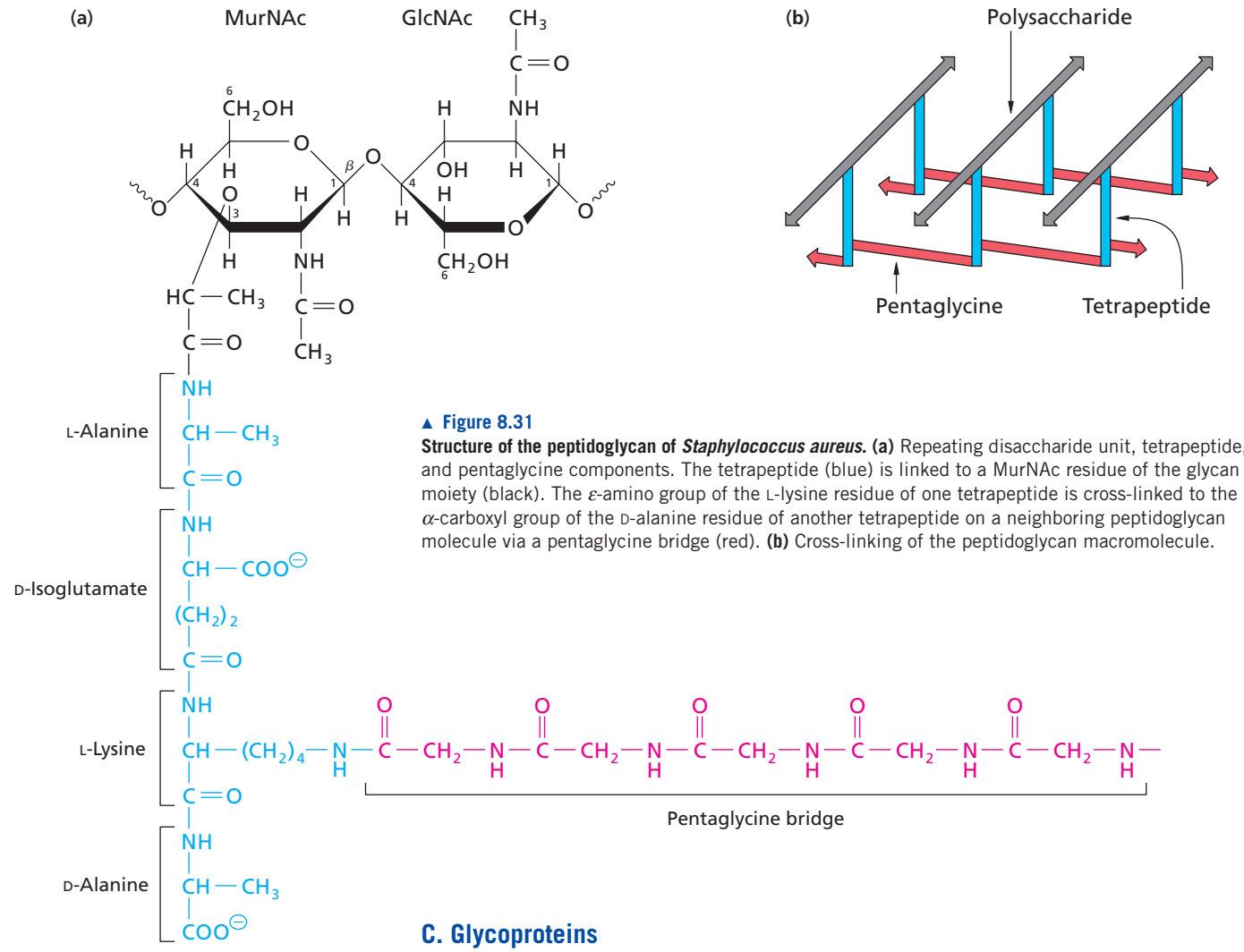


▲ ***Staphylococcus aureus* cells.** These bacterial cells have extensive polysaccharide capsules that protect them from their host's immune system.

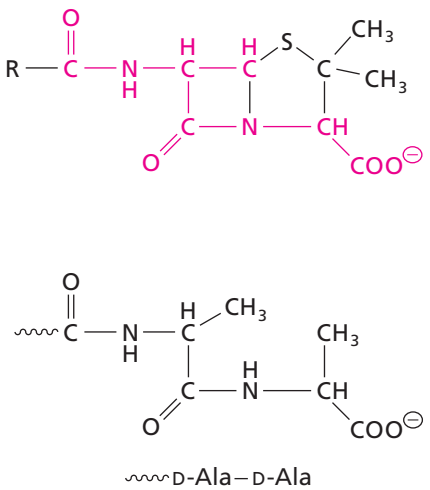


▲ **The Gram stain.** The Gram staining procedure distinguishes between gram-positive bacteria (left, purple) and gram-negative bacteria (right, pink).

Downloaded From Vet-e-Books.com



▲ **Figure 8.31**
Structure of the peptidoglycan of *Staphylococcus aureus*. (a) Repeating disaccharide unit, tetrapeptide, and pentaglycine components. The tetrapeptide (blue) is linked to a MurNAc residue of the glycan moiety (black). The ε-amino group of the L-lysine residue of one tetrapeptide is cross-linked to the α-carboxyl group of the D-alanine residue of another tetrapeptide on a neighboring peptidoglycan molecule via a pentaglycine bridge (red). **(b)** Cross-linking of the peptidoglycan macromolecule.



▲ **Figure 8.32**
Structures of penicillin and –D-Ala–D-Ala. The portion of penicillin that resembles the dipeptide is shown in red. R can be a variety of substituents.

C. Glycoproteins

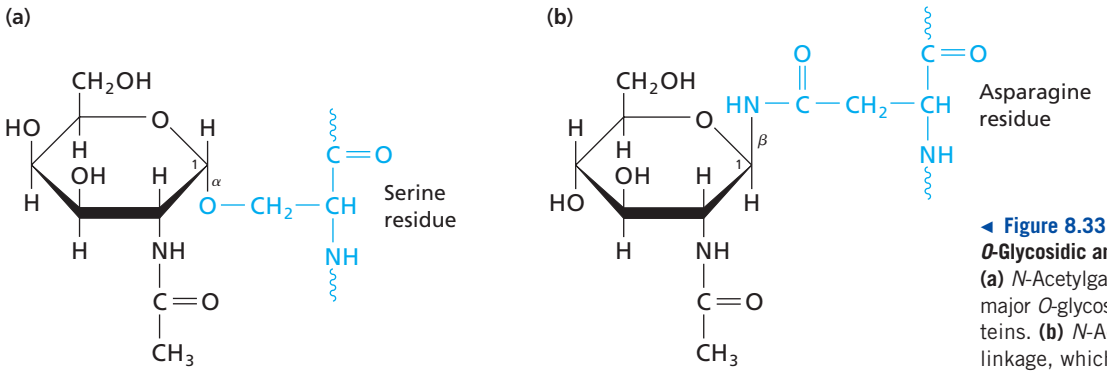
Glycoproteins, like proteoglycans, are proteins that contain covalently bound oligosaccharides (i.e., proteins that are glycosylated). In fact, proteoglycans are a type of glycoprotein. The carbohydrate chains of a glycoprotein vary in length from one to more than 30 residues and can account for as much as 80% of the total mass of the molecule. Glycoproteins are an extraordinarily diverse group of proteins that includes enzymes, hormones, structural proteins, and transport proteins.

The oligosaccharide chains of different glycoproteins exhibit great variability in composition. The composition of oligosaccharide chains can vary even among molecules of the same protein, a phenomenon called *microheterogeneity*.

Several factors contribute to the structural diversity of the oligosaccharide chains of glycoproteins.

1. An oligosaccharide chain can contain several different sugars. Eight sugars predominate in eukaryotic glycoproteins: the hexoses L-fucose, D-galactose, D-glucose, and D-mannose; the hexosamines *N*-acetyl-D-galactosamine and *N*-acetyl-D-glucosamine; the nine-carbon sialic acids (usually *N*-acetylneuraminic acid); and the pentose D-xylose. Many different combinations of these sugars are possible.
2. The sugars can be joined by either α- or β-glycosidic linkages.
3. The linkages can also join various carbon atoms in the sugars. In hexoses and hexosamines, the glycosidic linkages always involve C-1 of one sugar but can involve C-2, C-3, C-4, or C-6 of another hexose or C-3, C-4, or C-6 of an amino sugar (C-2 is usually *N*-acetylated in this class of sugar). C-2 of sialic acid, not C-1, is linked to other sugars.
4. Oligosaccharide chains of glycoproteins can contain up to four branches.

Downloaded From Vet-eBooks.com



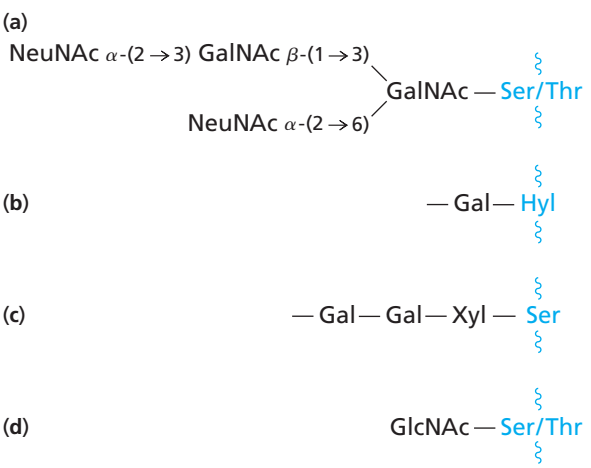
◀ **Figure 8.33**
***O*-Glycosidic and *N*-glycosidic linkages.** **(a)** *N*-Acetylgalactosamine–serine linkage, the major *O*-glycosidic linkage found in glycoproteins. **(b)** *N*-Acetylglucosamine–asparagine linkage, which characterizes N-linked glycoproteins. The *O*-glycosidic linkage is α, whereas the *N*-glycosidic linkage is β.

The astronomical number of possible oligosaccharide structures afforded by these four factors is not realized in cells because cells do not possess specific glycosyltransferases to catalyze the formation of all possible glycosidic linkages. In addition, individual glycoproteins—through their unique conformations—modulate their own interactions with the glycosylating enzymes so that most glycoproteins possess a heterogeneous but reproducible oligosaccharide structure.

The oligosaccharide chains of most glycoproteins are either *O*- or *N*-linked. In **O-linked oligosaccharides**, a GalNAc residue is typically linked to the side chain of a serine or threonine residue. In **N-linked oligosaccharides**, a GlcNAc residue is linked to the amide nitrogen of an asparagine residue. The structures of an *O*-glycosidic and an *N*-glycosidic linkage are compared in Figure 8.33. Additional sugar residues can be attached to the GalNAc or the GlcNAc residue. An individual glycoprotein can contain both *O*- and *N*-linked oligosaccharides and some glycoproteins contain a third type of linkage. In these glycoproteins, the protein is attached to ethanolamine that is linked to a branched oligosaccharide to which lipid is also attached (Section 9.10).

There are four important subclasses of *O*-glycosidic linkages in glycoproteins.

1. The most common *O*-glycosidic linkage is the GalNAc-Ser/Thr linkage mentioned above. Other sugars—for example, galactose and sialic acid—are frequently linked to the GalNAc residue (Figure 8.34a).
2. Some of the 5-hydroxylysine (Hyl) residues of collagen (Figure 4.35) are joined to D-galactose via an *O*-glycosidic linkage (Figure 8.34b). This structure is unique to collagen.
3. The glycosaminoglycans of certain proteoglycans are joined to the core protein via a Gal–Gal–Xyl–Ser structure (Figure 8.34c).
4. In some proteins, a single residue of GlcNAc is linked to serine or threonine (Figure 8.34d).



◀ **Figure 8.34**
Four subclasses of *O*-glycosidic linkages. **(a)** Example of a typical linkage in which *N*-acetylgalactosamine (GalNAc) with attached residues is linked to a serine or threonine residue. **(b)** Linkage found in collagen, where a galactose residue, usually attached to a glucose residue, is linked to hydroxylysine (Hyl). **(c)** Trisaccharide linkage found in certain proteoglycans. **(d)** GlcNAc linkage found in some proteins.

Downloaded From Vet-e-Books.com

BOX 8.3 ABO BLOOD GROUP

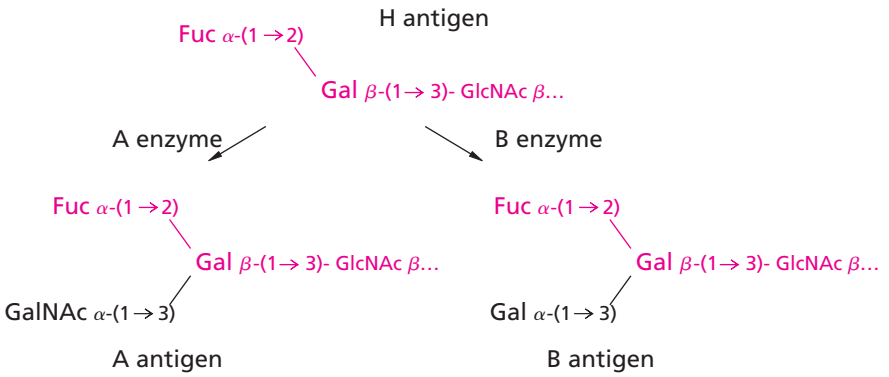
The ABO blood group was first discovered in 1901 by Karl Landsteiner, who received the Nobel Prize in Physiology or Medicine in 1930. Most primates display three different kinds of *O*- or *N*-linked oligosaccharides on their cell surfaces. The core structure of these oligosaccharides is called H antigen. It consists of various combinations of galactose (Gal), fucose (Fuc), *N*-acetylglucosamine (GlcNac), and *N*-acetylneuraminic acid (sialic acid, NeuNAc). These monosaccharides are linked in various ways to form a short branched structure that exhibits considerable microheterogeneity. One of the most common H antigen structures is shown in the figure.

The core structure (H antigen) can be modified in various ways. The addition of a GalNAc residue in α-(1 → 3) linkage forms A antigen. This reaction is catalyzed by A enzyme. The addition of Gal in α-(1 → 3) linkage is catalyzed by B enzyme.

If only A antigen is present, a person will have A blood type. If only B antigen is present, the blood type will be B. The AB blood type indicates that both A antigen and B antigen are present on cell surfaces. If neither GalNAc or Gal have been added to the H antigen structure, then neither A antigen nor B antigen will be present and the blood type is O.

The ABO blood group is determined by a single gene on chromosome 9. Human (and other primate) populations contain many alleles of this gene. The original gene encoded A enzyme, which transfers GalNAc. Variants of this gene have altered the specificity of the enzyme so that it no longer recognizes GalNAc but, instead, transfers Gal. These B enzymes differ by several amino acid residues from the allele that encodes the A enzyme. The structures of both types of glycosyltransferase enzymes have been solved and they reveal that only a single amino acid substitution is required to change the specificity from *N*-acetylaminogalactosyltransferase to galactosyltransferase.

The chromosome 9 locus can also contain several alleles that encode nonfunctional proteins. One of the most common mutations is a single base pair deletion near the *N*-terminus



O-Linked oligosaccharides may account for 80% of the mass of mucins. These large glycoproteins are found in mucus, the viscous fluid that protects and lubricates the epithelium of the gastrointestinal, genitourinary, and respiratory tracts. The oligosaccharide chains of mucins contain an abundance of NeuNAc residues and sulfated sugars. The negative charges of these residues are responsible in part for the extended shape of mucins, which contributes to the viscosity of solutions containing mucins.

The biosynthesis of the oligosaccharide chains of glycoproteins requires a battery of specific enzymes in distinct compartments of the cell. In the stepwise synthesis of *O*-linked oligosaccharides, glycosyltransferases catalyze the addition of glycosyl groups donated by nucleotide–sugar coenzymes. The oligosaccharide chains are assembled by addition of the first sugar molecule to the protein, followed by subsequent single-sugar additions to the nonreducing end.

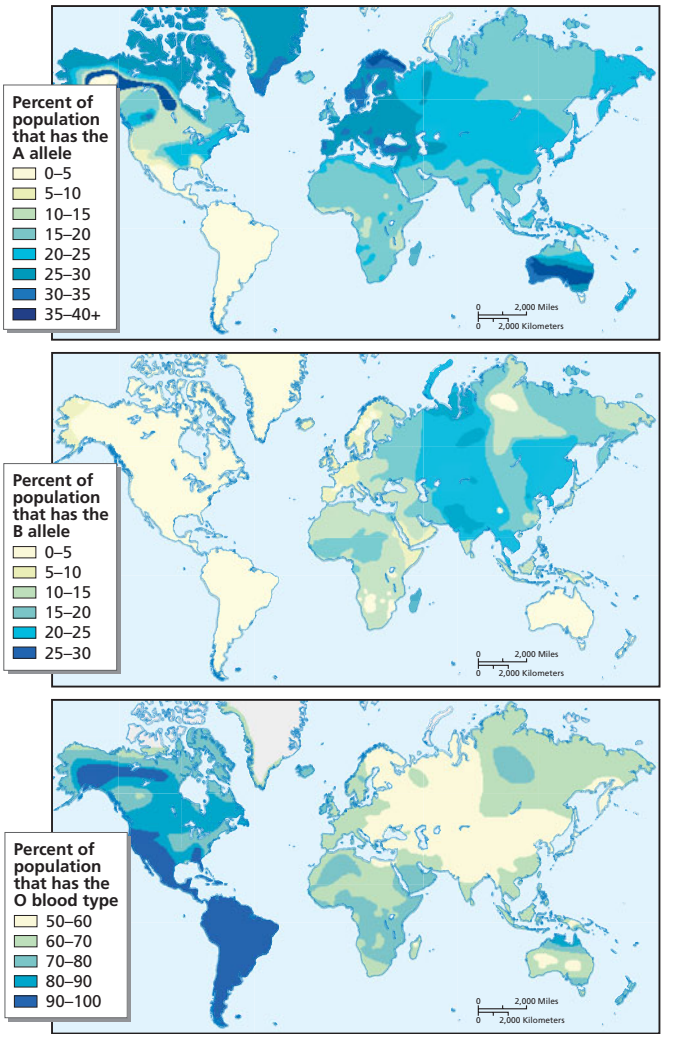
N-Linked oligosaccharides, like *O*-linked oligosaccharides, exhibit great variety in sugar sequence and composition. Most *N*-linked oligosaccharides can be divided into

Downloaded From Vet-e-Books.com

of the coding region. This deletion shifts the reading frame for translation (Section 22.1) making it impossible to synthesize a functional enzyme of either type. This is another example of a human pseudogene. People who are homozygous for these nonfunctional O alleles will not synthesize either A antigen or B antigen and their blood type will be O. (See the Online Medelian Inheritance in Man (OMIM: **ncbi.nlm.nih.gov/omim**) database entry 110300 for an excellent and complete summary of all ABO variants.)

All of your blood cells display some of the unmodified core oligosaccharide (H antigen) even if your blood type is A, B, or AB. This is because not all of the H antigen structures are modified. Under normal circumstances, human plasma will not contain antibodies against H antigen. However, O-type individuals will have antibodies against A antigen and B antigen because these structures are recognized as nonself. If O-type individuals receive a blood transfusion from someone with A, B, or AB blood, they will mount an immune response and reject it. Similarly, if you have A-type blood you will have anti-B antibodies and cannot receive a transfusion from someone with B or AB blood type.

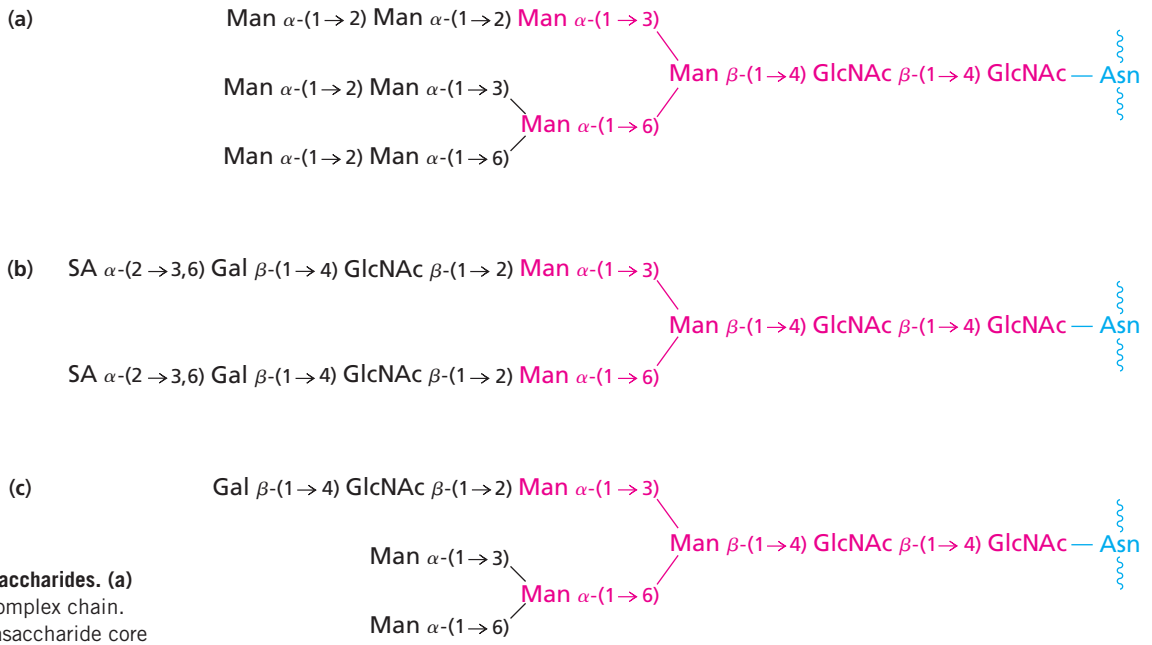
The O allele (pseudogene) is the most common allele in most human populations and the B allele is the most rare. Some Native American populations are homogeneous for the O allele and everyone has type O blood. Type O individuals are perfectly normal, indicating that the absence of the A and B oligosaccharide structures has no effect on normal growth and development (i.e., the allele is neutral in most environments). However, there are some correlations between blood type and disease. People with type O blood, for example, are more susceptible to cholera caused by infections of the bacterium *Vibrio cholerae*. Such selective pressures may be responsible for maintaining the frequencies of A and B alleles in some populations.



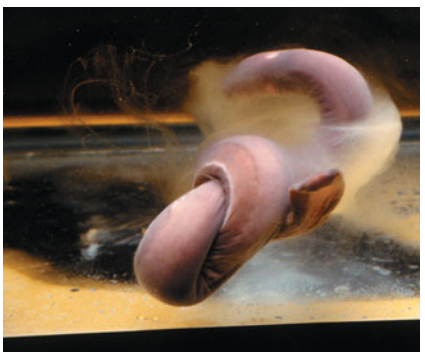
▲ **ABO blood group: distribution of alleles in humans.**

three subclasses: high mannose, complex, and hybrid (Figure 8.35). The appearance of a common core pentasaccharide ($GlcNAc_2Man_3$) in each class reflects a common initial pathway for biosynthesis. The synthesis of *N*-linked oligosaccharides begins with the assembly of a compound consisting of a branched oligosaccharide with 14 residues (nine of which are mannose residues) linked to the lipid dolichol. The entire oligosaccharide chain is transferred to an asparagine residue of a newly synthesized protein, after which the chain is trimmed by the action of glycosidases. High-mannose chains represent an early stage in the biosynthesis of *N*-linked oligosaccharides. Complex oligosaccharide chains result from further removal of sugar residues from high-mannose chains and the addition of other sugar residues, such as fucose, galactose, GlcNAc, and sialic acid (a phenomenon called oligosaccharide processing). These additional sugar residues are donated by nucleotide sugars in reactions catalyzed by glycosyltransferases as in the synthesis of *O*-linked oligosaccharides. In certain cases, a glycoprotein can contain a hybrid oligosaccharide chain, a branched oligosaccharide in which one branch is of the high-mannose type and the other is of the complex type.

Downloaded From Vet-eBooks.com



▲ **Figure 8.35**
Structures of *N*-linked oligosaccharides. (a) High-mannose chain. **(b)** Complex chain. **(c)** Hybrid chain. The pentasaccharide core common to all *N*-linked structures is shown in red. SA represents sialic acid, usually NeuNAc.



▲ **Mucins.** Mucins are heavily glycosylated proteins secreted by the epithelial cells of animals. You are probably familiar with the mucins secreted by cells lining your mouth (saliva), nasal cavity ("snot"), and intestine. The mucin shown here is being secreted by a hagfish.

The synthesis of glycoproteins is discussed in Section 22.10.

Most glycoproteins are secreted from the cell or are bound to the outer surface of the plasma membrane. There are very few glycoproteins in the cytoplasm. With rare exceptions, none of the basic metabolic enzymes are glycosylated. The addition of oligosaccharide chains is tightly coupled to sorting and secretion in eukaryotic cells. The oligosaccharides are attached to specific proteins in the lumen of the endoplasmic reticulum and the groups are modified by various glycosyltransferase enzymes as the proteins move from the ER through the Golgi to the cell surface. The structure of the linked oligosaccharide serves as a marker for sorting proteins into various compartments. For example, some proteins are targeted to the lysosomes, depending on the structure of the oligosaccharide, while others are marked for secretion.

In addition to their roles as markers in sorting and secretion, the presence of one or more oligosaccharide chains on a protein can alter its physical properties, including its size, shape, solubility, electric charge, and stability. Biological properties that can be altered include rate of secretion, rate of folding, and immunogenicity. In a few cases, specific roles for the oligosaccharide chains of glycoproteins have been identified. For example, a number of mammalian hormones are dimeric glycoproteins whose oligosaccharide chains facilitate assembly of the dimer and confer resistance to proteolysis. Also, the recognition of one cell by another that occurs during cell migration or oocyte fertilization can depend in part on the binding of proteins on the surface of one cell to the carbohydrate portions of certain glycoproteins on the surface of the other cell.

Summary

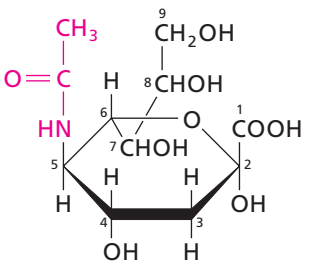
1. Carbohydrates include monosaccharides, oligosaccharides, and polysaccharides. Monosaccharides are classified as aldoses or ketoses or their derivatives.
2. A monosaccharide is designated D or L, according to the configuration of the chiral carbon farthest from the carbonyl carbon atom. Each monosaccharide has 2^n possible stereoisomers, where n is the number of chiral carbon atoms. Enantiomers are nonsuperimposable mirror images of each other. Epimers differ in configuration at only one of several chiral centers.
3. Aldoses with at least five carbon atoms and ketoses with at least six carbon atoms exist principally as cyclic hemiacetals or hemiketals known as furanoses and pyranoses. In these ring structures, the configuration of the anomeric (carbonyl) carbon is designated either α or β. Furanoses and pyranoses can adopt several conformations.
4. Derivatives of monosaccharides include sugar phosphates, deoxy sugars, amino sugars, sugar alcohols, and sugar acids.

Downloaded From Vet-eBooks.com

5. Glycosides are formed when the anomeric carbon of a sugar forms a glycosidic linkage with another molecule. Glycosides include disaccharides, polysaccharides, and some carbohydrate derivatives.
6. Homoglycans are polymers containing a single type of sugar residue. Examples of homoglycans include the storage polysaccharides starch and glycogen and the structural polysaccharides cellulose and chitin.
7. Hetero glycans contain more than one type of sugar residue. They are found in glycoconjugates such as proteoglycans, peptidoglycans, and glycoproteins.
8. Proteoglycans are proteins linked to chains of repeating disaccharides. Proteoglycans are prominent in the extracellular matrix and in connective tissues such as cartilage.
9. The cell walls of many bacteria are made of peptidoglycans that are heteroglycan chains linked to peptides. Peptidoglycan molecules are extensively cross-linked, essentially converting peptidoglycan into a rigid macromolecule that defines the shape of a bacterium and protects the plasma membrane.
10. Glycoproteins are proteins containing covalently bound oligosaccharides. The oligosaccharide chains of most glycoproteins are either *O*-linked to serine or threonine residues or *N*-linked to asparagine residues and exhibit great variety in structure and sugar composition.

Problems

1. Identify each of the following:
   (a) Two aldoses whose configuration at carbons 3, 4, and 5 matches that of D-fructose.
   (b) The enantiomer of D-galactose.
   (c) An epimer of D-galactose that is also an epimer of D-mannose.
   (d) A ketose that has no chiral centers.
   (e) A ketose that has only one chiral center.
   (f) Monosaccharide residues of cellulose, amylose, and glycogen.
   (g) Monosaccharide residues of chitin.
2. Draw Fischer projections for (a) L-mannose, (b) L-fucose (6-deoxy-L-galactose), (c) D-xylitol, and (d) D-iduronate.
3. Describe the general structural features of glycosaminoglycans.
4. Honey is an emulsion of microcrystalline D-fructose and D-glucose. Although D-fructose in polysaccharides exists mainly in the furanose form, solution or crystalline D-fructose (as in honey) is a mixture of several forms with β-D-fructopyranose (67%) and β-D-fructofuranose (25%) predominating. Draw the Fischer projection for D-fructose and show how it can cyclize to form both of the cyclized forms above.
5. Sialic acid (*N*-acetyl-α-D-neuraminic acid) is often found in *N*-linked oligosaccharides that are involved in cell-cell interactions. Cancer cells synthesize much greater amounts of sialic acid than normal cells. Derivatives of sialic acid have been proposed as anticancer agents to block cell-surface interactions between normal and cancerous cells. Answer the following questions about the structure of sialic acid.
   (a) Is it an α or a β anomeric form?
   (b) Will sialic acid mutorotate between α and β anomeric forms?
   (c) Is this a "deoxy" sugar?
   (d) Will the open-chain form of sialic acid be an aldehyde or a ketone?
   (e) How many chiral carbons are there in the sugar ring?



Sialic acid

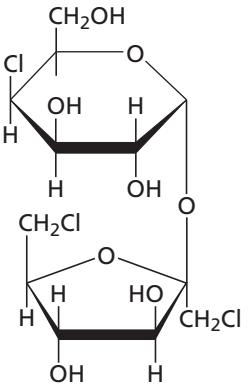
6. How many stereoisomers are possible for glucopyranose and for fructofuranose? How many are D sugars in each case, and how many are L sugars?
7. Draw the structure of each of the following molecules and label each chiral carbon with an asterisk:
   (a) α-D-Glucose 1-phosphate.
   (b) 2-Deoxy-β-D-ribose 5-phosphate.
   (c) D-Glyceraldehyde 3-phosphate.
   (d) L-Glucuronate.
8. In aqueous solution, almost all D-glucose molecules (>99%) are in the pyranose form. Other aldoses have a greater proportion of molecules in the open-chain form. D-Glucose may have evolved to be the predominant hexose because it is less likely than its isomers to react with and damage cellular proteins. Explain why D-glucose reacts less than other aldoses with the amino groups of proteins.
9. Why is the β-D-glucopyranose form of glucose more abundant than α-D-glucopyranose in aqueous solution?
10. The relative orientations of substituents on ribose rings are determined by the conformation of the ring itself. If the ribose is part of a polymeric molecule, then ring conformation will affect overall polymer structure. For example, the orientation of ribose phosphate substituents connecting monomeric nucleoside units is important in determining the overall structure of nucleic acid molecules. In one major form of DNA (B-DNA), the ribofuranose rings adopt an envelope conformation in which C-2′ carbon is above the plane defined by C-1, C-3, C-4, and the ring oxygen (C-2′ endo conformation). Draw the envelope structure of D-ribose 5-phosphate with a nucleoside base (B) attached in a β-anomeric position at the C-1 carbon.
11. In a procedure for testing blood glucose, a drop of blood is placed on a paper strip impregnated with the enzyme glucose oxidase and all the reagents necessary for the reaction

$$\beta\text{-}D\text{-Glucose} + O_2 \longrightarrow D\text{-Gluconolactone} + H_2O_2$$

The H_2O_2 produced causes a color change on the paper that indicates how much glucose is present. Since glucose oxidase is specific for the β anomer of glucose, why can the total blood glucose be measured?
12. Sucralose (registered under the brand name Splenda®) is a nonnutritive (noncaloric) sweetener that is approximately 600 times sweeter than sugar. Since sucralose is heat stable, it can be used in cooking and baking. The structure of sucralose is shown below.

Downloaded From Vet-eBooks.com

Name the disaccharide that is used as a starting substrate for the synthesis of sucralose. What chemical modifications have been made to the starting disaccharide?



13. Draw Haworth projections for the following glycosides:
    (a) Isomaltose [α-D-glucopyranosyl-(1 → 6)-α-D-glucopyranose].
    (b) Amygdalin, a compound in the pits of certain fruits, which has a —CH(CN)C_6H_5 group attached to C-1 of β-D-glucopyranosyl-(1 → 6)-β-D-glucopyranose.
    (c) The *O*-linked oligosaccharide in collagen (β-D-galactose attached to a 5-hydroxylysine residue)
14. Keratan sulfate is a glycosaminoglycan composed primarily of the following repeating disaccharide unit:—Gal β(1 → 4) GlcNAc6S β(1 → 3)—. The acetylated sugar has a sulfate ester on C-6. Keratan sulfate is found in cornea, bone, and cartilage aggregated with other glycosaminoglycans such as chondroitin sulfate. Draw a Haworth projection of the repeating disaccharide unit found in keratan sulfate.
15. A number of diseases result from hereditary deficiencies in specific glycosidases. In these diseases, certain glycoproteins are incompletely degraded and oligosaccharides accumulate in tissues. Which of the *N*-linked oligosaccharides in Figure 8.35 would be affected by deficiencies of the following enzymes?
    (a) *N*-Acetyl-β-glucosaminyl asparagine amidase
    (b) β-Galactosidase
    (c) Sialidase
    (d) Fucosidase
16. A carbohydrate–amino acid polymer that is a potent inhibitor of influenza virus has been synthesized. The virus is thought to be inactivated when multiple sialyl groups bind to viral surface proteins. Draw the chemical structure of the carbohydrate portion of this polymer (below, where X represents the rest of the polymer).

    NeuNAc α-(2 → 3) Gal β-(1 → 4) Glu β-(1 →)-X

17. Imagine that you could take a pill containing β-glucosidase. If, after taking this pill, you ate this textbook, what would it taste like? Would it taste any different if you could marinate it overnight in a solution containing β-glucosidase? Should publishers use flavored ink in order to encourage students to eat their textbooks?

Selected Readings

General

Collins, P. M., ed. (1987). *Carbohydrates* (London and New York: Chapman and Hall).

El Khadem, H. S. (1988). *Carbohydrate Chemistry: Monosaccharides and Their Derivatives* (Orlando, FL: Academic Press).

Li, X., Glaser, D., Li, W., Johnson, W. E., O'Brien, S. J., Beauchamp, G. K., and Brand, J. G. (2009). Analyses of sweet receptor gene (Tas1r2) and preference for sweet stimuli in species of Carnivora. *J. Hered.* 100(Supplement 1):S90–S100.

Li, X., Li, W., Wang, H., Cao, J., Maehashi, K., Huang, L., Bachmanov, A. A., Reed, D. R., Legrand-Defretin, V., Beauchamp, G. K., and Brand, J. G. (2005). Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. *PloS Genet.* 1(1): e3. DOI:10.1371/journal.pgen.0010003

Nodulation Factors

Dénarié, J., and Debellé, F. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* 65:503–535.

Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., and Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nature Communications.* DOI:10.1038/ncomms1009

Mergaert, P., Van Montagu, M., and Holsters, M. (1997). Molecular mechanisms of Nod factor diversity. *Mol. Microbiol.* 25:811–817.

Thoden, J. B., Kim, J., Raushel, F. M., and Holden, H. M. (2002). Structural and kinetic studies of sugar binding to galactose mutarotase from *Lactococcus lactis*. *J. Biol. Chem.* 277:45458–45465.

Proteoglycans

Heinegård, D., and Oldberg, Å. (1989). Structure and biology of cartilage and bone matrix noncollagenous macromolecules. *FASEB J.* 3:2042–2051.

Iozzo, R. V. (1999). The biology of the small leucine-rich proteoglycans: functional network of interactive proteins. *J. Biol. Chem.* 274:18843–18846.

Iozzo, R. V., and Murdoch, A. D. (1996). Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* 10:598–614.

Kjellén, L., and Lindahl, U. (1991). Proteoglycans: structures and interactions. *Annu. Rev. Biochem.* 60:443–475.

Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* 75:39–68.

Glycoproteins

Drickamer, K., and Taylor, M. E. (1998). Evolving views of protein glycosylation. *Trends Biochem. Sci.* 23:321–324.

Dwek, R. A., Edge, C. J., Harvey, D. J., Wormald, M. R., and Parekh, R. B. (1993). Analysis of glycoprotein-associated oligosaccharides. *Annu. Rev. Biochem.* 62:65–100.

Fudge, D. S., Levy, N., Chiu, S., and Gosline, J. M. (2005). Composition, morphology and mechanics of hagfish slime. *J. Exp. Biol.* 208:4613–4625.

Downloaded From Vet-e-Books.com

Lairson, L. L., Henrissat, B., Davies, G., and Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* 77:521–555.

Lechner, J., and Wieland, F. (1989). Structure and biosynthesis of prokaryotic glycoproteins. *Annu. Rev. Biochem.* 58:173–194.

Marionneau, S., Caileau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B. Ruvoën, N., Clément, M., and Le Pendu, J. (2001). ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie.* 83:565–573.

Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002). The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat. Struct. Biol.* 9:685–690.

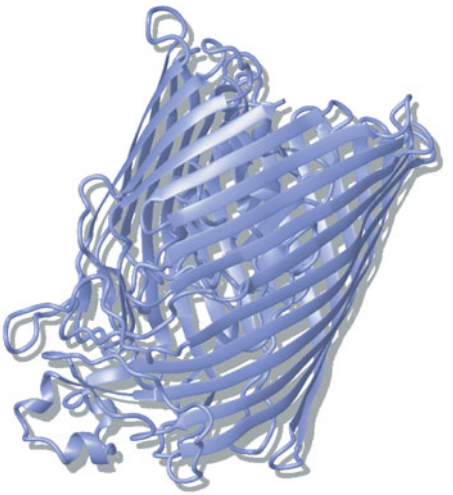
Rademacher, T. W., Parekh, R. B., and Dwek, R. A. (1988). Glycobiology. *Annu. Rev. Biochem.* 57:785–838

Rudd, P. M., and Dwek, R. A. (1997). Glycosylation: heterogeneity and the 3D structure of proteins. *Crit. Rev. Biochem. Mol. Biol.* 32:1–100.

Strous, G. J., and Dekker, J. (1992). Mucin-type glycoproteins. *Crit. Rev. Biochem. Mol. Biol.* 27:57–92.

Downloaded From Vet-eBooks.com





Lipids and Membranes

In this chapter, we consider **lipids**, (*lipo*-, fat) a third major class of biomolecules. Lipids—like proteins and carbohydrates—are essential components of all living organisms. However, unlike these other types of biomolecules, lipids have widely varied structures. They are often defined as water-insoluble (or only sparingly soluble) organic compounds found in biological systems but that's a very broad definition. Lipids are very soluble in nonpolar organic solvents. They are either hydrophobic (nonpolar) or amphipathic (containing both nonpolar and polar regions).

We begin this chapter with a discussion of the structures and functions of the different classes of lipids. In the second part of the chapter, we examine the structures of biological membranes whose properties as cellular barriers depend on the properties of their lipids. Finally, we describe the principles of membrane transport and transmembrane signaling pathways.

In this article, we therefore present and discuss a fluid mosaic model of membrane structure, and propose that it is applicable to most biological membranes, such as plasmalemmal and intracellular membranes, including the membranes of different cell organelles such as mitochondria and chloroplasts.

—S.J. Singer and G.L. Nicholson (1972)

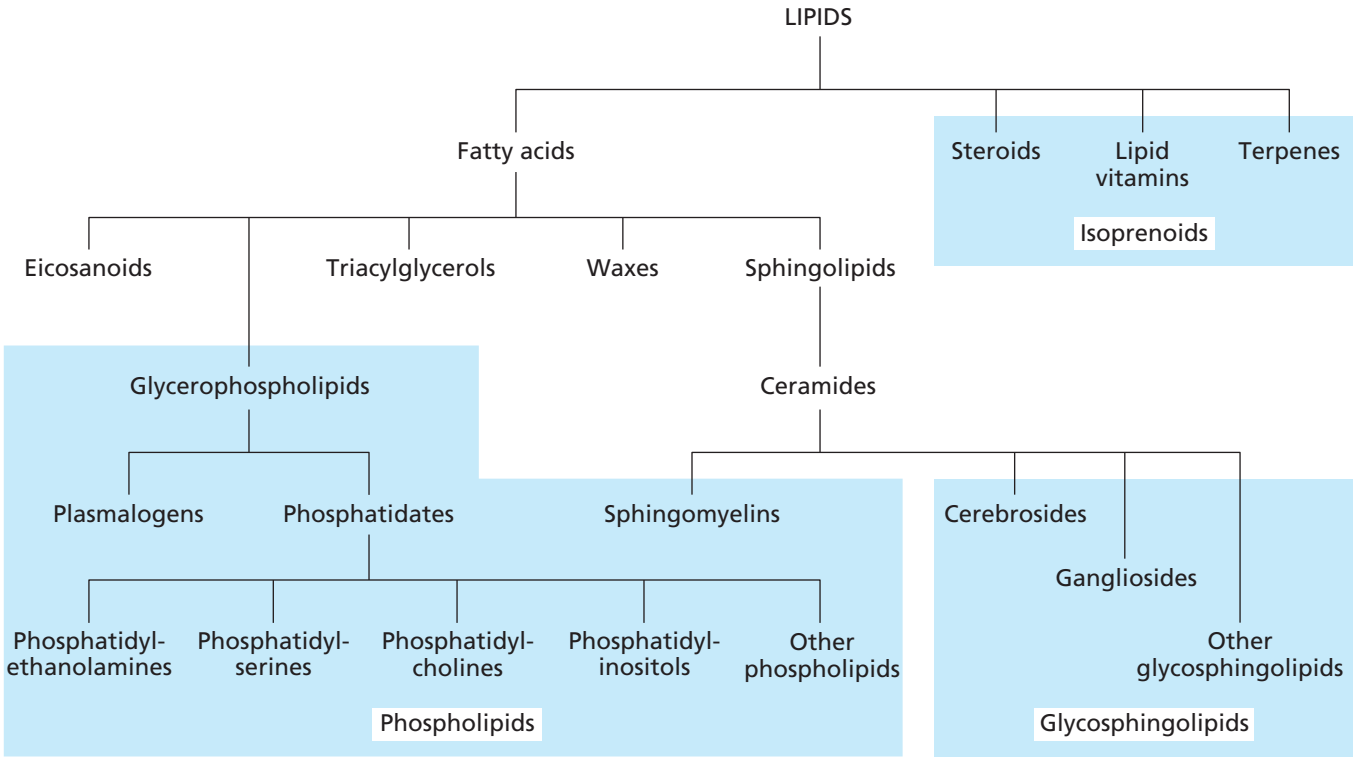
9.1 Structural and Functional Diversity of Lipids

Figure 9.1 shows the major types of lipids and their structural relationships to one another. The simplest lipids are the **fatty acids** that have the general formula R—COOH, where R represents a hydrocarbon chain composed of various lengths of —CH_2-(methylene) units. Fatty acids are components of many more complex types of lipids, including triacylglycerols, glycerophospholipids, and sphingolipids. Lipids containing phosphate groups are called **phospholipids** and lipids containing both sphingosine and carbohydrate groups are called **glycosphingolipids**. Steroids, lipid vitamins, and terpenes are related to the five-carbon molecule isoprene and are therefore called **isoprenoids**. The name *terpenes* has been applied to all isoprenoids but usually is restricted to those that occur in plants.

Lipids have diverse biological functions as well as diverse structures. Biological membranes contain a variety of amphipathic lipids including glycerophospholipids and

Top: Ribbon structure of the transmembrane portion of porin FhuA from *Escherichia coli* (see Figure 9.28).

Downloaded From Vet-e-Books.com



▲ **Figure 9.1**
Structural relationships of the major classes of lipids. Fatty acids are the simplest lipids. Many other types of lipids either contain or are derived from fatty acids. Glycerophospholipids and sphingomyelins contain phosphate and are classified as phospholipids. Cerebrosides and gangliosides contain sphingosine and carbohydrate and are classified as glycosphingolipids. Steroids, lipid vitamins, and terpenes are called isoprenoids because they are related to the five-carbon molecule isoprene rather than to fatty acids.

sphingolipids. In some organisms, triacylglycerols (fats and oils) function as intracellular storage molecules for metabolic energy. Fats also provide animals with thermal insulation and padding. Waxes in cell walls, exoskeletons, and skins protect the surfaces of some organisms. Some lipids have highly specialized functions. For example, steroid hormones regulate and integrate a host of metabolic activities in animals and eicosanoids participate in the regulation of blood pressure, body temperature, and smooth-muscle contraction in mammals. Gangliosides and other glycosphingolipids are located at the cell surface and can participate in cellular recognition.

9.2 Fatty Acids

Fatty acid biosynthesis is discussed in Chapter 16.

More than 100 different fatty acids have been identified in various species. Fatty acids differ from one another in the length of their hydrocarbon tails, the number of carbon–carbon double bonds, the positions of the double bonds in the chains, and the number of branches. Some of the fatty acids commonly found in mammals are shown in Table 9.1.

All fatty acids have a carboxyl group (—COOH) at their "head." This is why they are acids. The pK_a of this group is about 4.5 to 5.0 so it is ionized at physiological pH (—COO^-). Fatty acids are a form of detergent because they have a long hydrophobic tail and a polar head (Section 2.4). As expected, the concentration of *free* fatty acid in cells is quite low because high concentrations of free fatty acids could disrupt membranes. Most fatty acids are components of more complex lipids. They are joined to other molecules by an ester linkage at the terminal carboxyl group.

Fatty acids can be referred to by either International Union of Pure and Applied Chemistry (IUPAC) names or common names. Common names are used for the most frequently encountered fatty acids.

The number of carbon atoms in the most abundant fatty acids ranges from 12 to 20 and is almost always an even number since fatty acids are synthesized by the sequential addition of two-carbon units. In IUPAC nomenclature, the carboxyl carbon is labeled C-1 and the remaining carbon atoms are numbered sequentially. In common

Downloaded From Vet-e-Books.com

Table 9.1 Some common fatty acids (anionic forms)

Number of carbons	Number of double bonds	Common name	IUPAC name	Molecular formula	Melting point, °C
12	0	Laurate	Dodecanoate	$CH_3(CH_2)_{10}COO^{\ominus}$	44
14	0	Myristate	Tetradecanoate	$CH_31(CH_21)_{12}COO^{\ominus}$	52
16	0	Palmitate	Hexadecanoate	$CH_31(CH_21)_{14}COO^{\ominus}$	63
18	0	Stearate	Octadecanoate	$CH_3(CH_2)_{16}COO^{\ominus}$	70
20	0	Arachidate	Eicosanoate	$CH_3(CH_2)_{18}COO^{\ominus}$	75
22	0	Behenate	Docosanoate	$CH_3(CH_2)_{20}COO^{\ominus}$	81
24	0	Lignocerate	Tetracosanoate	$CH_3(CH_2)_{22}COO^{\ominus}$	84
16	1	Palmitoleate	*cis*-Δ^9-Hexadecenoate	$CH_3(CH_2)_5CH{=}CH(CH_2)_7COO^{\ominus}$	−0.5
18	1	Oleate	*cis*-Δ^9-Octadecenoate	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COO^{\ominus}$	13
18	2	Linoleate	cis, *cis*-$\Delta^{9,12}$-Octadecadienoate	$CH_3(CH_2)_4(CH{=}CHCH_2)_2(CH_2)_6COO^{\ominus}$	−9
18	3	Linolenate	all *cis*-$\Delta^{9,12,15}$-Octadecatrienoate	$CH_3CH_2(CH{=}CHCH_2)_3(CH_2)_6COO^{\ominus}$	−17
20	4	Arachidonate	all *cis*-$\Delta^{5,8,11,14}$-Eicosatetraenoate	$CH_3(CH_2)_4(CH{=}CHCH_2)_4(CH_2)_2COO^{\ominus}$	−49

nomenclature, Greek letters are used to identify the carbon atoms. The carbon adjacent to the carboxyl carbon (C-2 in IUPAC nomenclature) is designated α, and the other carbons are lettered β, γ, δ, and ε and so on (Figure 9.2). The Greek letter ω (omega) specifies the carbon atom farthest from the carboxyl group, whatever the length of the hydrocarbon tail. (ω is the last letter in the Greek alphabet.)

Fatty acids without a carbon–carbon double bond are classified as **saturated**, whereas those with at least one carbon–carbon double bond are classified as **unsaturated**. Unsaturated fatty acids with only one carbon–carbon double bond are called **monounsaturated** and those with two or more are called **polyunsaturated**. The configuration of the double bonds in unsaturated fatty acids can be either *cis* or *trans* . The configuration is usually *cis* in naturally occurring fatty acids (see Box. 9.2).

The positions of double bonds are indicated by the symbol Δ^n in IUPAC nomenclature, where the superscript *n* indicates the lower-numbered carbon atom of each

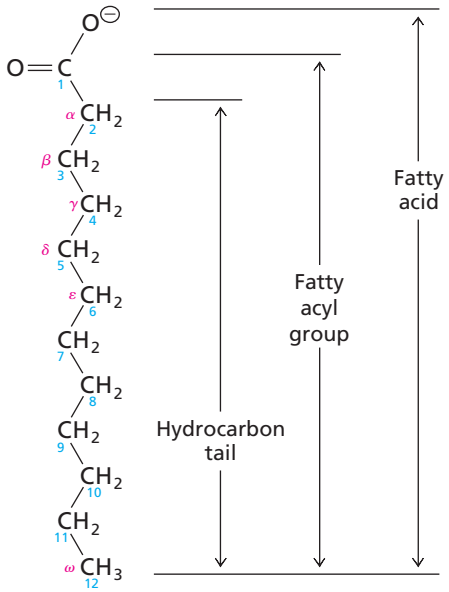
BOX 9.1 COMMON NAMES OF THE FATTY ACIDS

Laurate present in oil from the laurel plant (*Laurus nobilis*) (1873)
Myristate oil from nutmeg (*Myristica fragrans*) (1848)
Palmitate from palm oil (1857)
Stearate from French *stéarique* referring to fat from steers, or tallow (1831)
Arachidate present in oil from peanuts (*Arachis hypogaea*) (1866)
Behenate a corruption of "ben" from ben-nut = seeds of the Horseradish tree (1873)
Lignocerate probably from Latin *lignum* ("wood") (~1900)
Oleate from Latin *oleum* ("oil") (1899)
Linoleate found in linseed oil (*lin* + *oleate*) (1857)



▲ **The African oil palm tree, *Elaeis guineensis*.** Palm oil is a complex mixture of saturated and unsaturated fatty acids but palmitate makes up 44% of the total. The presence of such a large amount of saturated fatty acid means that palm oil is a semisolid at room temperature. It can never be "virgin" or "extra virgin" (see Box 16.6).

Downloaded From Vet-e-Books.com



◀ **Figure 9.2**
Structure and nomenclature of fatty acids. Fatty acids consist of a long hydrocarbon tail terminating with a carboxyl group. Since the pK_a of the carboxyl group is approximately 4.5 to 5.0, fatty acids are anionic at physiological pH. In IUPAC nomenclature, carbons are numbered beginning with the carboxyl carbon. In common nomenclature, the carbon atom adjacent to the carboxyl carbon is designated α, and the remaining carbons are lettered β, γ, δ, and so on. The carbon atom farthest from the carboxyl carbon is designated the ω carbon, whatever the length of the tail. The fatty acid shown, laurate (or dodecanoate), has 12 carbon atoms and contains no carbon–carbon double bonds.

double-bonded pair (Table 9.1). The double bonds of most polyunsaturated fatty acids are separated by a methylene group and are therefore not conjugated.

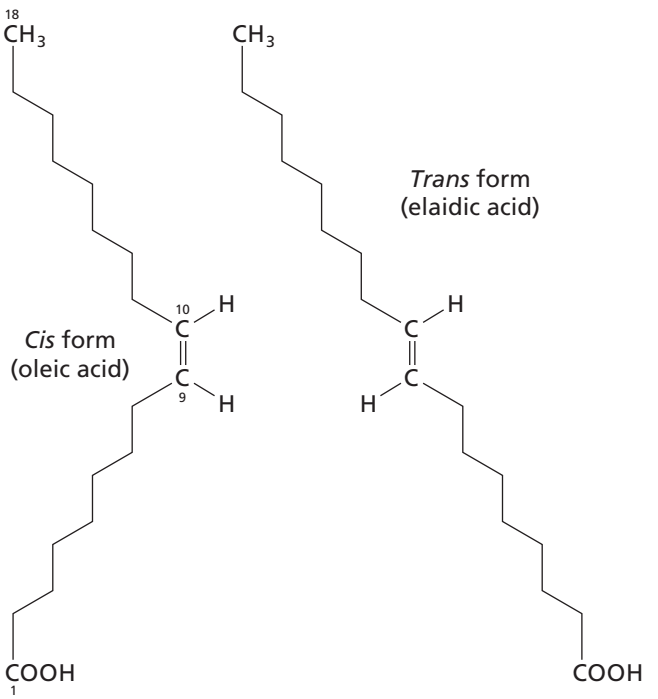
A shorthand notation for identifying fatty acids uses two numbers separated by a colon—the first refers to the number of carbon atoms in the fatty acid and the second refers to the number of carbon–carbon double bonds, with their positions indicated as superscripts following a Greek symbol, Δ. In this notation, palmitate is written as 16:0, oleate as 18:1 Δ^9, and arachidonate as 20:4 $\Delta^{5,8,11,14}$. Unsaturated fatty acids can

BOX 9.2 *TRANS* FATTY ACIDS AND MARGARINE

The configuration of most double bonds in unsaturated fatty acids is *cis* but some fatty acids in the human diet have the *trans* configuration. *Trans* fatty acids can come from animal sources such as dairy products and ruminant meats. However, most of the edible *trans* fatty acids consumed in Western industrialized countries are present as hydrogenated vegetable oils in some margarines or shortenings. Dietary *trans* monounsaturated fatty acids can increase plasma levels of cholesterol and triglycerides and their ingestion may increase the risk of cardiovascular disease. More work is required to establish the exact level of risk.

Plant oils such as corn oil and sunflower oil can be converted to "spreadable" semisolid substances known as margarines. Margarines can be produced by the partial or complete hydrogenation of double bonds in plant oils. The hydrogenation process itself not only saturates the carbon–carbon double bonds of fatty acid esters but can also change the configuration of the remaining double bonds from *cis* to *trans*. The physical properties of these *trans* fatty acids are similar to those of saturated fatty acids.

In order to reduce consumption of *trans* fatty acids, many margarines are now produced from plant oils without hydrogenation by adding other edible components such as skim milk powder.



▲ ***Cis* and *trans* forms of Δ^9-octadecanoate.** (Left) Oleate (*cis*-Δ^9-octadecanoate). (Right) the *trans* configuration after hydrogenation.

Downloaded From Vet-e-Books.com

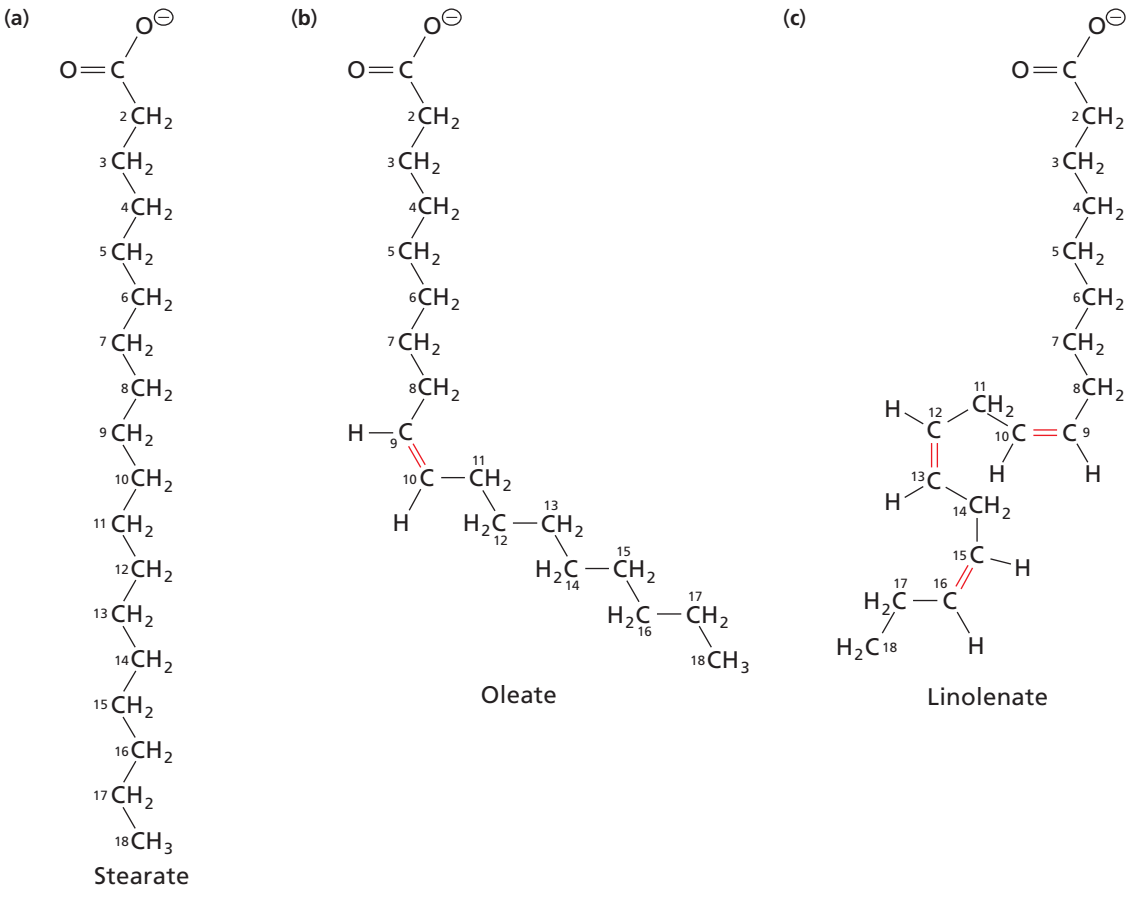
also be described by the location of the last double bond in the chain. This double bond is usually found three, six, or nine carbon atoms from the end of the chain. Such fatty acids are called ω - 3 (e.g., 18:3 $\Delta^{9,12,15}$), ω - 6 (e.g., 18:2 $\Delta^{9,12}$), or ω - 9 (e.g., 18:1 Δ^{9}).

The physical properties of saturated and unsaturated fatty acids differ considerably. Typically, saturated fatty acids are waxy solids at room temperature (22°C) whereas unsaturated fatty acids are liquids at this temperature. The length of the hydrocarbon chain of a fatty acid and its degree of unsaturation influence the melting point. Compare the melting points listed in Table 9.1 for the saturated fatty acids laurate (12:0), myristate (14:0), and palmitate (16:0). As the lengths of the hydrocarbon tails increase, the melting points of the saturated fatty acids also increase. The number of van der Waals interactions among neighboring hydrocarbon tails increases as the tails get longer so more energy is required to disrupt the interactions.

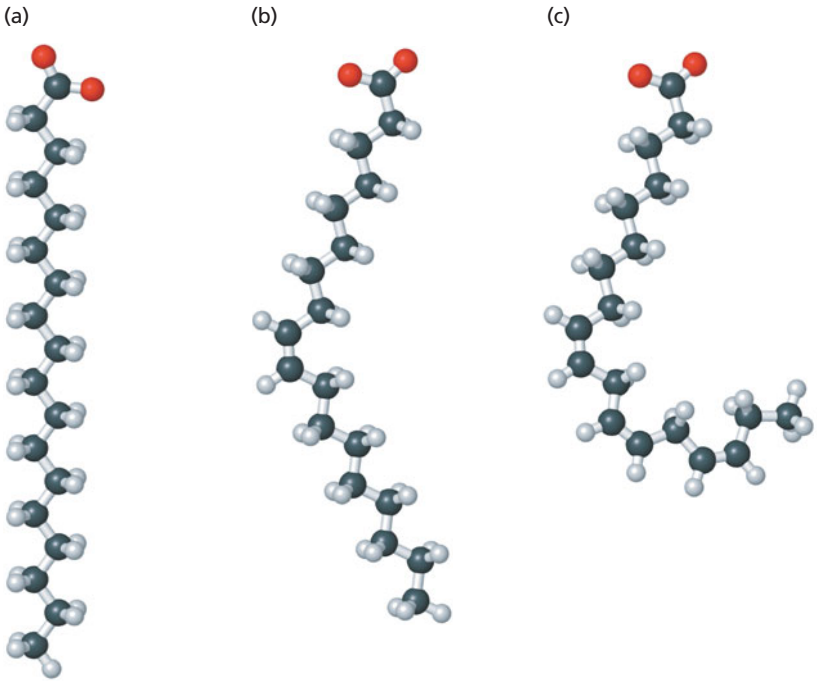
Compare the structures of stearate (18:0), oleate (18:1), and linolenate (18:3) in Figures 9.3 and 9.4. The saturated hydrocarbon tail of stearate is flexible since rotation can occur around every carbon–carbon bond. In a crystal of stearic acid, the hydrocarbon chains are extended and pack together closely. The presence of *cis* double bonds in oleate and linolenate produces pronounced bends in the hydrocarbon chains since rotation around double bonds is hindered. These bends prevent close packing and extensive van der Waals interactions among the hydrocarbon chains. Consequently, *cis* unsaturated fatty acids have lower melting points than saturated fatty acids. As the degree of unsaturation increases, fatty acids become more fluid. Note that stearic acid (melting point 70°C) is a solid at body temperature but oleic acid (melting point 13°C) and linolenic acid (melting point −17°C) are both liquids.

As mentioned earlier, free fatty acids occur only in trace amounts in living cells. Most fatty acids are esterified to glycerol or other backbone compounds to form more complex lipid molecules. In esters and other derivatives of carboxylic acids, the RC═O moiety contributed by the acid is called the acyl group. In common nomenclature,

▼ **Figure 9.3**
Structures of three C_{18} fatty acids. (a) Stearate (octadecanoate), a saturated fatty acid. **(b)** Oleate (*cis*-Δ^9-octadecenoate) a monounsaturated fatty acid. **(c)** Linolenate (all-*cis*-$\Delta^{9,12,15}$-octadecatrienoate) a polyunsaturated fatty acid. The *cis* double bonds produce kinks in the tails of the unsaturated fatty acids. Linolenate is a very flexible molecule that can assume a variety of conformations.



Downloaded From Vet-e-Books.com



◀ **Figure 9.4**
Stearate (left), oleate (center), and linolenate (right). Color key: carbon, grey; hydrogen, white; oxygen, red.

complex lipids that contain specific fatty acyl groups are named after the parent fatty acid. For example, esters of the fatty acid laurate are called lauroyl esters, and esters of linoleate are called linoleoyl esters. (A lauryl group is the alcohol analog of the lauroyl acyl group). The relative abundance of particular fatty acids varies with the type of organism, type of organ (in multicellular organisms), and food source. The most abundant fatty acids in animals are usually oleate (18:1), palmitate (16:0), and stearate (18:0).

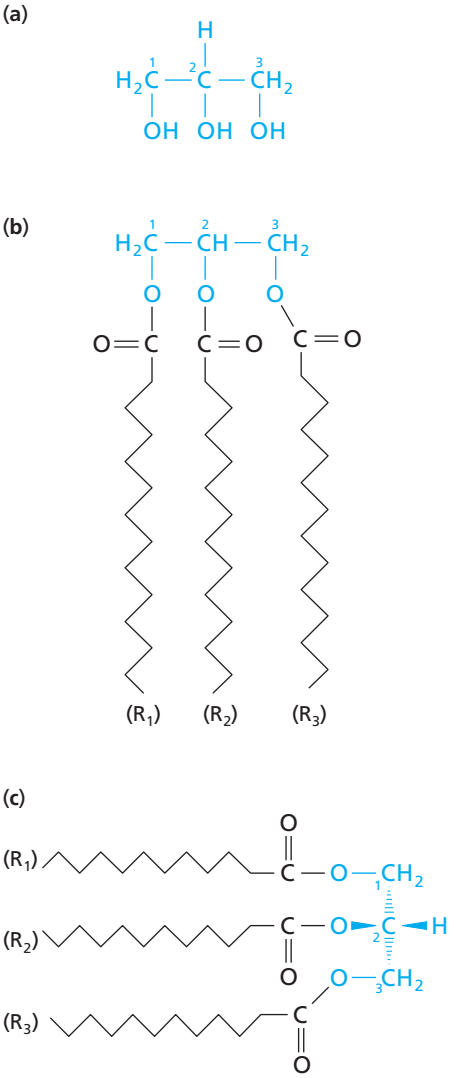
Mammals require certain dietary polyunsaturated fatty acids that they cannot synthesize, such as linoleate (18:2 $\Delta^{9,12}$) and linolenate (18:3 $\Delta^{9,12,15}$). These fatty acids are called essential fatty acids. Mammals can synthesize other polyunsaturated fatty acids from an adequate supply of linoleate and linolenate. (Recall that many vitamins are also essential components of the mammalian diet because mammals cannot synthesize them. In addition to vitamins and essential fatty acids, we will see in Chapter 17 that many amino acids cannot be synthesized in mammals.)

Linolenate is an omega-3 (ω-3) fatty acid since the last double bond is three carbon atoms from the tail end of the molecule. Omega-3 fatty acids are very popular dietary supplements. They are enriched in fish oils, which is why many people recommend that you include fish and fish oils in your diet. Linolenate is an essential fatty acid so your diet must include an adequate supply of this omega-3 fatty acid. This adequate amount is readily supplied in the typical diet of people all over the world, which is why essential fatty acid deficiency is rare. The market for supplemental omega-3 fatty acids is driven by other factors. The most important benefit is protection against cardiovascular disease. The scientific evidence indicates that extra amounts of omega-3 fatty acids provide a small benefit in terms of reducing the risk of heart attacks, particularly a second heart attack. None of the other claims are based on reproducible double-blind test results after controlling for other factors. Eating fish, for example, will not make you smarter.

Many fatty acids besides those listed in Table 9.1 are present in nature. For example, fatty acids containing cyclopropane rings are found in bacteria. Branched-chain fatty acids are common components of bacterial membranes and also occur on the feathers of ducks. Many other fatty acids are rare and have highly specialized functions.

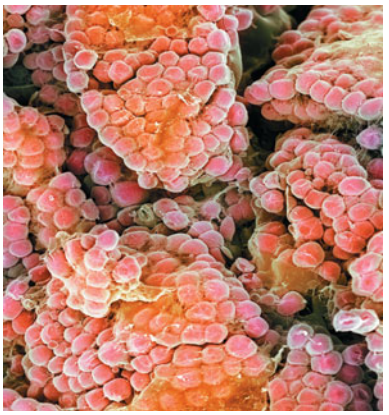
9.3 Triacylglycerols

As their name implies, **triacylglycerols** (historically called triglycerides) are composed of three fatty acyl residues esterified to glycerol, a three-carbon sugar alcohol (Figure 9.5). Triacylglycerols are very hydrophobic.



▲ **Figure 9.5**
Structure of a triacylglycerol. Glycerol **(a)** is the backbone to which three fatty acyl residues are esterified **(b)**. Although glycerol is not chiral, C-2 of a triacylglycerol is chiral when the acyl groups bound to C-1 and C-3 (R_1 and R_3) differ. The general structure of a triacylglycerol is shown in **(c)**, oriented for comparison with the structure of L-glyceraldehyde (Figure 8.1). This orientation allows stereospecific numbering of glycerol derivatives with C-1 at the top and C-3 at the bottom.

Downloaded From Vet-e-Books.com



▲ **Figure 9.6**
Adipocytes. This is a colorized scanning electron micrograph of clusters of adipocytes. A fat droplet occupies most of the volume of each adipocyte.

The structures and functions of lipoproteins are discussed in Section 16.1B.

Fats and oils are mixtures of triacylglycerols. They can be solids (fats) or liquids (oils), depending on their fatty acid compositions and on the temperature. Triacylglycerols containing only saturated long chain fatty acyl groups tend to be solids at body temperature and those containing unsaturated or short chain fatty acyl groups tend to be liquids. A sample of naturally occurring triacylglycerols can contain as many as 20 to 30 molecular species that differ in their fatty acid constituents. Tripalmitin, found in animal fat, contains three residues of palmitic acid. Triolein, which contains three oleic acid residues, is the principal triacylglycerol in olive oil.

In most cells, triacylglycerols coalesce as fat droplets. These droplets are sometimes seen near mitochondria in cells that rely on fatty acids for metabolic energy. In mammals, most fat is stored in adipose tissue that is composed of specialized cells known as adipocytes. Each adipocyte contains a large fat droplet that accounts for nearly the entire volume of the cell (Figure 9.6). Although distributed throughout the bodies of mammals, most adipose tissue occurs just under the skin and in the abdominal cavity. Extensive subcutaneous fat serves both as a storage depot for energy and as thermal insulation and is especially pronounced in aquatic mammals.

9.4 Glycerophospholipids

Triacylglycerols are not found in biological membranes. The most abundant lipids in most membranes are **glycerophospholipids** (also called phosphoglycerides). Glycerophospholipids, like triacylglycerols, have a glycerol backbone. The simplest glycerophospholipids are the, phosphatidates—they consist of two fatty acyl groups esterified to C-1 and C-2 of glycerol 3-phosphate (Table 9.2). Note that there are three fatty acyl groups esterified to glycerol in triacylglycerols whereas there are only two fatty acyl groups (R_1 and R_2) in the glycerophospholipids. The distinguishing feature of the glycerophospholipids is the presence of a phosphate group on C-3 of the glycerol backbone. The structures of glycerophospholipids can be drawn as derivatives of L-glycerol 3-phosphate with the C-2 substituent on the left in a Fischer projection, as in Table 9.2. For simplicity, we usually show these compounds as stereochemically uncommitted structures.

KEY CONCEPT

Glycerophospholipids have polar heads and long, hydrophobic fatty acid tails.

Phosphatidates are present in small amounts as intermediates in the biosynthesis or breakdown of more complex glycerophospholipids. In most glycerophospholipids, the phosphate group is esterified to both glycerol and another compound bearing an —OH group. Table 9.2 lists some common types of glycerophospholipids. Note that glycerophospholipids are amphipathic molecules with a polar head and long, nonpolar tails. The structures of three types of glycerophospholipids—phosphatidylethanolamine, phosphatidylserine, and phosphatidylcholine—are shown in Figure 9.7.

Each type of glycerophospholipid consists of a family of molecules with the same polar head group and different fatty acyl chains. For example, human red blood cell membranes contain at least 21 different species of phosphatidylcholine that differ from one another in the fatty acyl chains esterified at C-1 and C-2 of the glycerol backbone. In general, glycerophospholipids have saturated fatty acids esterified to C-1 and unsaturated fatty acids esterified to C-2. The major membrane glycerophospholipids in *Escherichia coli* are phosphatidylethanolamine and phosphatidylglycerol.

KEY CONCEPT

Many important lipids are derivatives of glycerol (see Box 16.1).

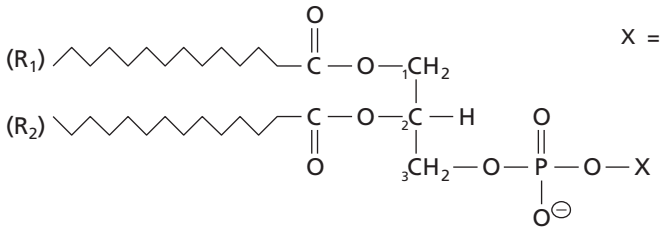


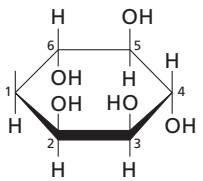
▲ **Yellow jacket wasp.** The venom of wasps, bees, and snakes contains phospholipases.

A variety of phospholipases can be used to dissect glycerophospholipid structures and determine the identities of their individual fatty acids. The specific positions of fatty acids in glycerophospholipids can be determined by using phospholipase A_1 and phospholipase A_2 that specifically catalyze the hydrolysis of the ester bonds at C-1 and C-2, respectively (Figure 9.8). Phospholipase A_2 is the major phospholipase in pancreatic juice and it is responsible for the digestion of membrane phospholipids in the diet. It is also present in snake, bee, and wasp venom. High concentrations of the products of the action of phospholipase A_2 can disrupt cell membranes. Thus, injection of snake venom into the blood can result in life-threatening lysis of the membranes of red blood cells. Phospholipase C catalyzes hydrolysis of the P—O bond between glycerol and

Downloaded From Vet-e-Books.com

Table 9.2 Some common types of glycerophospholipids



Precursor of X (HO—X)	Formulas of —O—X	Name of resulting glycerophospholipid
Water	—H	Phosphatidate
Choline	$-CH_2CH_2\overset{\oplus}{N}(CH_3)_3$	Phosphatidylcholine
Ethanolamine	$-CH_2CH_2\overset{\oplus}{N}H_3$	Phosphatidylethanolamine
Serine	$-CH_2-CH(\overset{\oplus}{N}H_3)-COO^{\ominus}$	Phosphatidylserine
Glycerol	$-CH_2CH(OH)-CH_2OH$	Phosphatidylglycerol
Phosphatidyl-glycerol	$-CH_2CH(OH)-CH_2-O-P(O^{\ominus})(=O)-O-CH_2-CH(OCOCH(OCOR_4)-CH_2OCOR_3)$	Diphosphatidylglycerol (Cardiolipin)
myo-Inositol		Phosphatidylinositol

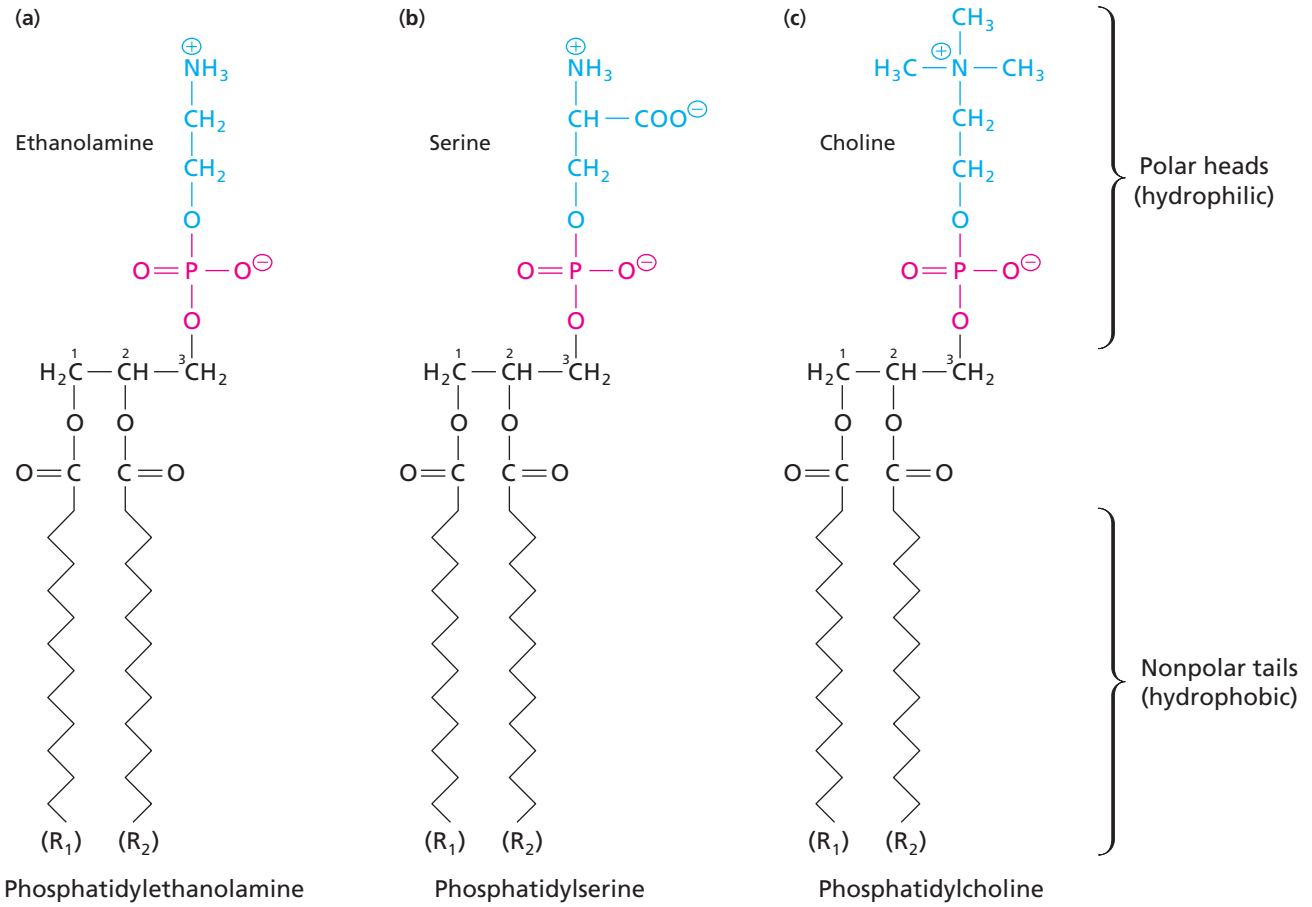
phosphate to liberate diacylglycerol. Phospholipase D converts glycerophospholipids to phosphatidates.

Plasmalogens are the other major type of glycerophospholipids. They differ from phosphatidates because the hydrocarbon substituent on the C-1 hydroxyl group of glycerol is attached by a vinyl ether linkage rather than an ester linkage (Figure 9.9). Ethanolamine or choline is commonly esterified to the phosphate group of plasmalogens. Plasmalogens account for about 23% of the glycerophospholipids in the human central nervous system and are also found in the membranes of peripheral nerve and muscle tissue.

9.5 Sphingolipids

Sphingolipids are the second most abundant lipids in plant and animal membranes. In mammals, sphingolipids are particularly abundant in tissues of the central nervous system. Most bacteria do not have sphingolipids. The structural backbone of sphingolipids is sphingosine (*trans*-4-sphingenine), an unbranched C_{18} alcohol with a *trans* double

Downloaded From Vet-e-Books.com

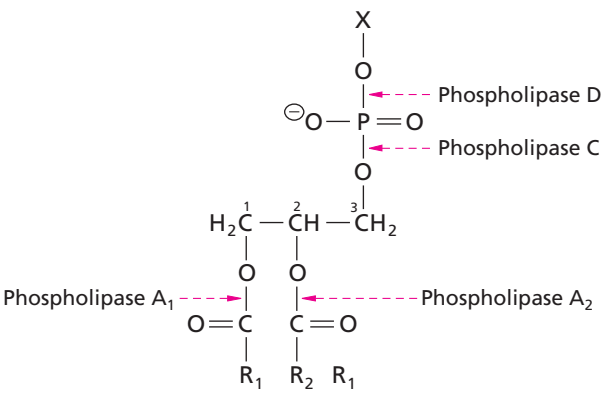


▲ **Figure 9.7**
Structures of (a) phosphatidylethanolamine, (b) phosphatidylserine, and (c) phosphatidylcholine. Functional groups derived from esterified alcohols are shown in blue. Since each of these lipids can contain many combinations of fatty acyl groups, the general name refers to a family of compounds, not to a single molecule.

bond between C-4 and C-5, an amino group at C-2, and hydroxyl groups at C-1 and C-3 (Figure 9.10a). **Ceramide** consists of a fatty acyl group linked to the C-2 amino group of sphingosine by an amide bond (Figure 9.10b). Ceramides are the metabolic precursors of all sphingolipids. The three major families of sphingolipids are the sphingomyelins, the cerebrosides, and the gangliosides. Of these, only sphingomyelins contain phosphate and are classified as phospholipids; cerebrosides and gangliosides contain carbohydrate residues and are classified as glycosphingolipids (Figure 9.1).

In **sphingomyelins**, phosphocholine is attached to the C-1 hydroxyl group of a ceramide (Figure 9.10c). Note the resemblance of sphingomyelin to phosphatidylcholine (Figure 9.7c)—both molecules are zwitterions containing choline, phosphate, and two long hydrophobic tails. Sphingomyelins are present in the plasma membranes of most mammalian cells and are a major component of the myelin sheaths that surround certain nerve cells.

Figure 9.8 ►
Action of four phospholipases. Phospholipases A_1, A_2, C, and D can be used to dissect glycerophospholipid structure. Phospholipases catalyze the selective removal of fatty acids from C-1 or C-2 or convert glycerophospholipids to diacylglycerols or phosphatidates.



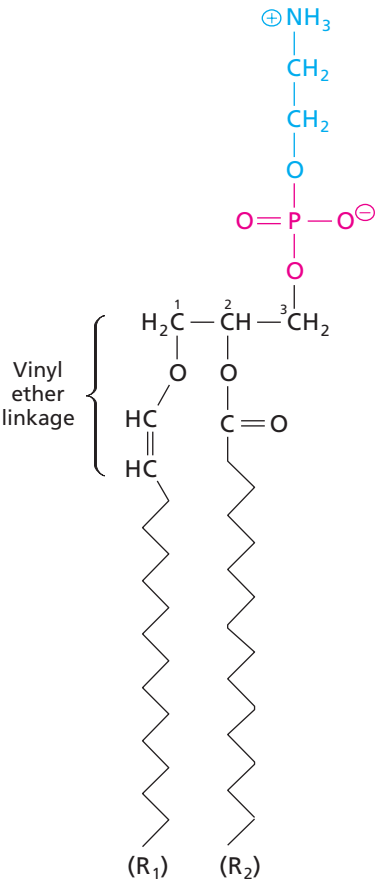
Downloaded From Vet-eBooks.com

Cerebrosides are glycosphingolipids that contain one monosaccharide residue attached by a β-glycosidic linkage to C-1 of a ceramide. Galactocerebrosides, also known as galactosylceramides, have a single β-D-galactosyl residue as a polar head group (Figure 9.11). Galactocerebrosides are abundant in nerve tissue and account for about 15% of the lipids of myelin sheaths. Many other mammalian tissues contain glucocerebrosides, ceramides with a β-D-glucosyl head group. In some glycosphingolipids, a linear chain of up to three more monosaccharide residues is attached to the galactosyl or glucosyl moiety of a cerebroside.

Gangliosides are more complex glycosphingolipids in which oligosaccharide chains containing *N*-acetylneuraminic acid (NeuNAc) are attached to a ceramide. NeuNAc (Figure 8.15), an acetylated derivative of neuraminic acid, makes the head groups of gangliosides anionic. The structure of a representative ganglioside, G_{M2}, is shown in Figure 9.12. The M in G_{M2} stands for monosialo (i.e., one NeuNAc residue); G_{M2} was the second monosialo ganglioside characterized, thus the subscript 2.

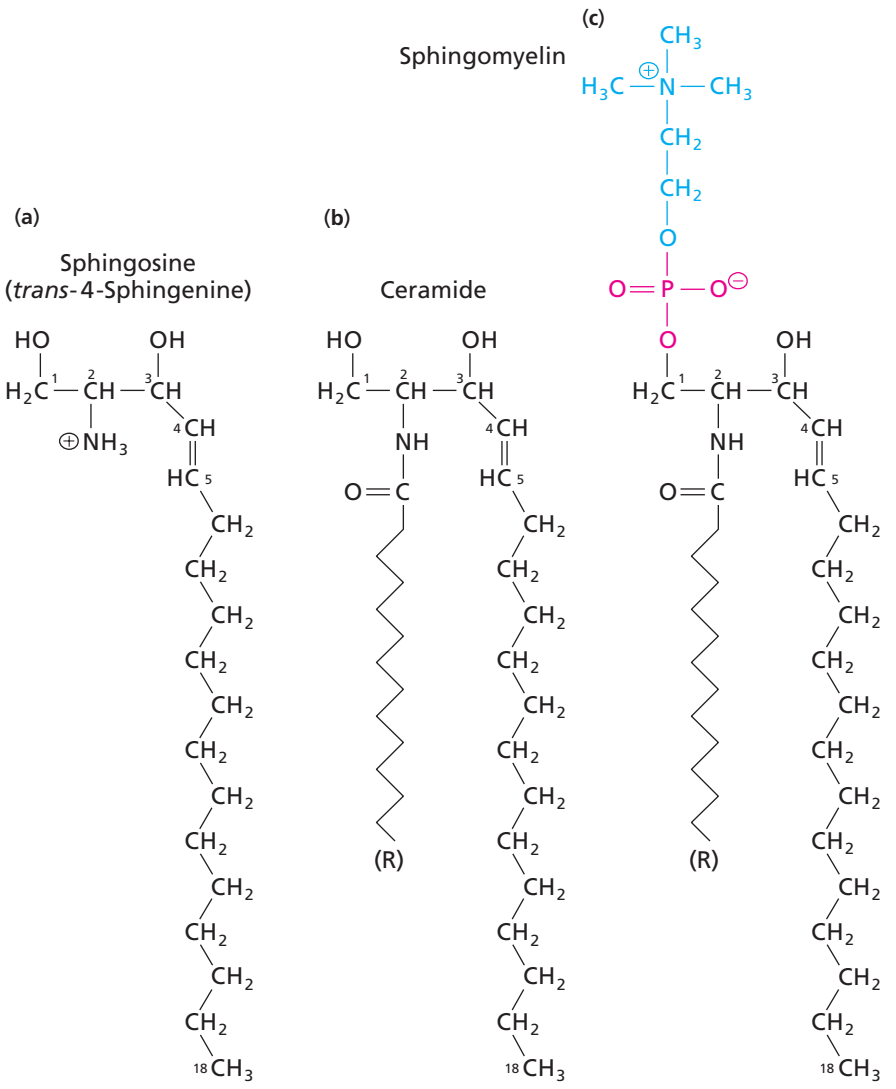
More than 60 varieties of gangliosides have been characterized. Their structural diversity results from variations in the composition and sequence of sugar residues. Ganglioside G_{M1}, for example, is similar to ganglioside G_{M2} shown in Figure 9.12 except that it has an additional β-D-galactose residue attached to the terminal *N*-acetyl-β-D-galactosamine residue via a β-(1 → 4) linkage. In all gangliosides, the ceramide is linked through its C-1 to a β-glucosyl residue, which in turn is bound to a β-galactosyl residue.

Gangliosides are present on cell surfaces with the two hydrocarbon chains of the ceramide moiety embedded in the plasma membrane and the oligosaccharides on the



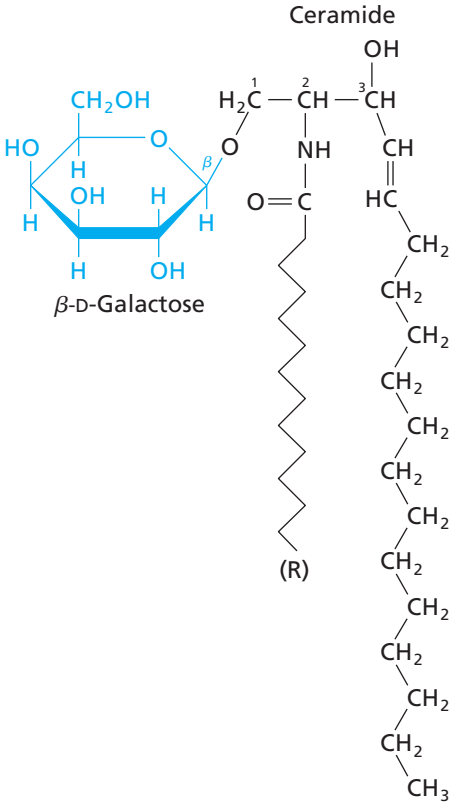
▲ **Figure 9.9**
Structure of an ethanolamine plasmalogen. A hydrocarbon is linked to the C-1 hydroxyl group of glycerol to form a vinyl ether.

Genetic defects associated with lipid metabolism are described in Chapter 16.

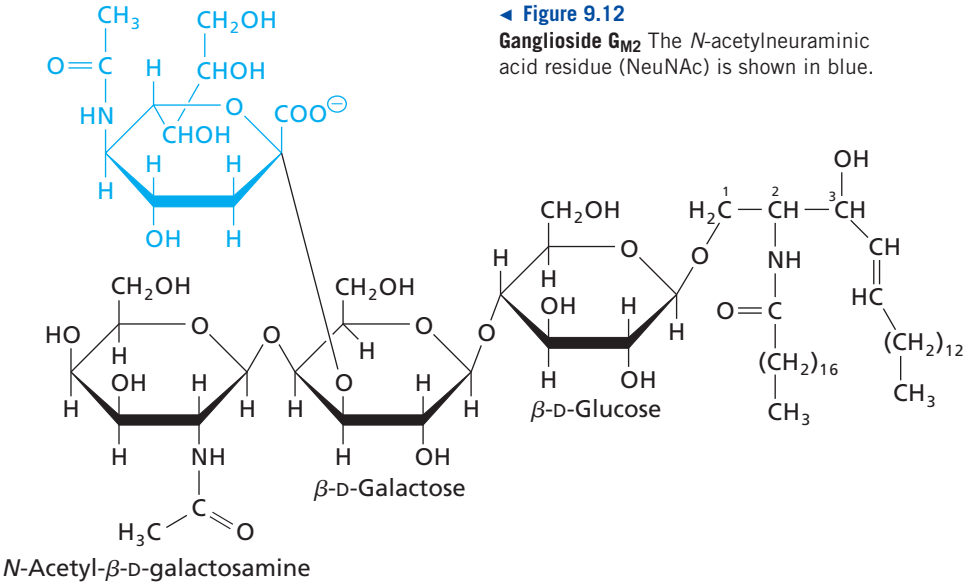


◀ **Figure 9.10**
Structures of sphingosine, ceramide, and sphingomyelin. (a) Sphingosine, the backbone for sphingolipids, is a long-chain alcohol with an amino group at C-2. **(b)** Ceramides have a long-chain fatty acyl group attached to the amino group of sphingosine. **(c)** Sphingomyelins have a phosphate group (red) attached to the C-1 hydroxyl group of a ceramide and a choline group (blue) attached to the phosphate.

Downloaded From Vet-e-Books.com



▲ **Figure 9.11**
Structure of a galactocerebroside. β-D-Galactose (blue) is attached to the C-1 hydroxyl group of a ceramide (black).

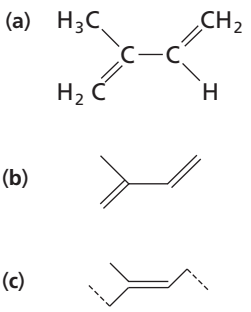


◄ **Figure 9.12**
Ganglioside G_{M2} The *N*-acetylneuraminic acid residue (NeuNAc) is shown in blue.

extracellular surface. Gangliosides and other glycosphingolipids are part of the cell surface repertoire of diverse oligosaccharide chains along with glycoproteins. Collectively, these markers provide cells with distinguishing surface markers that can serve in cellular recognition and cell-to-cell communication. Structures similar to the ABO blood group antigens on the surface of human cells (Box. 8.3) can be oligosaccharide components of glycosphingolipids in addition to being linked to proteins to form glycoproteins.

Genetically inherited defects in ganglioside metabolism are responsible for a number of debilitating and often lethal diseases, such as Tay-Sachs disease and generalized gangliosidosis. Certain rare genetic defects lead to deficiencies of enzymes responsible for the degradation of sphingolipids in the lysosomes of cells. In Tay-Sachs disease, there is a deficiency of a hydrolase that catalyzes removal of *N*-acetylgalactosamine from G_{M2}. Accumulation of G_{M2} causes lysosomes to swell leading to tissue enlargement. In the central nervous tissue, where there is little room for expansion, nerve cells die causing blindness, mental retardation, and death.

The exposed carbohydrates on the cell surface also provide convenient receptors for bacteria, viruses, and toxins. For example, cholera toxin, produced by the bacterium *Vibrio cholerae*, binds to the ganglioside G_{M1} of intestinal epithelial cells. Binding stimulates entry of the toxin into the cells where it interferes with normal signaling pathways leading to massive efflux of fluid into the intestine. This often produces death by dehydration.



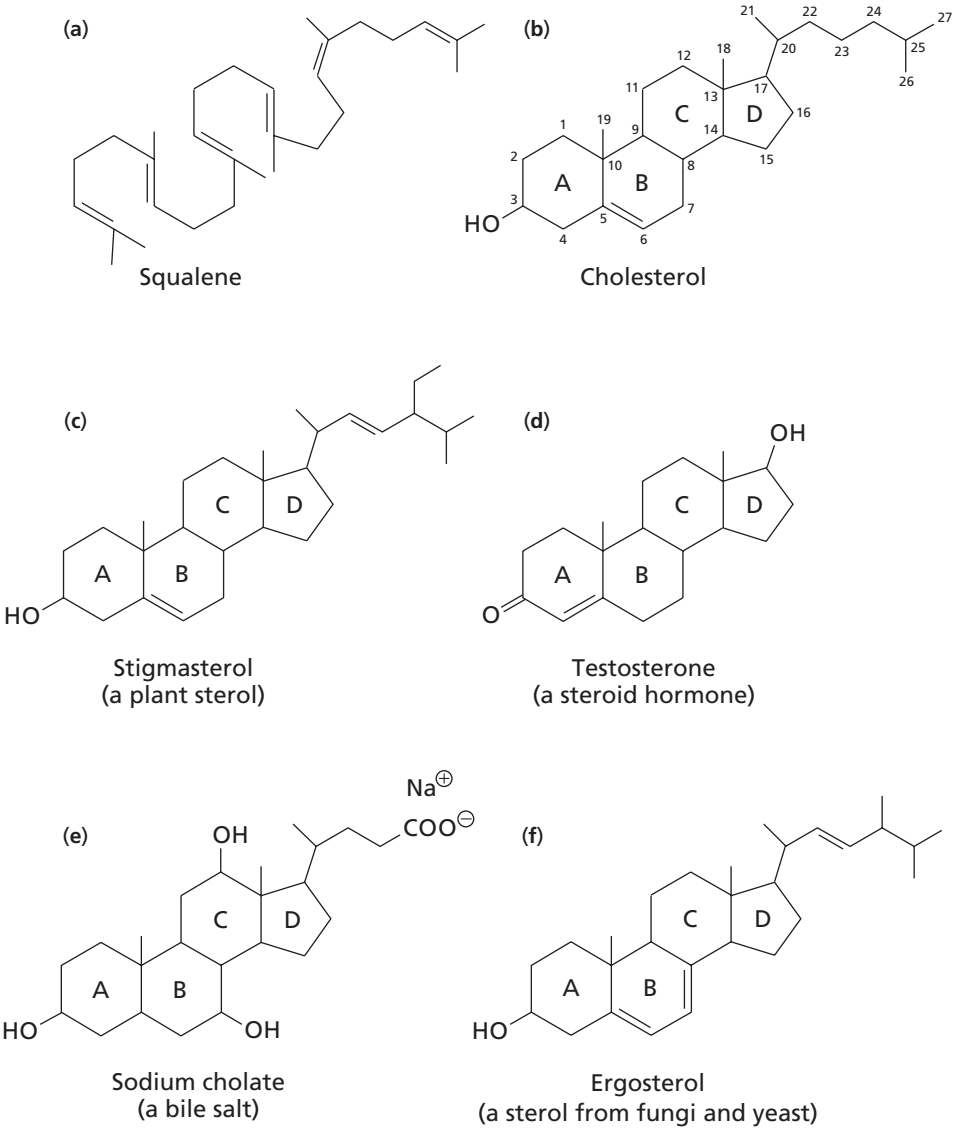
▲ **Figure 9.13**
Isoprene (2-methyl-1,3-butadiene), the basic structural unit of isoprenoids. (a) Chemical structure. **(b)** Carbon backbone. **(c)** Isoprene unit where dashed lines represent covalent bonds to a adjacent units.

9.6 Steroids

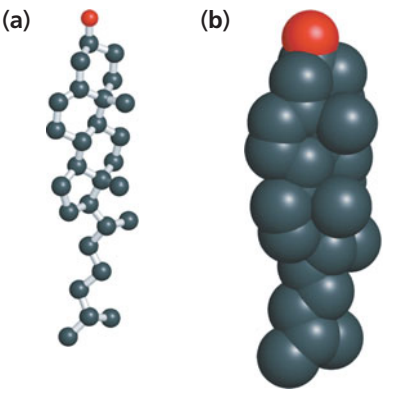
Steroids are a third class of lipids found in the membranes of eukaryotes and, very rarely, in bacteria. Steroids, along with lipid vitamins and terpenes, are classified as isoprenoids because their structures are related to the five-carbon molecule isoprene (Figure 9.13). Steroids contain four fused rings: three six-carbon rings designated A, B, and C and a five-carbon D ring. The characteristic ring structure is derived from squalene (Figure 9.14a). Substituents of the nearly planar ring system can point either down (the α configuration) or up (the β configuration). The structures of several steroids are shown in Figure 9.14.

The steroid cholesterol is an important component of animal plasma membranes but is less common in plants and absent from prokaryotes, protists, and fungi. These species have other steroids (e.g., stigmasterol, ergosterol) that are very similar to cholesterol. Cholesterol is actually a **sterol** because it has a hydroxyl group at C-3. Other steroids include the sterols of plants, fungi, and yeast (which also have a hydroxyl group at C-3); mammalian steroid hormones (such as estrogens, androgens, progestins, and

Downloaded From Vet-e-Books.com



◀ **Figure 9.14**
Structures of several steroids. Squalene **(a)** is the precursor of most steroids. Steroids contain four fused rings (lettered A, B, C, and D). **(b)** Cholesterol. **(c)** Stigmasterol, a common component of plant membranes. **(d)** Testosterone, a steroid hormone involved in male development in animals. **(e)** Sodium cholate, a bile salt, which aids in the digestion of lipids. **(f)** Ergosterol, a compound from fungi and yeast.



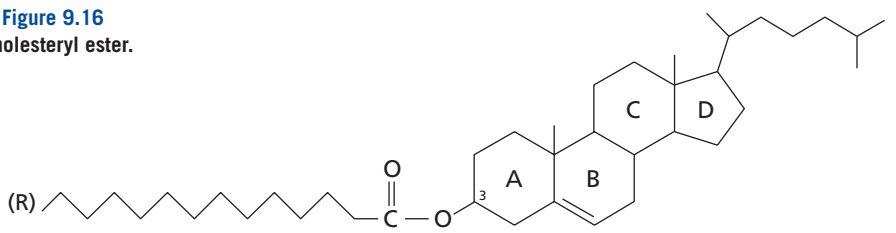
▲ **Figure 9.15**
Cholesterol. (a) Ball-and-stick model with the oxygen atom (red) at the top. Hydrogen atoms are not shown. The fused ring system of cholestrol is almost planar. **(b)** Space-filling model.

adrenal corticosteroids); and bile salts. These steroids differ in the length of the side chain attached to C-17 and in the number and placement of methyl groups, double bonds, hydroxyl groups, and in some cases, keto groups. Prokaryotes use squalene and some related nonsteroid lipids that do not have the complete ring structure of the steroids.

Cholesterol plays an essential role in mammalian biochemistry. It is not only a component of certain membranes but is also a precursor of steroid hormones and bile salts. The fused ring system of cholesterol, shown from the side in Figure 9.15, makes it less flexible than most other lipids. As a result, cholesterol modulates the fluidity of mammalian cell membranes, as we will see later in this chapter.

Steroids are far more hydrophobic than glycerophospholipids and sphingolipids. For example, free cholesterol's maximal concentration in water is only 10^{-8} M. Esterification of a fatty acid to the C-3 hydroxyl group forms a cholesteryl ester (Figure 9.16).

▼ **Figure 9.16**
Cholesteryl ester.



Downloaded From Vet-e-Books.com

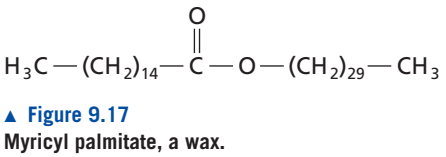
Because the 3-acyl group of the ester is nonpolar, a cholesteryl ester is even more hydrophobic than cholesterol itself. Cholesterol is converted to cholesteryl esters for storage in cells or for transport through the bloodstream. Because they are essentially insoluble in water, cholesterol and its esters must be complexed with phospholipids and amphipathic proteins in lipoproteins for transport (Section 16.1B).

9.7 Other Biologically Important Lipids

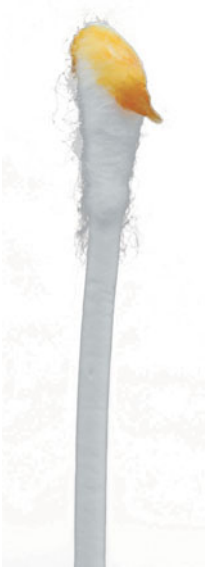
There are many kinds of lipids not found in membranes. These include diverse compounds such as waxes, eicosanoids, and some isoprenoids. Non-membrane lipids have a variety of specialized functions—some of which we have already encountered (e.g., lipid vitamins).

Waxes are nonpolar esters of long-chain fatty acids and long chain monohydroxylic alcohols. For example, myricyl palmitate, a major component of beeswax, is the ester of palmitate (16:0) and the 30-carbon myricyl alcohol (Figure 9.17). The hydrophobicity of myricyl palmitate makes beeswax very insoluble and its high melting point (due to the long, saturated hydrocarbon chains) makes beeswax hard and solid at typical outdoor temperatures. Waxes are widely distributed in nature. They provide protective waterproof coatings on the leaves and fruits of certain plants and on animal skin, fur, feathers, and exoskeletons. Ear wax, also known as cerumen (from the Latin word *cera*, "wax"), is secreted by cells lining the auditory canal. It serves to lubricate the canal and trap particles that could damage the eardrum. Ear wax is a complex mixture made up mostly of long chain fatty acids, cholesterol, and ceramides. It also contains squalene, triacylglycerols, and true waxes (about 10% of the weight).

Eicosanoids are oxygenated derivatives of C_{20} polyunsaturated fatty acids such as arachidonic acid. Some examples of eicosanoids are shown in Figure 9.18. Eicosanoids participate in a variety of physiological processes and can also mediate many potentially pathological responses. **Prostaglandins** are eicosanoids that have a cyclopentane ring.

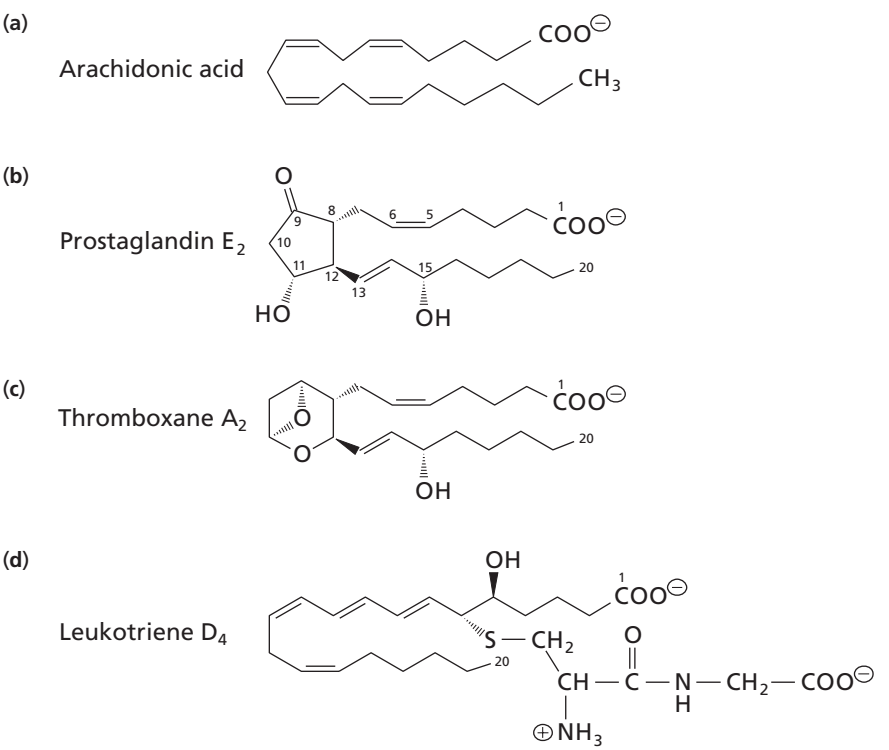


▲ **Figure 9.17**
Myricyl palmitate, a wax.





▲ **Earwax and beeswax are two examples of naturally occurring waxes.**



▲ **Figure 9.18**
Structures of arachidonic acid (a) and three eicosanoids derived from it. Arachidonate is a C_{20} polyunsaturated fatty acid with four *cis* double bonds.

Downloaded From Vet-eBooks.com

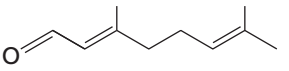
Prostaglandin E_2 can cause constriction of blood vessels, and thromboxane A_2 is involved in the formation of blood clots that in some cases can block the flow of blood to the heart or brain. Leukotriene D_4, a mediator of smooth-muscle contraction, also provokes the bronchial constriction seen in asthmatics. Aspirin (acetylsalicylic acid) alleviates pain, fever, swelling, and inflammation by inhibiting the synthesis of prostaglandins (Box. 16.1).

Some nonmembrane lipids are related to isoprene (Figure 9.13) but they are not steroids. We encountered several of these lipids in Chapter 7. The lipid vitamins A, E, and K are isoprenoids that contain long hydrocarbon chains or fused rings (Section 7.14). Vitamin D is an isoprenoid derivative of cholesterol. There are several carotenes related to retinol (vitamin A). The hydrophobic chain of ubiquinone contains 6–10 isoprenoid units (Section 7.15).

Simple isoprenoids are often called terpenes. They have structures that reveal their formation from isoprene units. Citral is a good example: it is present in many plants and imparts a strong lemon odor (Figure 19.19a). Other isoprenoids are bactoprenol (undecaprenyl alcohol) (Figure 9.19b) and juvenile hormone I (Figure 9.19c) that regulates the expression of genes required for development in insects. Isoprenoids similar to bactoprenol are important lipids in archaebacteria, where they replace fatty acids in most membrane phospholipids (see Box 9.5).

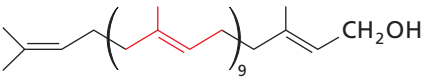
Terpenes can be extensively modified to form a more complex class of lipid called terpenoids. Many of these are cyclic compounds like limonene, which is responsible for the smell of oranges (Figure 19.19d). Gibberellins are multi-ring terpenoids that function as growth hormones in plants (Figure 19.19e).

(a)



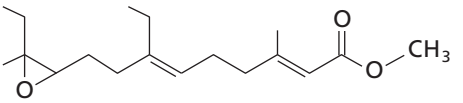
Citral

(b)



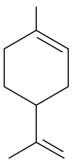
Bactoprenol
(Undecaprenyl alcohol)

(c)



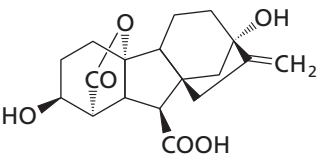
Juvenile hormone I

(d)



Limonene

(e)



Gibberellin GA1

▲ **Figure 9.19**
Some isoprenoids. Note the isoprene unit (red) in bactoprenol.

9.8 Biological Membranes

Biological membranes define the external boundaries of cells and separate compartments within cells. They are essential components of all living cells. A typical membrane consists of two layers of lipid molecules and many embedded proteins.

Biological membranes are not merely passive barriers to diffusion. They have a wide variety of complex functions. Some membrane proteins serve as selective pumps controlling the transport of ions and small molecules into and out of the cell. Membranes are also responsible for generating and maintaining the proton concentration gradients essential for the production of ATP. Receptors in membranes recognize extracellular signals and communicate them to the cell interior.

Many cells have membranes with specialized structures. For example, many bacteria have double membranes: an outer membrane and an inner plasma membrane. The liquid in the periplasmic space between these two membranes contains proteins that carry specific solutes to transport proteins in the inner membrane. The solutes then pass through the inner membrane by an ATP-dependent process. A mitochondrion's smooth outer membrane has proteins that form aqueous channels while its convoluted inner membrane is selectively permeable and has many membrane-bound enzymes. The nucleus also has a double membrane—nuclear contents interact with the cytosol through nuclear pores. The single membrane of the endoplasmic reticulum is highly convoluted. Its extensive network in eukaryotic cells is involved in the synthesis of transmembrane and secreted proteins and of lipids for many membranes.

In this section, we explore the structure of biological membranes. In the remaining sections of this chapter, we discuss the properties and functions of biological membranes.

A. Lipid Bilayers

We saw earlier that detergents in aqueous solutions can spontaneously form monolayers or micelles (Section 2.4). Like detergents, amphipathic glycerophospholipids and glycosphingolipids can form monolayers under some conditions. In cells, these lipids do not pack well into micelles but rather tend to form **lipid bilayers** (Figure 9.20). Lipid bilayers are the main structural component of all biological membranes, including plasma membranes and the internal membranes of eukaryotic cells. The noncovalent

Downloaded From Vet-e-Books.com

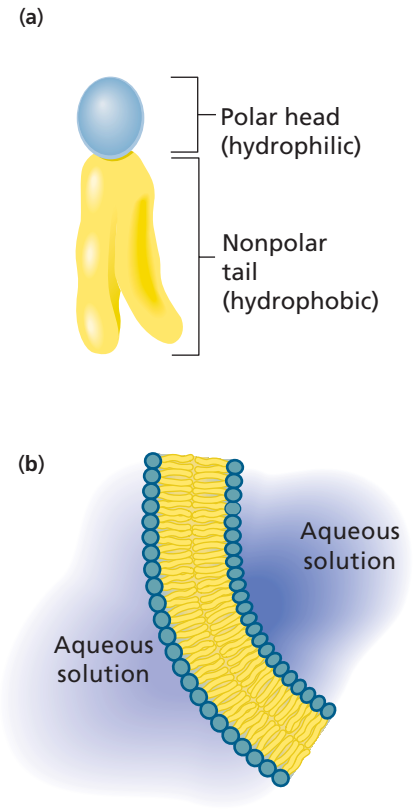
BOX 9.3 GREGOR MENDEL AND GIBBERELLINS

Gregor Mendel studied seven traits in order to come up with the basic laws of heredity. One of the traits was stem length (*Le/le*). The *Le* gene has been cloned and sequenced (Lester et al., 1997). It encodes the enzyme gibberellin 3β-hydroxylase, an enzyme required for the synthesis of the terpenoid gibberellin GA1. The production of gibbberellin GA1 by the normal gene stimulates growth producing a tall pea plant. The mutant gene produces a less active enzyme that synthesizes less hormone and plants homozygous for the mutant allele (*le*) are short.

The mutation is a single nucleotide substitution that converts an alanine codon into a threonine codon (A229T). Another one of Mendel's seven traits is described in Box. 15.3.



► **The stem length mutation.** Tall plants (left) are normal. Mutations in the stem length gene (*Le*) produce short plants (right).



▲ **Figure 9.20**
Membrane lipid and bilayer. (a) An amphipathic membrane lipid. **(b)** Cross-section of a lipid bilayer. The hydrophilic head groups (blue) of each leaflet face the aqueous medium, and the hydrophobic tails (yellow) pack together in the interior of the bilayer.

interactions among lipid molecules in bilayers make membranes flexible and allow them to self-seal. Triacylglycerols, which are very hydrophobic rather than amphipathic, cannot form bilayers and cholesterol, although slightly amphipathic, does not form bilayers by itself.

A lipid bilayer is typically about 5 to 6 nm thick and consists of two sheets, or monolayers (also called leaflets). In each sheet, the polar head groups of amphipathic lipids are in contact with the aqueous medium and the nonpolar hydrocarbon tails point toward the interior of the bilayer (Figure 9.20).

The spontaneous formation of lipid bilayers is driven by the hydrophobic interactions (Section 2.5D). When lipid molecules associate, the entropy of the solvent molecules increases and this favors formation of the lipid bilayer.

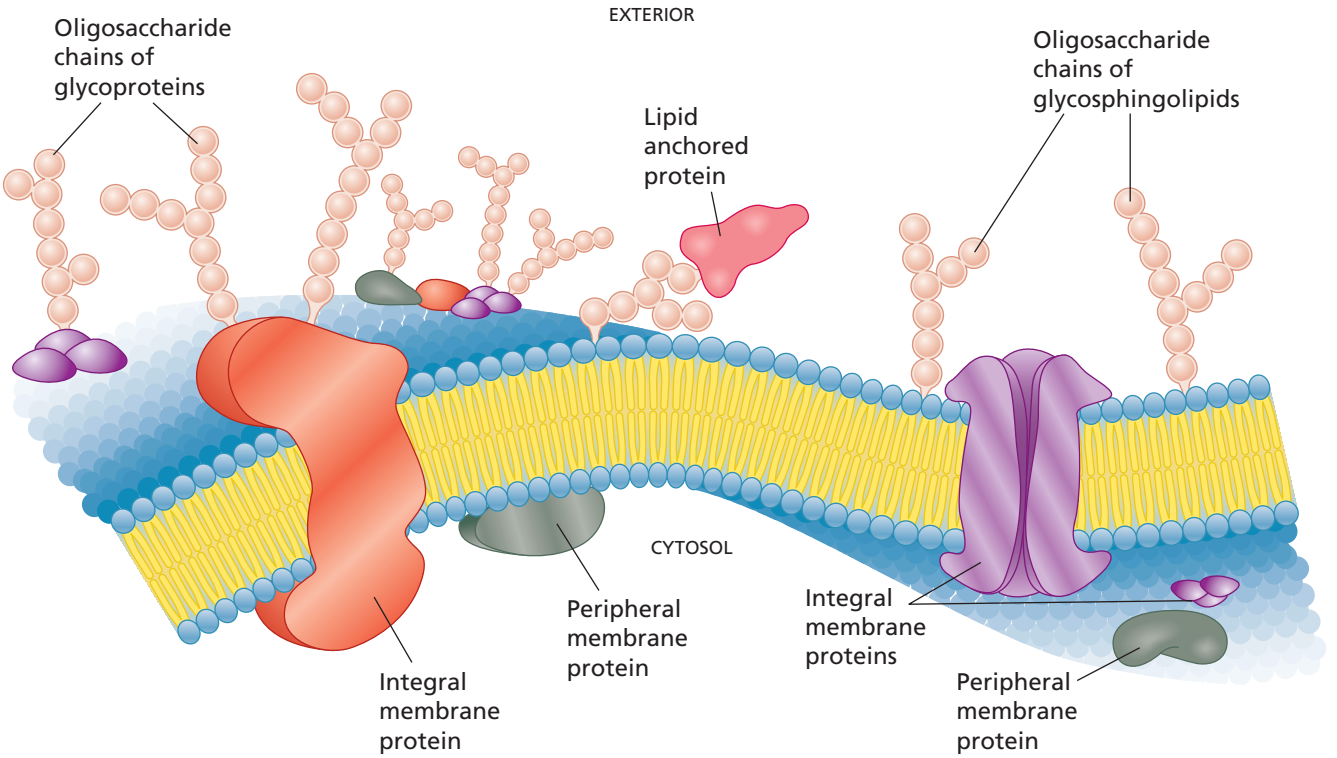
B. Three Classes of Membrane Proteins

Cellular and intracellular membranes contain specialized membrane-bound proteins. These proteins are divided into three classes based on their mode of association with the lipid bilayer: integral membrane proteins, peripheral membrane proteins, and lipid anchored membrane proteins (Figure 9.21).

Integral membrane proteins, also referred to as transmembrane proteins, contain hydrophobic regions embedded in the hydrophobic core of the lipid bilayer. Integral membrane proteins usually span the bilayer completely, with one part of the protein exposed on the outer surface and one part exposed on the inner surface. Some integral membrane proteins are anchored by only a single membrane-spanning portion of the polypeptide chain, whereas other membrane proteins have several transmembrane segments connected by loops at the membrane surface. The membrane-spanning segment is often an α helix containing approximately 20 amino acid residues.

One of the best characterized integral membrane proteins is bacteriorhodopsin (Figure 9.22a). This protein is found in the cytoplasmic membrane of the halophilic (salt-loving) bacterium *Halobacterium halobium*, where it helps harness light energy used in the synthesis of ATP. Bacteriorhodopsin consists of a bundle of seven α helices. The exterior surface of the helical bundle is hydrophobic and interacts directly with lipid molecules in the membrane. The interior surface contains charged amino acid side chains that bind the pigment molecule. Bacteriorhodopsin is one of several α-helical membrane proteins whose structures are known in detail. These α-helix bundle

Downloaded From Vet-e-Books.com

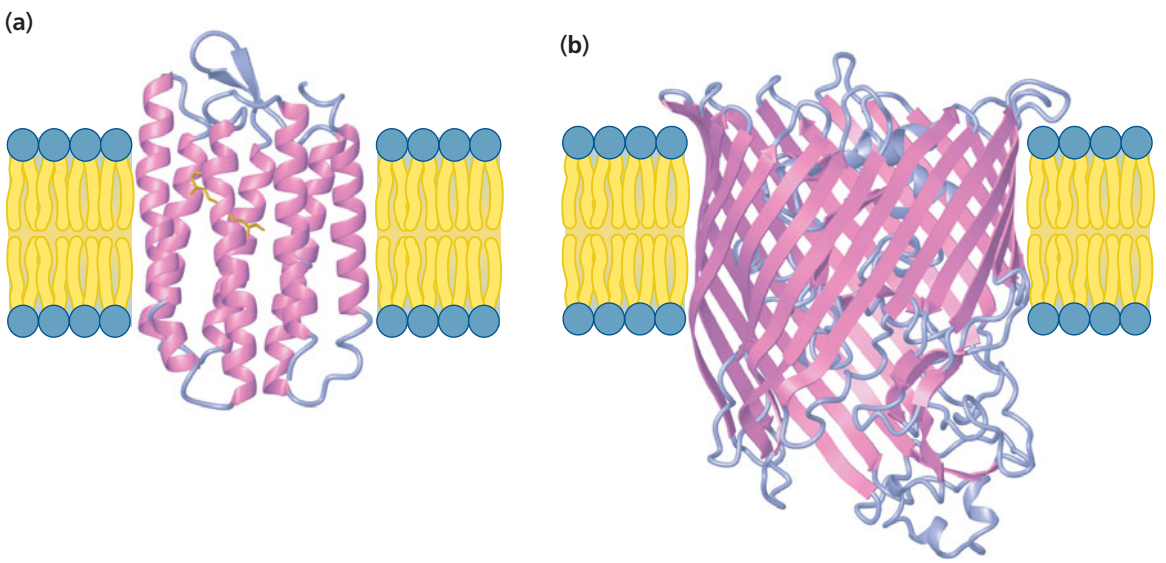


▲ **Figure 9.21**
Structure of a typical eukaryotic plasma membrane. A lipid bilayer forms the basic matrix of biological membranes, and proteins (some of which are glycoproteins) are associated with it in various ways. The oligosaccharides of glycoproteins and glycolipids face the extracellular space.

proteins make up one of the two major classes of integral membrane proteins. The other class is the β-barrel proteins (see below).

In the absence of data on three-dimensional structure, the presence of transmembrane α-helical regions in membrane proteins can often be predicted by searching amino acid sequences for regions that are hydrophobic (i.e., that have high hydropathy values) (Section 3.2G) and a tendency to be present in α-helices (Section 4.4). Various prediction algorithms have been developed over the years and they are currently able to detect 70% of known transmembrane α-helices. These predictions are important because it is still very difficult to crystallize membrane proteins in order to determine their true structure.

▼ **Figure 9.22**
Integral membrane proteins. (a) Bacteriorhodopsin: seven membrane-spanning α helices, connected by loops, form a bundle that spans the bilayer. The light-harvesting prosthetic group is shown in yellow. [PDB 1FBB]. **(b)** Porin FhuA from *Escherichia coli*: this porin forms a channel for the passage of protein-bound iron into the bacterium. The channel is formed from 22 antiparallel β strands that form a β-barrel. [PDB 1BY3].



Downloaded From Vet-e-Books.com

Protein folding is another example of an entropically driven assembly reaction (Section 4.11A).

We consider the functions of some of these membrane proteins later in this chapter. We will also encounter membrane proteins in other chapters, including those on membrane-associated electron transport (Chapter 14), photosynthesis (Chapter 15), and protein synthesis (Chapter 22).

The function of bacteriorhodopsin is described in Section 15.2.

Some prenyl-decorated proteins will be encountered in the discussion of signal transduction (Section 9.12).

Many integral membrane proteins have a β barrel fold (Figure 4.23b). The exterior surface of the β strands contacts the membrane lipids and the center of the barrel often serves as a pore or channel for passing molecules from one side of the membrane to the other. The *E. coli* porin, FhuA, is a typical example of this type of integral membrane protein (Figure 9.22b).

Peripheral membrane proteins are associated with one face of the membrane through charge–charge interactions and hydrogen bonding with integral membrane proteins or with the polar head groups of membrane lipids. Peripheral membrane proteins are more readily dissociated from membranes by changes in pH or ionic strength.

Lipid anchored membrane proteins are tethered to a membrane through a covalent bond to a lipid anchor. In the simplest lipid anchored membrane proteins, an amino acid side chain is linked by an amide or ester bond to a fatty acyl group, often from myristate or palmitate. The fatty acid is inserted into the cytoplasmic leaflet of the bilayer, anchoring the protein to the membrane (Figure 9.23a). Proteins of this type are found in viruses and eukaryotic cells.

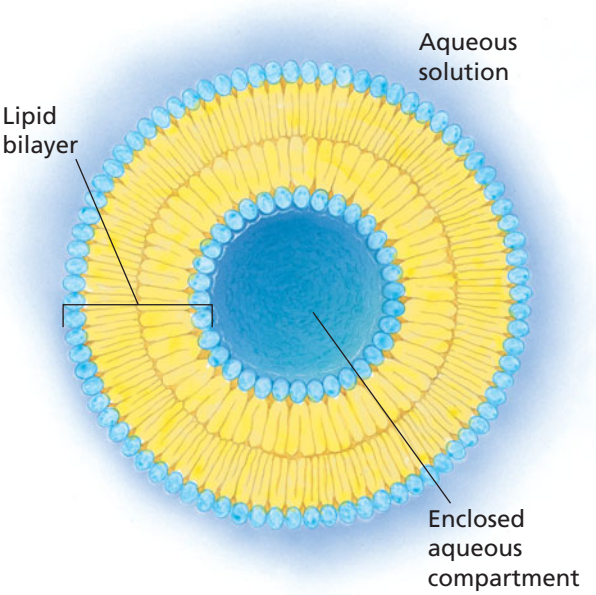
Other lipid anchored membrane proteins are covalently linked to an isoprenoid chain (either 15- or 20-carbon) through the sulfur atom of a cysteine residue at or near the C-terminus of the protein (Figure 9.23b). These *prenylated proteins* are found on the cytoplasmic face of both plasma membranes and intracellular membranes.

Many eukaryotic lipid anchored proteins are linked to a molecule of glycosylphosphatidylinositol (Figure 9.23c). The membrane anchor is the 1,2-diacylglycerol portion of the glycosylphosphatidylinositol. A glycan of varied composition is attached to the inositol by a glucosamine residue, a mannose residue links the glycan to a phosphoethanolamine residue, and the C-terminal α-carboxyl group of the protein is linked to the ethanolamine by an amide bond. Over 100 different proteins are known to be associated with membranes by a glycosylphosphatidylinositol anchor. These proteins have a variety of functions and they are present only in the outer monolayer of the plasma membrane. They are found in the cholesterol-sphingolipid rafts described in Section 9.9.

All three types of lipid anchors are covalently linked to amino acid residues post-translationally, that is, after the protein has been synthesized. Like integral membrane proteins, most lipid anchored proteins are permanently associated with the membrane, although the proteins themselves do not interact with the membrane. Once released by treatment with phospholipases, the proteins behave like soluble proteins.

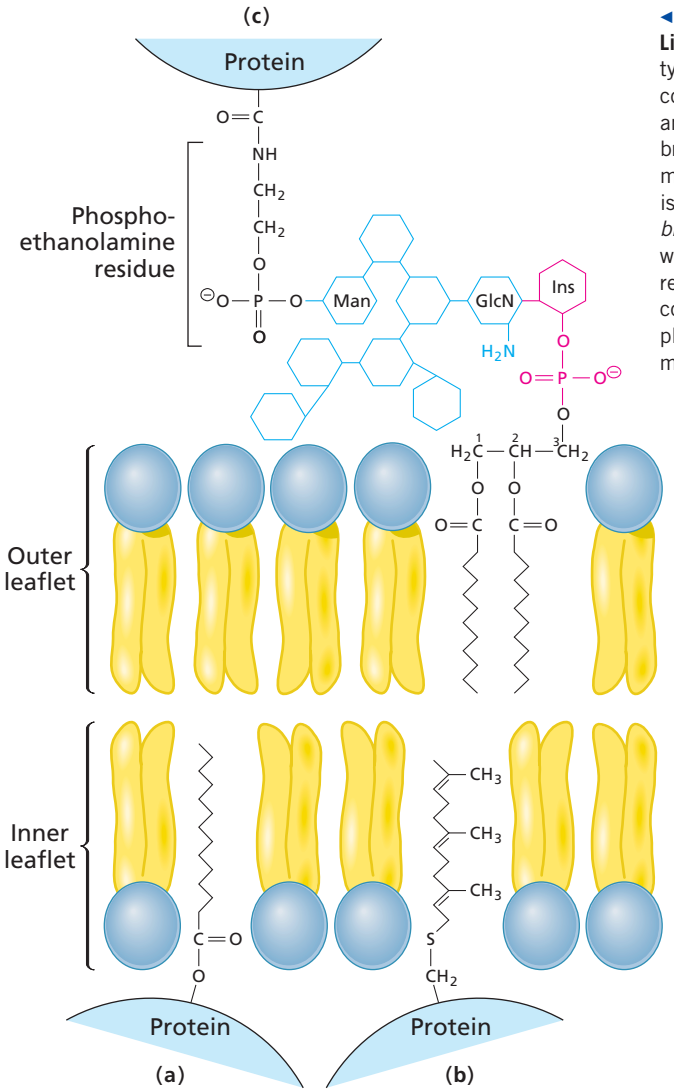
BOX 9.4 NEW LIPID VESICLES, OR LIPOSOMES

Synthetic vesicles (often called liposomes) consisting of phospholipid bilayers that enclose an aqueous compartment can be formed in the laboratory. In order to minimize unfavorable contact between the hydrophobic edge of the bilayer and the aqueous solution, lipid bilayers tend to close up to form these spherical structures. The vesicles are generally quite stable and impermeable to many substances. Liposomes whose aqueous inner compartment contains drug molecules can be used to deliver drugs to particular tissues in the body, provided that specific targeting proteins are present in the liposome membrane. Synthetic bilayers are an important experimental tool in the investigation of cellular membranes. An example of such an experiment is described in Box. 15.3.



► **Schematic cross-section of a lipid vesicle, or liposome.** The bilayer is made up of two leaflets. In each leaflet, the polar head groups of the amphipathic lipids extend into the aqueous medium and the nonpolar hydrocarbon tails point inward and are in van der Waals contact with each other.

Downloaded From Vet-e-Books.com



◀ **Figure 9.23**
Lipid anchored membrane proteins attached to the plasma membrane. The three types of anchors can be found in the same membrane, but they do not form a complex as shown here. **(a)** A fatty acyl anchored protein. **(b)** A prenyl anchored membrane protein. Note that fatty acyl and prenyl anchored membrane proteins can also occur on the cytoplasmic (outer) leaflet of intracellular membranes. **(c)** Protein anchored by glycosylphosphatidylinositol. Shown here is the variant surface glycoprotein of the parasitic protozoan *Trypanosoma brucei*. The protein is covalently bound to a phosphoethanolamine residue, which in turn is bound to a glycan. The glycan (blue) includes a mannose residue to which the phosphoethanolamine residue is attached and a glucosamine residue that is attached to the phosphoinositol group (red) of phosphatidylinositol. Abbreviations: GlcN, glucosamine; Ins, inositol; Man, mannose.





The total number of membrane proteins in a typical cell isn't known for certain but they are likely to represent a significant fraction of the proteome. In *E. coli*, for example, there appear to be roughly 1000 membrane proteins of all types. Since the total number of proteins is about 4000 (Chapter 4), membrane proteins account for about 25% of the total. This fraction is probably higher in multicellular eukaryotes because there are many more membrane proteins involved in cell-cell interactions and intracellular signaling.

Different membranes have different proteins (and lipids). In some cases a cell or compartment is enclosed by a double membrane consisting of two separate lipid bilayers (Figure 9.24). In the case of mitochondria and *E. coli*, the inner membranes have many more membrane proteins than the outer membranes.

Figure 9.24 ▶
Double membrane of mitochondria and many bacteria. The plasma membrane of most eukaryotic cells is a single lipid bilayer. Within eukaryotic cells the nucleus and major organelles such as mitochondria (top right) are bounded by double membranes. In bacteria, the gram-negative bacteria have a double membrane consisting of an inner and outer lipid bilayer as shown for *E. coli* (bottom right). It's not surprising that mitochondria (and chloroplasts) have a double membrane since they are derived from gram-negative bacteria that use the double membrane as part of the energy-producing mechanism of electron transport and ATP synthesis (Chapter 14).

Downloaded From Vet-e-Books.com

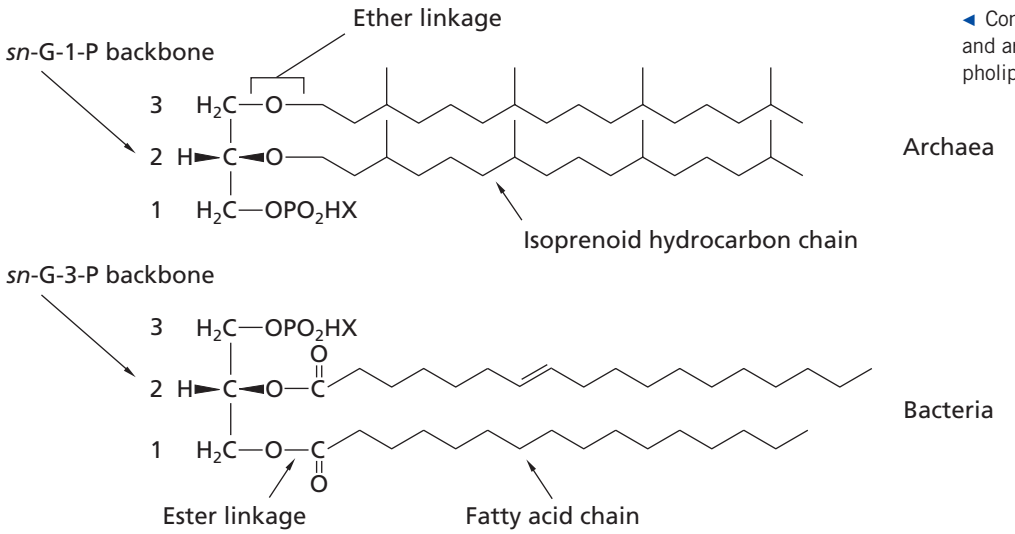
BOX 9.5 SOME SPECIES HAVE UNUSUAL LIPIDS IN THEIR MEMBRANES

Many species have unusual lipids in some of their membranes. The unusual lipids are sometimes confined to genera or families and sometimes entire orders share some distinctive lipid compositions. Within the eukaryotes, there are some lipids found only in some classes of animals and not others or in some classes of plants and not others. There are even distinctive lipid compositions in some entire kingdoms such as plants, animals, or fungi.

Prokaryotes are a very diverse group with many varieties of lipids. Major groups such as cyanobacteria, mycoplasma, and gram positive bacteria, can have quite characteristic lipid compositions in their membranes.

The archaebacteria (or Archaea) have glycerophospholipids that are quite unusual and distinctive. The glycerol phosphate backbone in archaebacterial glycerophospholipids is *sn*-glycerol-1-phosphate, a stereoisomer of the one found in other species (*sn*-glycerol-3-phosphate). (see Box 16.1) The hydrocarbon chains are attached to the glycerol backbone via ether linkages, not ester linkages, and the hydrocarbon chains in archaebacteria are often isoprenoid derivatives, not fatty acid derivatives.

There are a few species of gram-negative bacteria that have mixtures of ether and ester linkages in their lipids but unusual lipid composition of archaebacteria argues strongly in favor of classifying them as a distinctive monophyletic group. As mentioned earlier (Section 1.5), some scientists argue that the distinctiveness of archaebacteria justifies creating a third domain of life but the current view favors a more complex web of life perspective.



◄ Comparison of typical bacterial and archaebacterial glycero phospholipids.

C. The Fluid Mosaic Model of Biological Membranes

A typical biological membrane contains about 25% to 50% lipid and 50% to 75% protein by mass. Carbohydrates are present as components of glycolipids and glycoproteins. The lipids are a complex mixture of phospholipids, glycosphingolipids (in animals), and cholesterol (in some eukaryotes). Cholesterol and some other lipids that do not form bilayers by themselves (about 30% of the total) are stabilized in a bilayer arrangement by the other 70% of lipids in the membrane (see next section).

The compositions of biological membranes vary considerably among species and even among different cell types in multicellular organisms. For example, the myelin membrane that insulates nerve fibers contains relatively little protein. In contrast, the inner mitochondrial membrane is rich in proteins reflecting its high level of metabolic activity. The plasma membrane of red blood cells is also exceptionally rich in proteins.

Each biological membrane has a characteristic lipid composition, in addition to having a characteristic lipid to protein ratio. Membranes in brain tissue, for example, have a relatively high content of phosphatidylserines whereas membranes in heart and lung cells have high levels of phosphatidylglycerols and sphingomyelins, respectively. Phosphatidylethanolamines constitute nearly 70% of the inner membrane lipids of *E. coli* cells. The outer membranes of gram-negative bacteria contain lipopolysaccharides.

Downloaded From Vet-e-Books.com

In addition to being distributed differentially among different tissues, phospholipids are also distributed asymmetrically between the inner and outer monolayers of a single biological membrane. In mammalian cells, for example, 90% of the sphingomyelin molecules are in the outer surface of the plasma membrane. Phosphatidylserines are also asymmetrically distributed in many cells, with 90% of the molecules in the cytoplasmic monolayer.

A biological membrane is thicker than a lipid bilayer—typically 6 to 10 nm thick. The **fluid mosaic model** proposed in 1972 by S. Jonathan Singer and Garth L. Nicolson is still generally valid for describing the arrangement of lipid and protein within a membrane. According to the fluid mosaic model, the membrane is a dynamic structure in which both proteins and lipids can rapidly and randomly diffuse laterally or rotate within the bilayer. Membrane proteins are visualized as icebergs floating in a highly fluid lipid bilayer sea (Figure 9.21). (Actually, some proteins are immobile and some lipids have restricted movement.)

KEY CONCEPT

Membranes consist of a lipid bilayer and embedded proteins. Lipids and proteins can diffuse rapidly within the membrane.

9.9 Membranes Are Dynamic Structures

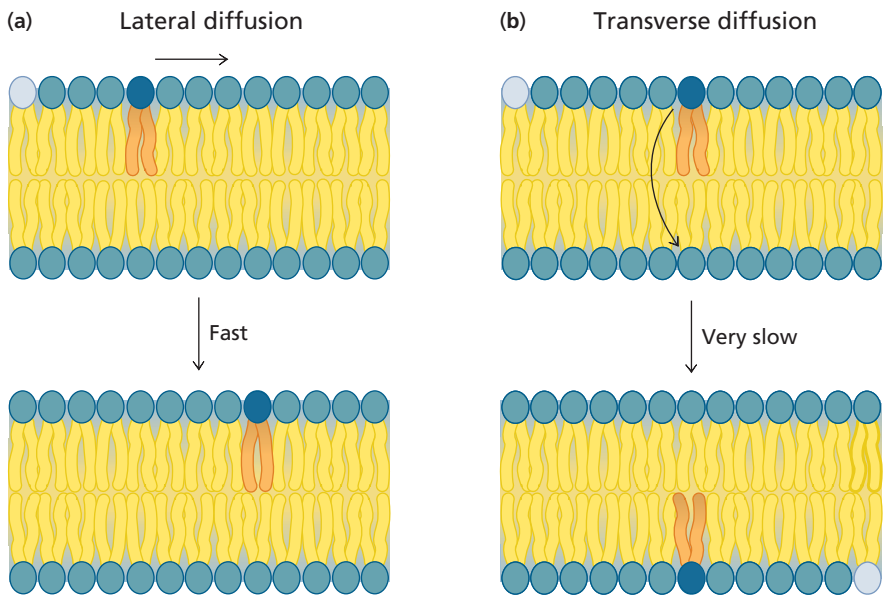
The lipids in a bilayer are in constant motion giving lipid bilayers many of the properties of fluids. A lipid bilayer can therefore be regarded as a two-dimensional solution. Lipids undergo several types of molecular motion within bilayers. The rapid movement of lipids within the plane of one monolayer is an example of two-dimensional lateral diffusion. A phospholipid molecule can diffuse from one end of a bacterial cell to the other (a distance of about 2 μm) in about 1 second at 37°C.

In contrast, transverse diffusion (or flip-flop) is the passage of lipids from one monolayer of the bilayer to the other. Transverse diffusion is much slower than lateral diffusion (Figure 9.25). The polar head of a phospholipid molecule is highly solvated and must shed its solvation sphere and penetrate the hydrocarbon interior of the bilayer in order to move from one leaflet to the other. The energy barrier associated with this movement is so high that transverse diffusion of phospholipids in a bilayer occurs at about one-billionth the rate of lateral diffusion. The very slow rate of transverse diffusion of membrane lipids is what allows the inner and outer layers of biological membranes to maintain different lipid compositions.

All cells synthesize new membrane by adding lipids and protein to preexisting membranes. As the plasma membrane is extended, the cell increases in size. Eventually the cell will divide and each daughter cell will inherit a portion (usually half) of the parental membranes. Internal membranes are extended and divide in the same manner.

You might have inherited lipid molecules from your grandmother! (see Problem 18).

In bacteria, lipid molecules are usually added to the cytoplasmic side of the lipid bilayer. Lipid asymmetry is generated by preferentially adding newly synthesized lipids to



◀ **Figure 9.25**
Diffusion of lipids within a bilayer. (a) Lateral diffusion of lipids is relatively rapid. **(b)** Transverse diffusion, or flip-flop, of lipids is very slow.

Downloaded From Vet-e-Books.com

only one of the monolayers. Since transverse diffusion is so slow, these newly synthesized molecules will not spread to the outer layer of the plasma membrane. This accounts for the enrichment of some types of lipids in the inner layer. Lipid asymmetry can also be generated and maintained by the activity of membrane-bound flipases and flopases—enzymes that use the energy of ATP to move specific phospholipids from one monolayer to the other. The activity of these enzymes accounts for the enrichment of certain types of phospholipid in the outer layer. Eukaryotic cells make their membrane lipids in an asymmetric arrangement in the endoplasmic reticulum or the Golgi apparatus. The membrane fragments flow from these organelles—retaining the asymmetry—to other membranes.

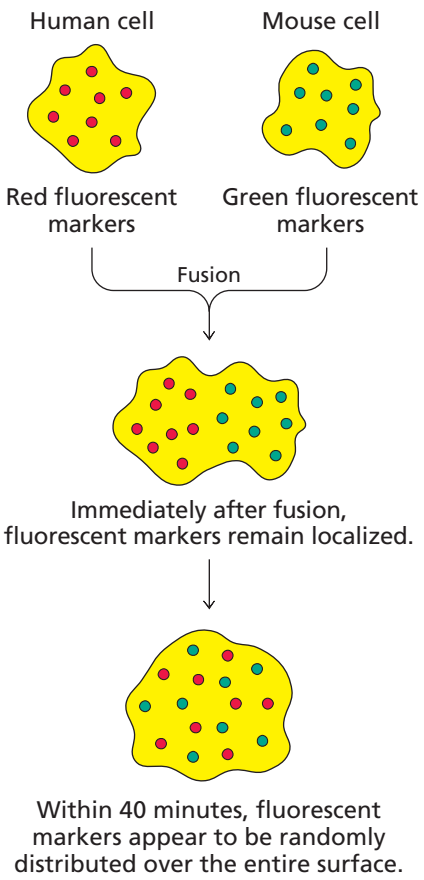
In 1970, L. D. Frye and Michael A. Edidin devised an elegant experiment to test whether membrane proteins diffuse within the lipid bilayer. Frye and Edidin fused mouse cells with human cells to form heterokaryons (hybrid cells). By using red fluorescence-labeled antibodies that specifically bind to certain proteins in human plasma membranes and green fluorescence-labeled antibodies that specifically bind to certain proteins in mouse plasma membranes, they observed the changes in the distribution of membrane proteins over time by immunofluorescence microscopy. The labeled proteins were intermixed within 40 minutes after cell fusion (Figure 9.26). This experiment demonstrated that at least some membrane proteins diffuse freely within biological membranes.

A few membrane proteins move laterally very rapidly but the majority of membrane proteins diffuse about 100 to 500 times more slowly than membrane lipids. The diffusion of some proteins is severely restricted by aggregation or by attachment to the cytoskeleton just beneath the membrane surface. Relatively immobile membrane proteins may act as fences or cages, restricting the movement of other proteins. The limited diffusion of membrane proteins produces protein patches, or domains—areas of membrane whose composition differs from that of the surrounding membrane.

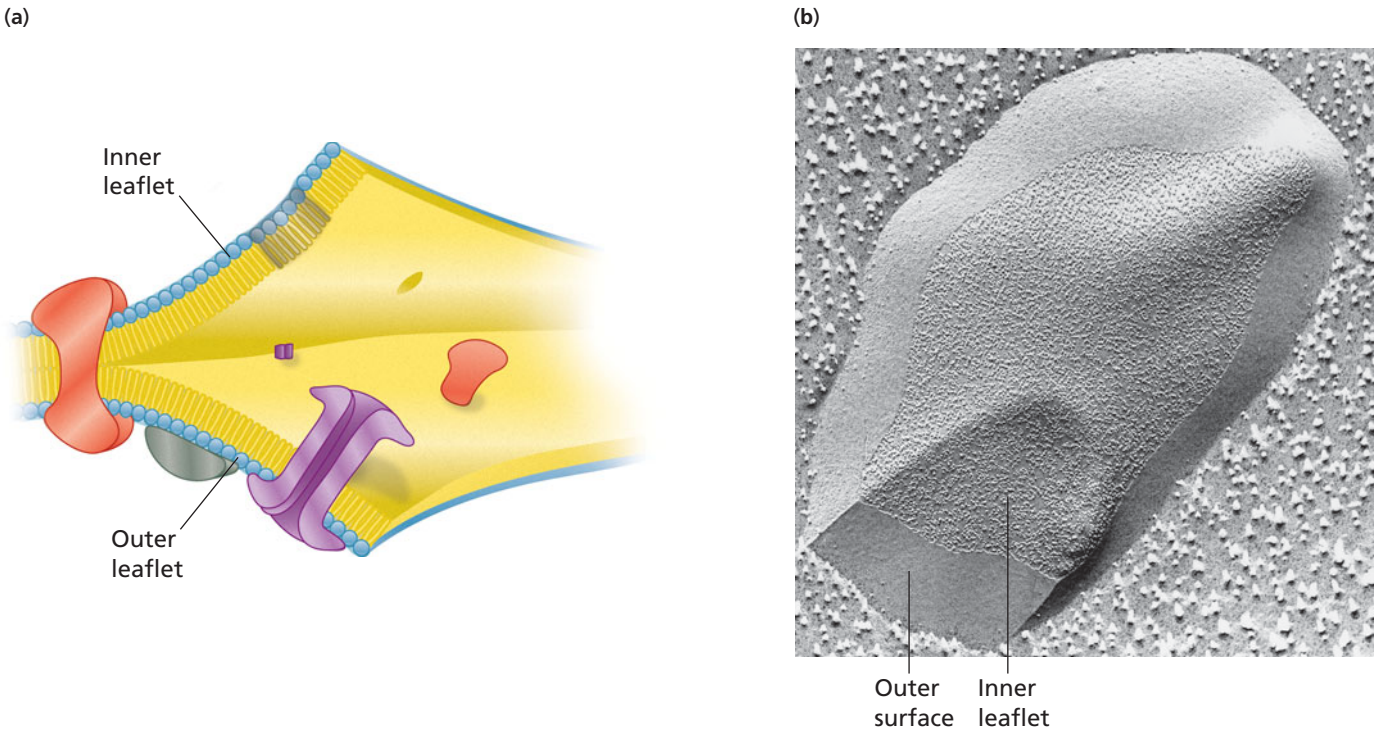
The distribution of membrane proteins can be visualized by *freeze-fracture electron microscopy*. In this technique, a membrane sample is rapidly frozen to the temperature of liquid nitrogen and then fractured with a knife. The membrane splits between the leaflets of the lipid bilayer where the intermolecular interactions are weakest (Figure 9.27a). Ice is evaporated in a vacuum and the exposed internal surface of the membrane is then coated with a thin film of platinum to make a metal replica for examination in an electron microscope. Membranes that are rich in membrane proteins contain pits and bumps indicating the presence of proteins. In contrast, membranes that contain no proteins are smooth. Figure 9.27b shows the bumpy surface of the inner monolayer of a red blood cell membrane exposed by removal of the outer layer.

The fluid properties of lipid bilayers depend on the flexibility of their fatty acyl chains. Saturated acyl chains are fully extended at low temperatures forming a crystalline array with maximal van der Waals contact between the chains. When the lipid bilayer is heated, a phase transition analogous to the melting of a crystalline solid occurs. The acyl chains of lipids in the resulting liquid crystalline phase are relatively disordered and loosely packed. During the phase transition, the thickness of the bilayer decreases by about 15% as the hydrocarbon tails become less extended because of rotation around C—C bonds (Figure 9.28). Bilayers composed of a single type of lipid undergo phase transition at a distinct temperature called the phase-transition temperature. When the lipids contain unsaturated acyl chains, the hydrophobic core of the bilayer is fluid well below room temperature (23°C). Biological membranes, which contain a heterogeneous mixture of lipids, change gradually from the gel to the liquid crystalline phase, typically over a temperature range of 10° to 40°C. Phase transitions in biological membranes can be localized so fluid- and gel-phase regions can coexist at certain temperatures.

The structure of a phospholipid has dramatic effects on its fluidity and phase-transition temperature. As we saw in Section 9.2, the hydrocarbon chain of a fatty acid with a *cis* double bond has a kink that disrupts packing and increases fluidity. Incorporating an unsaturated fatty acyl group into a phospholipid lowers the phase-transition temperature. Changes in membrane fluidity affect the membrane transport and catalytic functions of membrane proteins so many organisms maintain membrane fluidity under different conditions by adjusting the ratio of unsaturated to saturated fatty acyl groups in membrane



▲ **Figure 9.26**
Diffusion of membrane proteins. Human cells whose membrane proteins had been labeled with a red fluorescent marker were fused with mouse cells whose membrane proteins had been labeled with a green fluorescent marker. The initially localized markers became dispersed over the entire surface of the fused cell within 40 minutes.

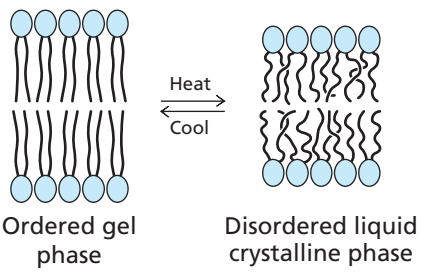


▲ **Figure 9.27**
Freeze fracturing a biological membrane.
(a) Splitting the lipid bilayer along the interface of the two leaflets. A platinum replica of the exposed internal surface is examined in an electron microscope. Membrane proteins appear as protrusions or cavities in the replica. **(b)** Electron micrograph of a freeze-fractured erythrocyte membrane. The bumps on the inner membrane surface show the locations of membrane proteins.

lipids. For example, when bacteria are grown at low temperatures, the proportion of unsaturated fatty acyl groups in membranes increases. Goldfish adapt to the temperature of the water in which they swim: as the environmental temperature drops, there is a rise in unsaturated fatty acids in goldfish intestinal membranes and whole brain. The lower melting point and greater fluidity of unsaturated fatty acyl groups preserve membrane fluidity allowing membrane processes to continue at colder temperatures.

Cholesterol accounts for 20% to 25% of the mass of lipids in a typical mammalian plasma membrane and significantly affects membrane fluidity. When the rigid cholesterol molecules intercalate between the hydrocarbon chains of the membrane lipids, the mobility of fatty acyl chains in the membrane is restricted and fluidity decreases at high temperatures (Figure 9.29). Cholesterol disrupts the ordered packing of the extended fatty acyl chains and thereby increases fluidity at low temperatures. Cholesterol in animal cell membranes thus helps maintain fairly constant fluidity despite fluctuations in temperature or degree of fatty acid saturation.

Cholesterol tends to associate with sphingolipids because they have long saturated fatty acid chains. The unsaturated chains of most glycerophospholipids produce kinks that don't easily accommodate cholesterol molecules in the membrane. Because of this preferential association, mammalian membranes consist of patches of cholesterol/sphingolipids regions surrounded by regions that have very little cholesterol. These patches are called **lipid rafts**. Certain membrane proteins may preferentially associate with lipid rafts. Thus, some membrane proteins may also have a patch-like distribution on the cell surface. Membrane proteins are thought to play an important role in maintaining the integrity of lipid rafts.



▲ **Figure 9.28**
Phase transition of a lipid bilayer. In the ordered gel state, the hydrocarbon chains are extended. Above the phase-transition temperature, rotation around C—C bonds disorders the chains in the liquid crystalline phase.

9.10 Membrane Transport

Plasma membranes physically separate a living cell from its environment. In addition, within both prokaryotic and eukaryotic cells there are membrane-bound compartments. The nucleus and mitochondria are obvious examples in eukaryotes.

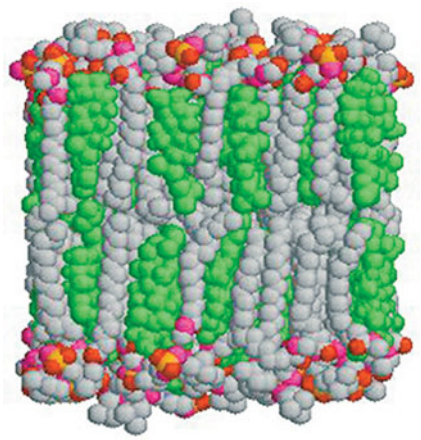
(a)



(b)



▲ **Goldfish adapt to water temperature. (a)** These goldfish (carp, *Carassius auratus*) have adapted to the water temperature in Kyoto, Japan, by adjusting the lipid composition of their membranes. **(b)** These Goldfish® do not adapt well to any water temperature.



▲ **Figure 9.29**
Model of a lipid membrane. Cholesterol molecules (green) are packed between phospholipid fatty acid chains (grey).

Membranes are selectively permeable barriers that restrict the free passage of most molecules. As a general rule, the permeability of molecules is related to their hydrophobicity and their tendency to dissolve in organic solvents. Thus, hexanoic acid, acetic acid, and ethanol are able to move across membranes quite readily. They have high permeability coefficients (Figure 9.30). Water, despite its strong polar character, is able to diffuse freely across lipid bilayers although, as the permeability coefficient indicates, its movement is still greatly restricted compared to organic solvents like hexanoic acid.

Small ions like Na^+, K^+, and Cl^- have very low permeability coefficients. They are unable to diffuse across a membrane because the hydrophobic core of the lipid bilayer presents an almost impenetrable barrier to most polar or charged species. H^+ ions have a much higher permeability coefficient although membranes still act as an effective barrier to protons.

As mentioned above, very hydrophobic molecules and some small uncharged molecules can move through biological membranes. Water, oxygen, and other small molecules must also be able to enter all cells and move freely between compartments inside eukaryotic cells even if they are not able to diffuse as quickly across membranes. Larger molecules, such as proteins and nucleic acids, have to be transported across membranes, including the membranes between compartments. Living cells move molecules across membranes using transport proteins (sometimes called pores, carriers, permeases, or pumps) and they transport macromolecules by endocytosis or exocytosis.

Nonpolar gases, such as O_2 and CO_2, and hydrophobic molecules, such as steroid hormones, lipid vitamins, and some drugs, enter and leave the cell by diffusing through the membrane moving from the side with the higher concentration to the side with the lower concentration. The rate of movement depends on the difference in concentrations, or the concentration gradient, between the two sides. Diffusion down a concentration gradient (i.e., downhill diffusion) is a spontaneous process driven by an increase in entropy and therefore a decrease in free energy (see below).

The traffic of other molecules and ions across membranes is mediated by three types of integral membrane proteins: channels and pores, passive transporters, and active transporters. These transport systems differ in their kinetic properties and energy requirements. For example, the rate of solute movement through pores and channels may increase with increasing solute concentration but the rate of movement through passive and active transporters may approach a maximum as the solute concentration increases (i.e., the transport protein becomes saturated). Some types of transport require a source of energy (Section C). The characteristics of membrane transport are summarized in Table 9.3. In this section, we describe the different membrane transport systems, as well as endocytosis and exocytosis.

A. Thermodynamics of Membrane Transport

Recall from Chapter 1 (Section 1.4C) that the actual Gibbs free energy change of a reaction is related to the standard Gibbs free energy change by the equation

$$\Delta G_{\text{reaction}} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln \frac{[C][D]}{[A][B]} \qquad \textbf{(9.1)}$$

where $\Delta G^{\circ\prime}_{\text{reaction}}$ represents the standard Gibbs free energy change for the reaction, [C][D] represents the concentrations of the products, and [A][B] represents the concentration of the reactants. The Gibbs free energy change associated with membrane transport depends only on the concentrations of the molecules on either side of the membrane.

For any molecule, A, the concentration on the inside of the membrane is $[A_{in}]$ and the concentration outside is $[A_{out}]$. The Gibbs free energy change associated with transporting molecules of A is

$$\Delta G_{\text{transport}} = RT \ln \frac{[A_{in}]}{[A_{out}]} = 2.303\ RT \log \frac{[A_{in}]}{[A_{out}]} \qquad \textbf{(9.2)}$$

Table 9.3 Characteristics of different types of membrane transport

	Protein carrier	Saturable with substrate	Movement relative to concentration gradient	Energy input required
Simple diffusion	No	No	Down	No
Channels and pores	Yes	No	Down	No
Passive transport	Yes	Yes	Down	No
Active transport				
Primary	Yes	Yes	Up	Yes (direct source)
Secondary	Yes	Yes	Up	Yes (ion gradient)

If the concentration of A inside the cell is much less than the concentration of A outside the cell then $\Delta G_{transport}$ will be negative and the flow of A into the cell will be thermodynamically favored. For exmple, if $[A_{in}] = 1$ mM and $[A_{out}] = 100$ mM, then at 25°C

$$\Delta G_{transport} = 2.303\ RT \log \frac{[A_{in}]}{[A_{out}]} = 2.303 \times 8.325 \times 298 \times (-2) = -11.4\ \text{kJ mol}^{-1} \quad (9.3)$$

Under these conditions, molecules of solute A will tend to flow into the cell in order to reduce the concentration gradient. Flow in the opposite direction is thermodynamically unfavorable since it is associated with a positive Gibbs free energy change ($\Delta G_{transport}$ = +11.4 kJ mol^{-1} for molecules moving from the inside of the cell to the outside).

Equation 9.2 only applies to uncharged molecules. In the case of ions, the Gibbs free energy change has to include a factor that takes into account the charge difference across a biological membrane. Most cells selectively export cations so the inside of a cell is negatively charged with respect to the outside. The charge difference across the membrane is

$$\Delta\Psi = \Psi_{in} - \Psi_{out} \quad (9.4)$$

where $\Delta\Psi$ is called the membrane potential (in volts). The Gibbs free energy change due to this electric potential is

$$\Delta G = zF\Delta\Psi \quad (9.5)$$

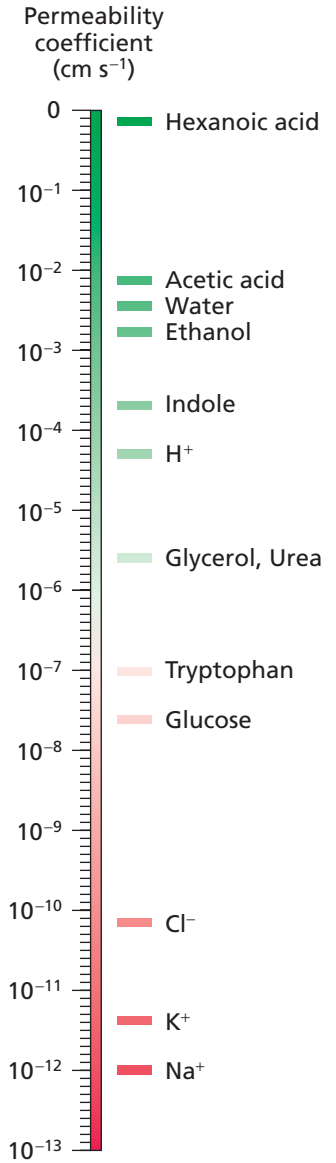
where z is the charge on the molecule being transported (e.g., +1, −1, +2, −2, etc.) and F is Faradays's constant (96,485 JV^{-1}mol^{-1}). Since the inside of the cell is negatively charged, the import of cations such as $Na^{\oplus}$ and $K^{\oplus}$ is thermodynamically favored by the membrane potential. The export of cations must be coupled to an energy-producing reaction since it is associated with a positive Gibbs free energy change.

Both the chemical (concentration) and electric (charge) effects have to be considered, for any transport process involving charged molecules. Thus,

$$\Delta G_{transport} = 2.303\ RT \log \frac{[A_{in}]}{[A_{out}]} + zF\Delta\Psi \quad (9.6)$$

B. Pores and Channels

Pores and *channels* are transmembrane proteins with a central passage for ions and small molecules. (Usually, the term *pore* is used for bacteria and *channel* for animals.) Solutes of the appropriate size, charge, and molecular structure can move rapidly

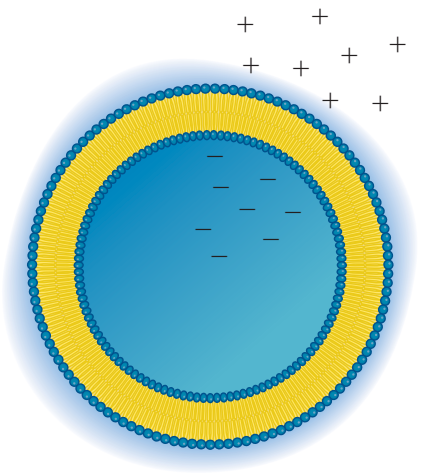


▲ **Figure 9.30**
Permeability coefficients of various molecules. Molecules with high permeability coefficients (top) are able to diffuse unaided across a membrane.

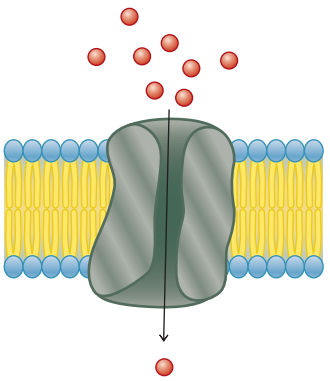
KEY CONCEPT

For a given solute, the Gibbs free energy change of transport depends on both the membrane potential and solute concentrations on either side of the membrane.

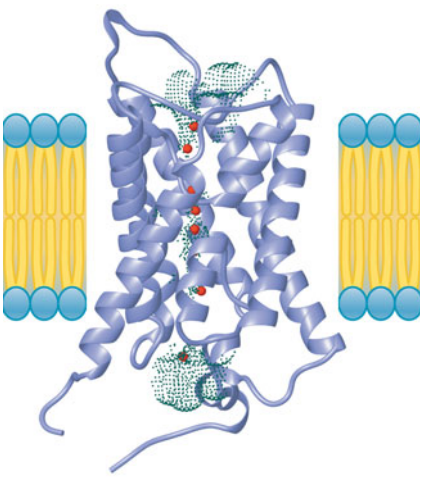
The importance of Equation 9.6 will become apparent when we describe chemiosmotic theory (Section 14.3).



▲ **Membrane potential.** In most cases the inside of a cell or membrane compartment is negative with respect to the outside and the membrane potential ($\Delta\psi$) is negative.



▲ **Figure 9.31**
Membrane transport through a pore or channel. A central passage allows molecules and ions of the appropriate size, charge, and geometry to traverse the membrane in either direction.



through the passage in either direction by diffusing down a concentration gradient (Figure 9.31). This process requires no energy. In general, the rate of movement of solute through a pore or channel is not saturable at high concentrations. For some channels, the rate may approach the diffusion controlled limit.

The outer membranes of some bacteria are rich in porins, a family of pore proteins that allow ions and many small molecules to gain access to specific transporters in the plasma membrane. Similar channels are found in the outer membranes of mitochondria. Porins are usually only weakly solute-selective. They can act as sieves that are permanently open or they can be regulated by the concentration of solutes. In contrast, plasma membranes also contain many channel proteins that are highly specific for certain ions and they open or close in response to a specific signal.

Aquaporin is an integral membrane protein that acts as a pore for water molecules. The channel through the middle of the protein will allow for passage of water molecules and other small uncharged molecules but it blocks passage of any charged molecules or large molecules. This channel is larger on the outside surface but narrows to a much smaller channel on the cytoplasmic side as shown for yeast aquaporin in Figure 9.32. Aquaporins are common in all species. They are required in cells where the rapid uptake of water is necessary because the rate of diffusion of water across the membrane is too slow. This is an example of a simple, somewhat specific, porin. It was discovered by Peter Agre, who received the Nobel Prize in Chemistry in 2003.

CorA is the primary Mg^{2+} pump in prokaryotic cells. It is highly selective for Mg^{2+} and permits the import of Mg^{2+} against a concentration gradient in response to the membrane potential. Positively charged ions "want" to flow into cells and the CorA pore allows passage of Mg^{2+} but not other ions. Mg^{2+} is essential for many cell functions. The rate of influx is regulated by the large cytoplasmic domain of CorA (Figure 9.33). It binds Mg^{2+} ions and when a sufficient number have bound, the pore is closed. Thus, influx of Mg^{2+} is controlled by the cytoplasmic concentration.

Membranes of nerve tissues have gated (i.e., controlled) potassium channels that selectively allow rapid outward transport of potassium ions. These channels permit $K^{\oplus}$ ions to pass through the membrane at least 10,000 times faster than the smaller $Na^{\oplus}$ ions. Crystallographic studies have shown that the potassium channel has a wide mouth (like a funnel) containing negatively charged amino acids to attract cations and repel anions. Hydrated cations are directed electrostatically to an electrically neutral constriction of the pore called the selectivity filter. Potassium ions rapidly lose some of their water of hydration and pass through the selectivity filter. Sodium ions apparently retain more water of hydration and therefore transit the filter much more slowly. The remainder of the channel has a hydrophobic lining. Based on comparisons of amino acid sequences, the general structural properties of the potassium channel seem to also apply to other types of channels and pores. Roderick MacKinnon shared the 2003 Nobel Prize in Chemistry with Peter Agre. MacKinnnon's work focused mainly on potassium channels.

C. Passive Transport and Facilitated Diffusion

Pore and channel proteins are examples of **passive transport** where the Gibbs free energy change for transport is negative and transport from one side of the membrane to the other is a spontaneous process. In *active transport* (see below), the solute moves against a concentration gradient and/or a charge difference. Active transport must be coupled to an energy-producing reaction in order to overcome the unfavorable Gibbs free energy change for unassisted transport. The simplest membrane transporters—whether active or passive—carry out **uniport**; that is, they carry only a single type of solute across the membrane (Figure 9.34a). Many transporters carry out the simultaneous transport of two different solute molecules. The process is called **symport** if both solutes are

◀ **Figure 9.32**
Fungal aquaporin. Aquaporin is an integral membrane protein with an α-helix bundle domain. The water channel (green dots) is open on the exterior surface and narrows to a tiny passage on the cytoplasmic side. *[Pichia pastoris* PDB 2W2E]

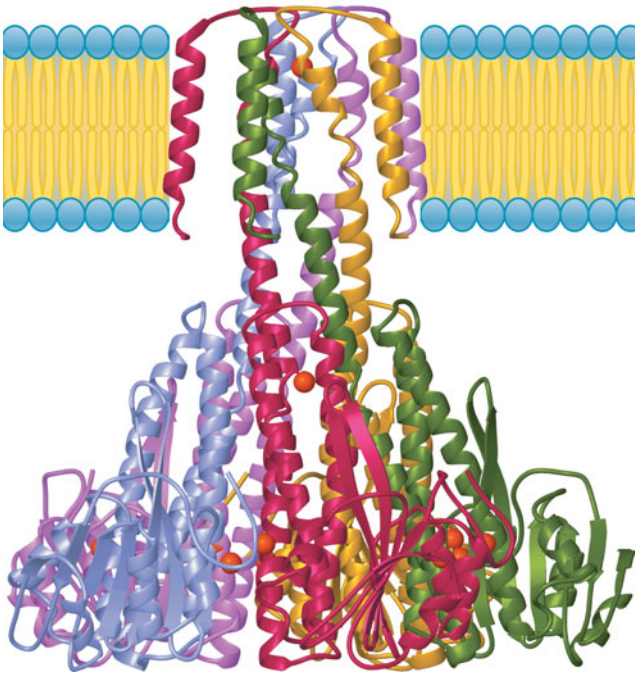


Figure 9.33 ▲
CorA, a magnesium pump. CorA is the prokaryotic magnesium pump. Mg^{2+} ions bind on the exterior surface and are transported through a highly selective channel in response to the membrane potential. The cytoplasmic domain binds Mg^{2+} ions and closes the pore in response to high internal concentrations of Mg^{2+}. This is the *Thermotoga maritima* version with each of the fire subunits in a different color. [PDB 2HN2]

transported in the same direction, (Figure 9.34b). If they are transported in opposite directions, the process is **antiport** (Figure 9.34c).

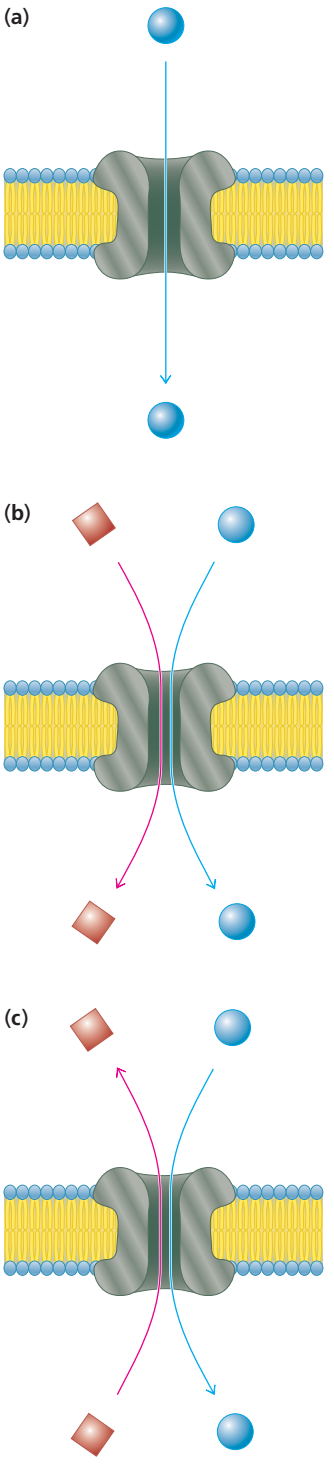
Passive transport includes simple diffusion across a membrane. When pores, channels, and transporters are involved, we call the process **facilitated diffusion**. Facilitated diffusion is still an example of passive transport since it does not require an energy source. The transport protein accelerates the movement of solute down its concentration gradient, or charge gradient, a process that would occur very slowly by diffusion alone. In this case, transport proteins are similar to enzymes because they increase the rate of a process that is thermodynamically favorable. For a simple passive uniport system, the initial rate of inward transport, like the initial rate of an enzyme-catalyzed reaction, depends on the external concentration of substrate. The equation describing this dependence is analogous to the Michaelis–Menten equation for enzyme catalysis (Equation 5.14).

$$v_0 = \frac{V_{max}[S]_{out}}{K_{tr} + [S]_{out}} \qquad (9.7)$$

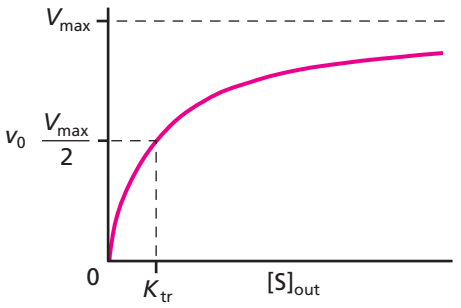
where v_0 is the initial rate of inward transport of the substrate at an external concentration $[S]_{out}$, V_{max} is the maximum rate of transport of the substrate, and K_{tr} is a constant analogous to the Michaelis constant (K_m) (i.e., K_{tr} is the substrate concentration at which the transporter is half-saturated). The lower the value of K_{tr}, the higher the affinity of the transporter for the substrate. The rate of transport is saturable, approaching a maximum value at a high substrate concentration (Figure 9.35).

As substrate accumulates inside the cell, the rate of outward transport increases until it equals the rate of inward transport, and $[S]_{in}$ equals $[S]_{out}$. At this point, there is no net change in the concentration of substrate on either side of the membrane, although substrate continues to move across the membrane in both directions.

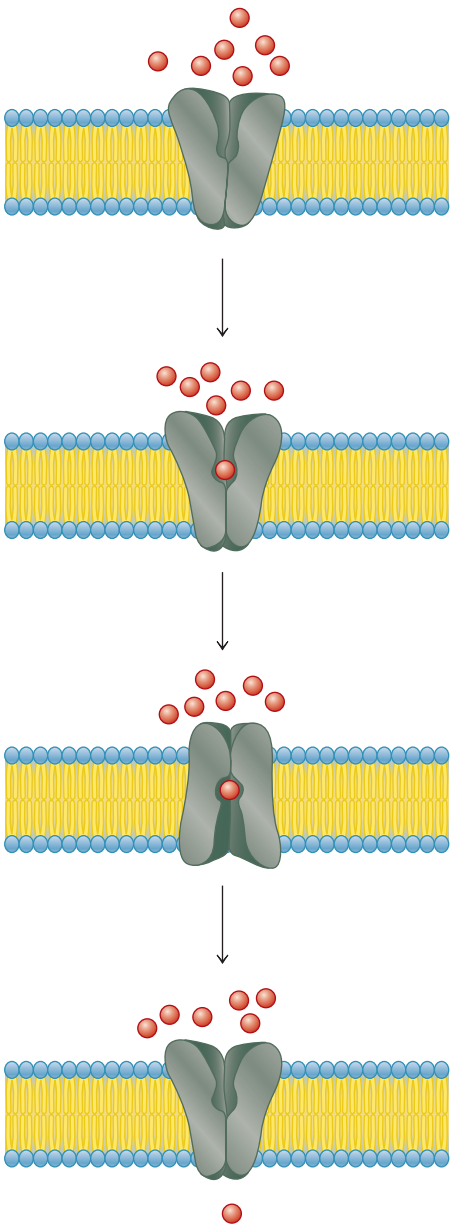
Models of transport protein operation suggest that some transporters undergo a conformational change after they bind their substrates. This conformational change allows the substrate to be released on the other side of the membrane; the transporter



▲ Figure 9.34
Types of passive and active transport. Although the transport proteins are depicted as having an open central pore, passive and active transporters actually undergo conformational changes when transporting their solutes. **(a)** Uniport. **(b)** Symport. **(c)** Antiport.



▲ **Figure 9.35**
Kinetics of passive transport. The initial rate of transport increases with substrate concentration until a maximum is reached. K_{tr} is the concentration of substrate at which the rate of transport is half-maximal.



◄ **Figure 9.36**
Passive and active transport protein function. The protein binds its specific substrate and then undergoes a conformational change, allowing the molecule or ion to be released on the other side of the membrane. Cotransporters have specific binding sites for each transported species.

then reverts to its original state (Figure 9.36). The conformational change in the transporter is often triggered by binding of the transported species, as in the induced fit of certain enzymes to their substrates (Section 6.9). In active transport, the conformational change can be driven by ATP or other sources of energy. Like enzymes, transport proteins can be susceptible to reversible and irreversible inhibition.

D. Active Transport

Active transport resembles passive transport in overall mechanism and kinetic properties. However, active transport requires energy to move a solute up its concentration gradient. In some cases, active transport of charged molecules or ions also results in a charge gradient across the membrane and active transport moves ions against the membrane potential.

Active transporters use a variety of energy sources, most commonly ATP. Ion-transporting ATPases are found in all organisms. These active transporters, which include $Na^{\oplus}$-$K^{\oplus}$ ATPase, and Ca^{2+} ATPase, create and maintain ion concentration gradients across the plasma membrane and across the membranes of internal organelles.

Primary active transport is powered by a direct source of energy such as ATP or light. For example, bacteriorhodopsin (Figure 9.22) uses light energy to generate a transmembrane proton concentration gradient that can be used for ATP formation. One primary active transport protein, P-glycoprotein, appears to play a major role in the resistance of tumor cells to multiple chemotherapeutic drugs. Multidrug resistance is a leading cause of failure in the clinical treatment of human cancers. P-Glycoprotein is an integral membrane glycoprotein (M_r 170,000) that is abundant in the plasma membrane of drug-resistant cells. Using ATP as an energy source, P-glycoprotein pumps a large variety of structurally unrelated nonpolar compounds, such as drugs, out of the cell up a concentration gradient. In this way, the cytosolic drug concentration is maintained at a level low enough to avoid cell death. The normal physiological function of P-glycoprotein appears to be removal of toxic hydrophobic compounds in the diet.

Secondary active transport is driven by an ion concentration gradient. The active uphill transport of one solute is coupled to the downhill transport of a second solute that was concentrated by primary active transport. For example, in *E. coli*, electron flow through a series of membrane-bound oxidation–reduction enzymes generates a higher extracellular concentration of protons. As protons flow back into the cell down their concentration gradient, lactose is also transported in, against its concentration gradient (Figure 9.37). The energy of the proton concentration gradient drives the secondary active transport of lactose. The symport of $H^{\oplus}$ and lactose is mediated by the transmembrane protein lactose permease.

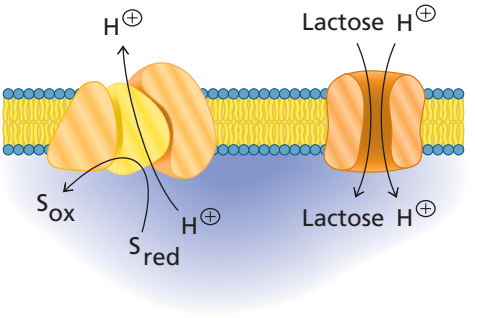
In large multicellular animals, secondary active transport is often powered by a sodium ion gradient. Most cells maintain an intracellular potassium ion concentration of about 140 mM in the presence of an extracellular concentration of about 5 mM. The cytosolic concentration of sodium ions is maintained at about 5 to 15 mM in the presence of an extracellular concentration of about 145 mM. These ion concentration gradients are maintained by $Na^{\oplus}$–$K^{\oplus}$ ATPase, an ATP-driven antiport system that pumps two $K^{\oplus}$ into the cell and ejects three $Na^{\oplus}$ for every molecule of ATP hydrolyzed (Figure 9.38). Each $Na^{\oplus}$–$K^{\oplus}$ ATPase can catalyze the hydrolysis of about 100 molecules of ATP per minute, a significant portion (up to one-third) of the total energy consumption of a typical animal cell. The $Na^{\oplus}$ gradient that is generated by $Na^{\oplus}$–$K^{\oplus}$ ATPase is the major source of energy for secondary active transport of glucose in intestinal cells. One glucose molecule is imported with each sodium ion that enters the cell. The energy released by the downhill movement of $Na^{\oplus}$ powers the uphill transport of glucose.

E. Endocytosis and Exocytosis

The transport we have discussed so far occurs by the flow of molecules or ions across an intact membrane. Cells also need to import and export molecules too large to be transported via pores, channels, or transport proteins. Prokaryotes possess specialized multicomponent export systems in their plasma and outer membranes that allow them to secrete certain proteins (often toxins or enzymes) into the extracellular medium. In eukaryotic cells, many—but not all—proteins (and certain other large substances) are moved into and out of the cell by **endocytosis** and **exocytosis**, respectively. In both cases, transport involves formation of a specialized type of lipid vesicle.

Endocytosis is the process by which macromolecules are engulfed by the plasma membrane and brought into the cell inside a lipid vesicle. Receptor-mediated endocytosis begins with the binding of macromolecules to specific receptor proteins in the plasma membrane of the cell. The membrane then invaginates, forming a vesicle that contains the bound molecules. As shown in Figure 9.39, the inside of such a membrane vesicle is equivalent to the outside of a cell; thus, substances inside the vesicle have not actually crossed the plasma membrane. Once inside the cell, the vesicle can fuse with an endosome (another type of vesicle) and then with a lysosome. Inside a lysosome, the endocytosed material and the receptor itself can be degraded. Alternatively, the ligand, the receptor, or both, can be recycled from the endosome back to the plasma membrane.

Exocytosis is similar to endocytosis except that the direction of transport is reversed. During exocytosis, materials destined for secretion from the cell are enclosed in vesicles by the Golgi apparatus (Section 1.8B). The vesicles then fuse with the plasma membrane releasing the vesicle contents into the extracellular space. The zymogens of digestive enzymes are exported from pancreatic cells in this manner (Section 6.7A).



▲ **Figure 9.37**
Secondary active transport in *Escherichia coli*. The oxidation of reduced substrates (S_{red}) generates a transmembrane proton concentration gradient. The energy released by protons moving down their concentration gradient drives the transport of lactose into the cell by lactose permease.

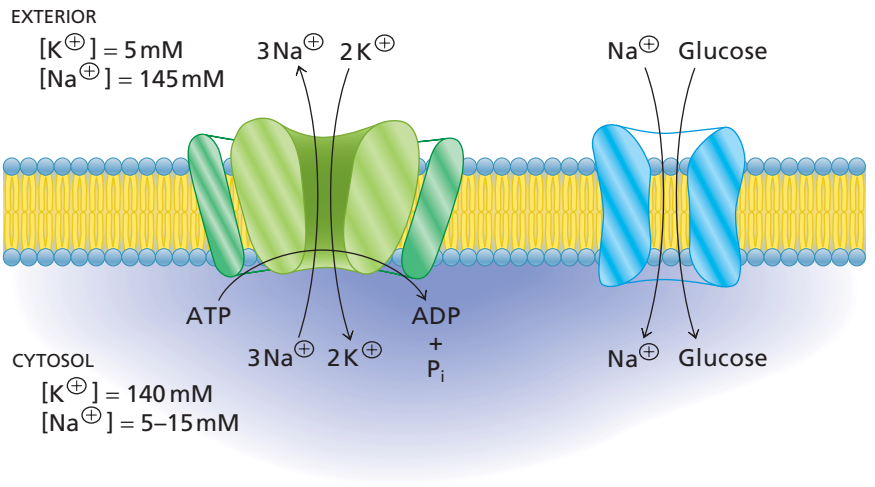
The secretory pathway in eukaryotic cells is described in Section 22.10.

9.11 Transduction of Extracellular Signals

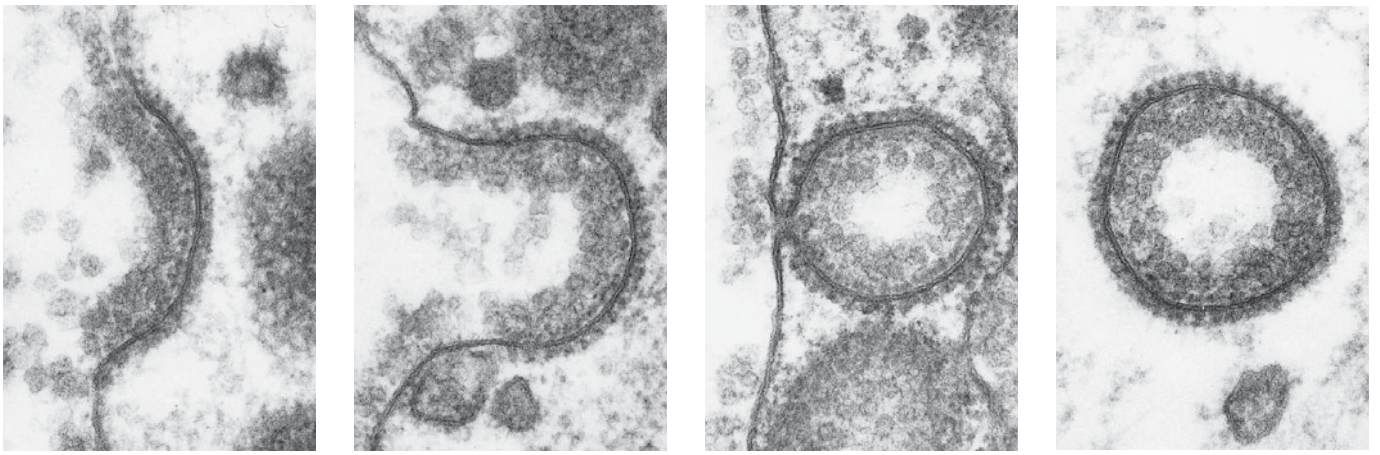
In order for a cell to interact with its external environment, it must detect molecules outside of the plasma membrane and convey that information to the inside of the cell. This process is called **signal transduction** and it is a very active field of research. In this section we'll cover the basic mechanism of the most common signaling pathways. As you learn more biochemistry, you'll encounter many variations of these themes.

A. Receptors

The plasma membranes of all cells contain specific receptors that allow the cell to respond to external chemical stimuli that cannot cross the membrane. For example,



◀ **Figure 9.38**
Secondary active transport in animals. The $Na^{\oplus}$–$K^{\oplus}$ ATPase generates a sodium ion gradient that drives secondary active transport of glucose in intestinal cells.



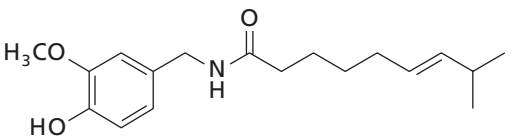
▲ **Figure 9.39**
Electron micrographs of endocytosis. Endocytosis begins with the binding of macromolecules to the plasma membrane of the cell. The membrane then invaginates forming a vesicle that contains the bound molecules. The inside of the vesicle is topologically equivalent to the outside of the cell.

bacteria can detect certain chemicals in their environment. A signal is passed via a cell surface receptor to the flagella, causing the bacterium to swim toward a potential food source. This is called positive **chemotaxis**. In negative chemotaxis, the bacteria swim away from toxic chemicals.

In multicellular organisms, stimuli such as *hormones*, *neurotransmitters* (substances that transmit nerve messages at synapses), and *growth factors* (proteins that regulate cell proliferation) are produced by specialized cells. These ligands can travel to other tissues where they bind to and produce specific responses in cells with the appropriate receptors on their surfaces. In this section, we see how the binding of water-soluble ligands to receptors elicits intracellular responses in mammals. These signal transduction pathways involve adenylyl cyclase, inositol phospholipids, and receptor tyrosine kinases.

BOX 9.6 THE HOT SPICE OF CHILI PEPPERS

Biochemists now know the mechanism by which spice from "hot" peppers exerts its action, causing a burning pain. The active factor in capsaicin peppers is a lipophilic vanilloid compound called capsaicin.

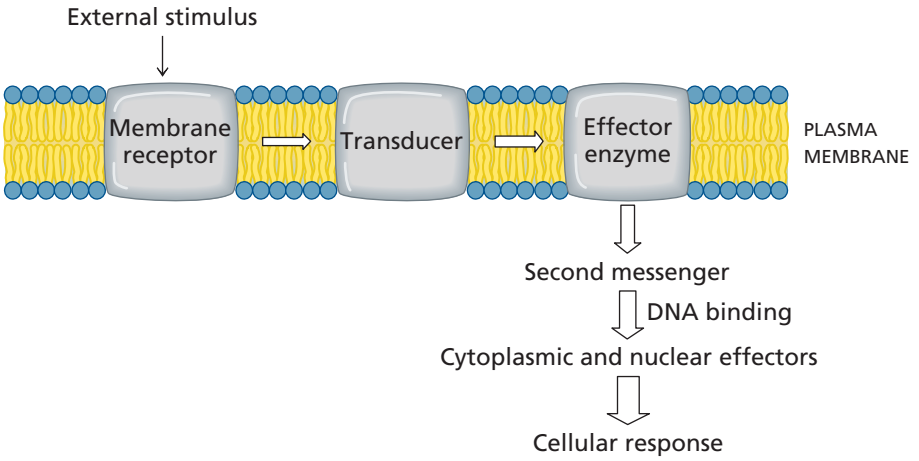


Capsaicin

A nerve cell protein receptor that responds to capsaicin has been identified and characterized. It is an ion channel and its amino acid sequence suggests that it has six transmembrane domains. Activation of the receptor by capsaicin causes the channel to open so that calcium and sodium ions can flow into the nerve cell and send an impulse to the brain. The receptor is activated not only by vanilloid spices but also by rapid increases in temperature. In fact, the main function of the receptor is detection of heat.



Chili peppers ►



◀ **Figure 9.40**
General mechanism of signal transduction across the plasma membrane of a cell.

A general mechanism for signal transduction is shown in Figure 9.40. A ligand binds to its specific receptor on the surface of the target cell. This interaction generates a signal that is passed through a membrane protein **transducer** to a membrane-bound **effector enzyme**. The action of the effector enzyme generates an intracellular **second messenger** that is usually a small molecule or ion. The diffusible second messenger carries the signal to its ultimate destination which may be in the nucleus, an intracellular compartment, or the cytosol. Ligand binding to a cell-surface receptor almost invariably results in the activation of protein kinases. These enzymes catalyze the transfer of a phosphoryl group from ATP to various protein substrates, many of which help regulate metabolism, cell growth, and cell division. Some proteins are activated by phosphorylation, whereas others are inactivated. A vast diversity of ligands, receptors, and transducers exists but only a few second messengers and types of effector enzymes are known.

Kinases were introduced in Section 6.9.

Receptor tyrosine kinases have a simpler mechanism for signal transduction. With these enzymes, the membrane receptor, transducer, and effector enzyme are combined in one enzyme. A receptor domain on the extracellular side of the membrane is connected to the cytosolic active site by a transmembrane segment. The active site catalyzes phosphorylation of its target proteins.

KEY CONCEPT

Membrane receptors are the primary step in carrying information across a membrane.

Amplification is an important feature of signaling pathways. A single ligand receptor complex can interact with a number of transducer molecules, each of which can activate several molecules of effector enzyme. Similarly, the production of many second messenger molecules can activate many kinase molecules that catalyze the phosphorylation of many target proteins. This series of amplification events is called a **cascade**. The cascade mechanism means that small amounts of an extracellular compound can affect large numbers of intracellular enzymes without crossing the plasma membrane or binding to each target protein.

The actions of the hormones insulin, glucagon, and epinephrine and the roles of transmembrane signaling pathways in the regulation of carbohydrate and lipid metabolism are described in Sections 11.5, 13.3, 13.7, 13.10, 16.1C, 16.4 (Box), and 16.7.

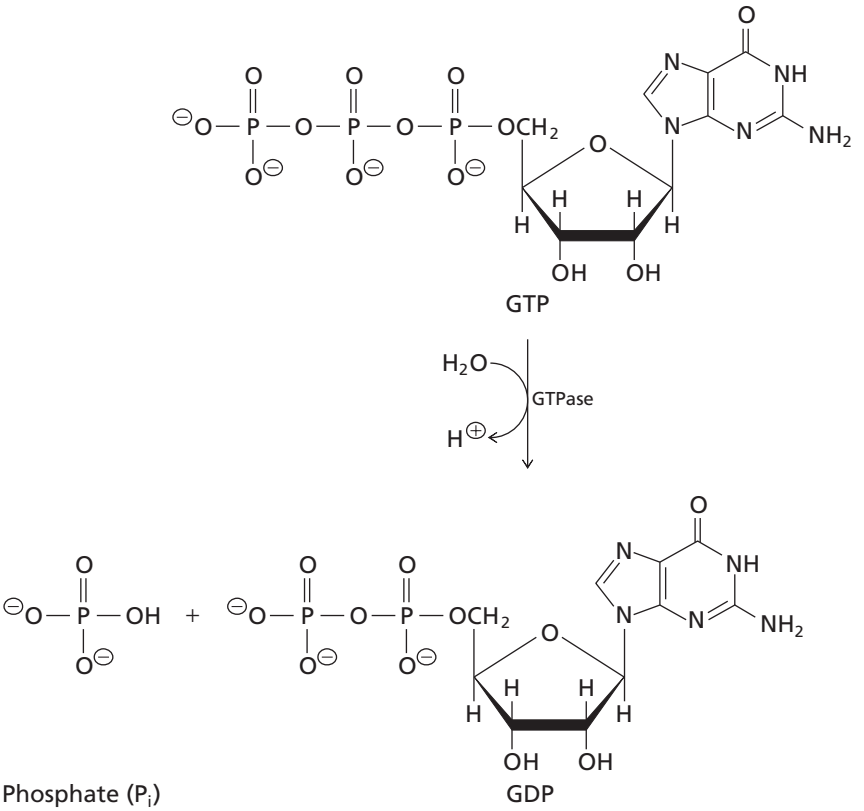
Not all chemical stimuli follow the general mechanism of signal transduction shown in Figure 9.40. For example, because steroid hormones are hydrophobic, they can diffuse across the plasma membrane into the cell where they can bind to specific receptor proteins in the cytoplasm. The steroid receptor complexes are then transferred to the nucleus. The complexes bind to specific regions of DNA called hormone response elements and thereby enhance or suppress the expression of adjacent genes.

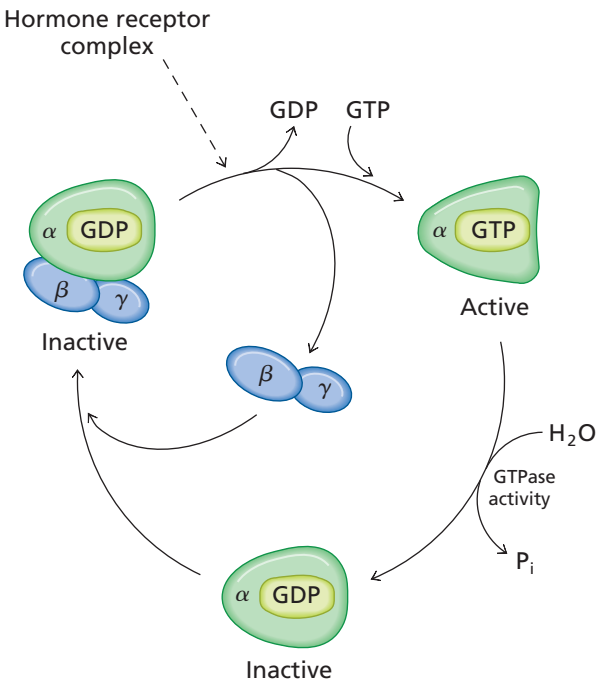
B. Signal Transducers

There are many kinds of receptors and many different transducers. Bacterial transducers are different than eukaryotic ones. There are some eukaryotic transducers found in most species. In this section, we'll concentrate on those general transducers.

Many membrane receptors interact with a family of guanine nucleotide binding proteins called **G proteins**. G proteins act as transducers—the agents that transmit external

Figure 9.41 ►
Hydrolysis of guanosine 5′-triphosphate (GTP) to guanosine 5′-diphosphate (GDP) and phosphate (P_i).





▲ Figure 9.42
G-protein cycle. G proteins undergo activation after binding to a receptor ligand complex and are slowly inactivated by their own GTPase activity. Both G_α–GTP/GDP and $G_{\beta\gamma}$ are membrane-bound.

stimuli to effector enzymes. G proteins have GTPase activity; that is, they slowly catalyze hydrolysis of bound guanosine 5′-triphosphate (GTP, the guanine analog of ATP) to guanosine 5′-diphosphate (GDP) (Figure 9.41). When GTP is bound to G protein it is active in signal tranduction and when G protein is bound to GDP it is inactive. The cyclic activation and deactivation of G proteins is shown in Figure 9.42. The G proteins involved in signaling by hormone receptors are peripheral membrane proteins located on the inner surface of the plasma membrane. Each protein consists of an α, a β, and a γ subunit. The α and γ subunits are lipid anchored membrane proteins; the α subunit is a fatty acyl anchored protein and the γ subunit is a prenyl anchored protein. The complex of $G_{\alpha\beta\gamma}$ and GDP is inactive.

When a hormone receptor complex diffusing laterally in the membrane encounters and binds $G_{\alpha\beta\gamma}$, it induces the G protein to change to an active conformation. Bound GDP is rapidly exchanged for GTP promoting the dissociation of G_α–GTP from $G_{\beta\gamma}$. Activated G_α–GTP then interacts with the effector enzyme. The GTPase activity of the G protein acts as a built-in timer since G proteins slowly catalyze the hydrolysis of GTP to GDP. When GTP is hydrolyzed the G_α–GDP complex reassociates with $G_{\beta\gamma}$ and the $G_{\alpha\beta\gamma}$–GDP complex is regenerated. G proteins have evolved into good switches because they are very slow catalysts, typically having a k_{cat} of only about 3 min^{-1}.

G proteins are found in dozens of signaling pathways including the adenylyl cyclase and the inositol-phospholipid pathways discussed below. An effector enzyme can respond to stimulatory G proteins (Gs) or inhibitory G proteins (Gi). The α subunits of different G proteins are distinct providing varying specificity but the β and γ subunits are similar and often interchangeable. Humans have two dozen α proteins, five β proteins, and six γ proteins.

C. The Adenylyl Cyclase Signaling Pathway

The cyclic nucleotides 3′,5′-cyclic adenosine monophosphate (cAMP) and its guanine analog, 3′,5′-cyclic guanosine monophosphate (cGMP), are second messengers that help transmit signals from external sources to intracellular enzymes. cAMP is produced from ATP by the action of adenylyl cyclase (Figure 9.43) and cGMP is formed from GTP in a similar reaction.

Many hormones that regulate intracellular metabolism exert their effects on target cells by activating the adenylyl cyclase signaling pathway. Binding of a hormone to a stimulatory receptor causes the conformation of the receptor to change promoting interaction between the receptor and a stimulatory G protein, G_s. The receptor ligand complex activates G_s that, in turn, binds the effector enzyme adenylyl cyclase and activates it by allosterically inducing a conformational change at its active site.

Adenylyl cyclase is an integral membrane enzyme whose active site faces the cytosol. It catalyzes the formation of cAMP from ATP. cAMP then diffuses from the membrane surface through the cytosol and activates an enzyme known as protein kinase A. This kinase is made up of a dimeric regulatory subunit and two catalytic subunits and is inactive in its fully assembled state. When the cytosolic concentration of cAMP increases as a result of signal transduction through adenylyl cyclase, four molecules of cAMP bind to the regulatory subunit of the kinase releasing the two catalytic subunits, which are enzymatically active (Figure 9.44). Protein kinase A, a serine-threonine protein kinase, catalyzes phosphorylation of the hydroxyl groups of specific serine and threonine residues in target enzymes. Phosphorylation of amino acid side chains on the target enzymes is reversed by the action of protein phosphatases that catalyze hydrolytic removal of the phosphoryl groups.

The ability to turn off a signal transduction pathway is an essential element of all signaling processes. For example, the cAMP concentration in the cytosol increases only transiently. A soluble cAMP phosphodiesterase catalyzes the hydrolysis of cAMP to AMP (Figure 9.43) limiting the lifetime of the second messenger. At high concentrations, the methylated purines caffeine and theophylline (Figure 9.45) inhibit cAMP phosphodiesterase, thereby decreasing the rate of conversion of cAMP to AMP. These inhibitors prolong and intensify the effects of cAMP and hence the activating effects of the stimulatory hormones.

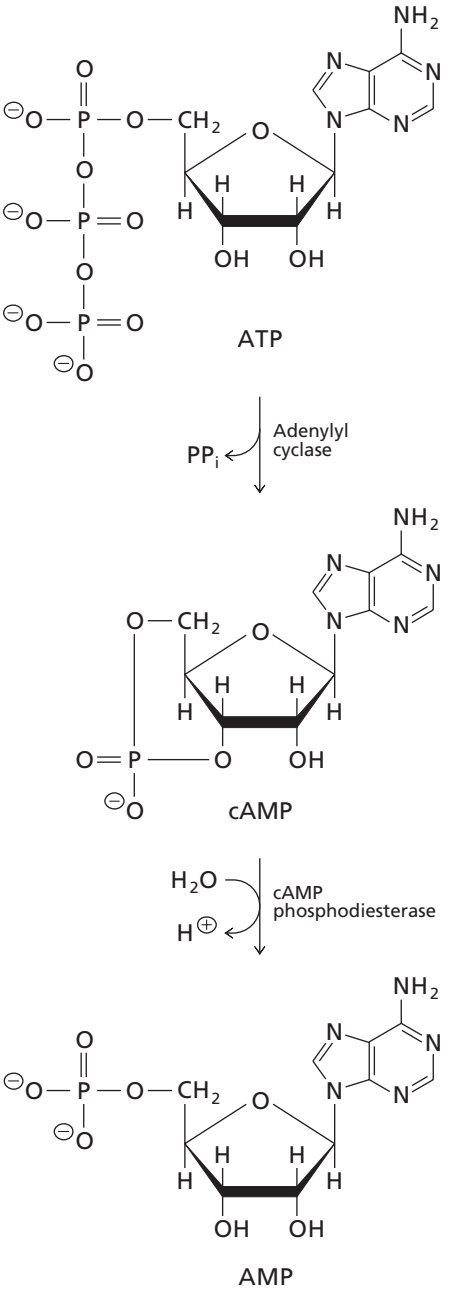
Hormones that bind to stimulatory receptors activate adenylyl cyclase and raise intracellular cAMP levels. Hormones that bind to inhibitory receptors inhibit adenylyl cyclase activity via receptor interaction with the transducer G_i. The ultimate response of a cell to a hormone depends on the type of receptors present and the type of G protein to which they are coupled. The main features of the adenylyl cyclase signaling pathway, including G proteins, are summarized in Figure 9.46.

D. The Inositol–Phospholipid Signaling Pathway

Another major signal transduction pathway produces two different second messengers, both derived from a plasma membrane phospholipid called phosphatidylinositol 4,5-*bis*phosphate (PIP_2) (Figure 9.47). PIP_2 is a minor component of plasma membranes located in the inner monolayer. It is synthesized from phosphatidylinositol by two successive phosphorylation steps catalyzed by ATP-dependent kinases.

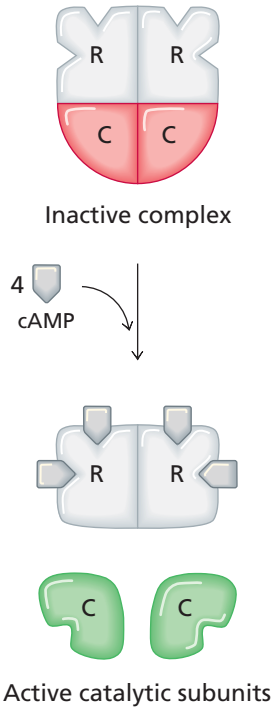
Following binding of a ligand to a specific receptor, the signal is transduced through the G protein G_q. The active GTP-bound form of G_q activates the effector enzyme phosphoinositide-specific phospholipase C that is bound to the cytoplasmic face of the plasma membrane. Phospholipase C catalyzes the hydrolysis of PIP_2 to inositol 1,4,5-*tris*phosphate (IP_3) and diacylglycerol (Figure 9.47). Both IP_3 and diacylglycerol are second messengers that transmit the original signal to the interior of the cell.

IP_3 diffuses through the cytosol and binds to a calcium channel in the membrane of the endoplasmic reticulum. This causes the calcium channel to open for a short time, releasing Ca^{2+} from the lumen of the endoplasmic reticulum into the cytosol. Calcium is also an intracellular messenger because it activates calcium-dependent protein

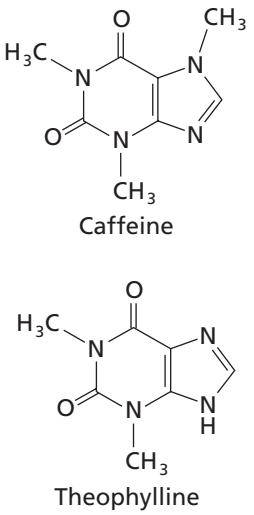


▲ **Figure 9.43**
Production and inactivation of cAMP. ATP is converted to cAMP by the transmembrane enzyme adenylyl cyclase. The second messenger is subsequently converted to 5′-AMP by the action of a cytosolic cAMP phosphodiesterase.

The response of *E. coli* to changes in glucose concentrations, modulated by cAMP, is described in Section 21.7B.



▲ **Figure 9.44**
Activation of protein kinase A. The assembled complex is inactive. When four molecules of cAMP bind to the regulatory subunit (R) dimer, the catalytic subunits (C) are released.



▲ **Figure 9.45**
Caffeine and theophylline.

kinases that catalyze phosphorylation of various protein targets. The calcium signal is short-lived since Ca^{2+} is pumped back into the lumen of the endoplasmic reticulum when the channel closes.

The other product of PIP_2 hydrolysis, diacylglycerol, remains in the plasma membrane. Protein kinase C, which exists in equilibrium between a soluble cytosolic form and a peripheral membrane form, moves to the inner face of the plasma membrane where it binds transiently and is activated by diacylglycerol and Ca^{2+}. Protein kinase C catalyzes phosphorylation of many target proteins altering their catalytic activity. Several protein kinase C isozymes exist, each with different catalytic properties and tissue distribution. They are members of the serine–threonine kinase family.

Signaling via the inositol–phospholipid pathway is turned off in several ways. First, when GTP is hydrolyzed, G_q returns to its inactive form and no longer stimulates phospholipase C. The activities of IP_3 and diacylglycerol are also transient. IP_3 is rapidly hydrolyzed to other inositol phosphates (which can also be second messengers) and inositol. Diacylglycerol is rapidly converted to phosphatidate. Both inositol and phosphatidate are recycled back to phosphatidylinositol. The main features of the inositol–phospholipid signaling pathway are summarized in Figure 9.48.

Phosphatidylinositol is not the only membrane lipid that gives rise to second messengers. Some extracellular signals lead to the activation of hydrolases that catalyze the conversion of membrane sphingolipids to sphingosine, sphingosine 1-phosphate, or ceramide. Sphingosine inhibits protein kinase C, and ceramide activates a protein kinase and a protein phosphatase. Sphingosine 1-phosphate can activate phospholipase

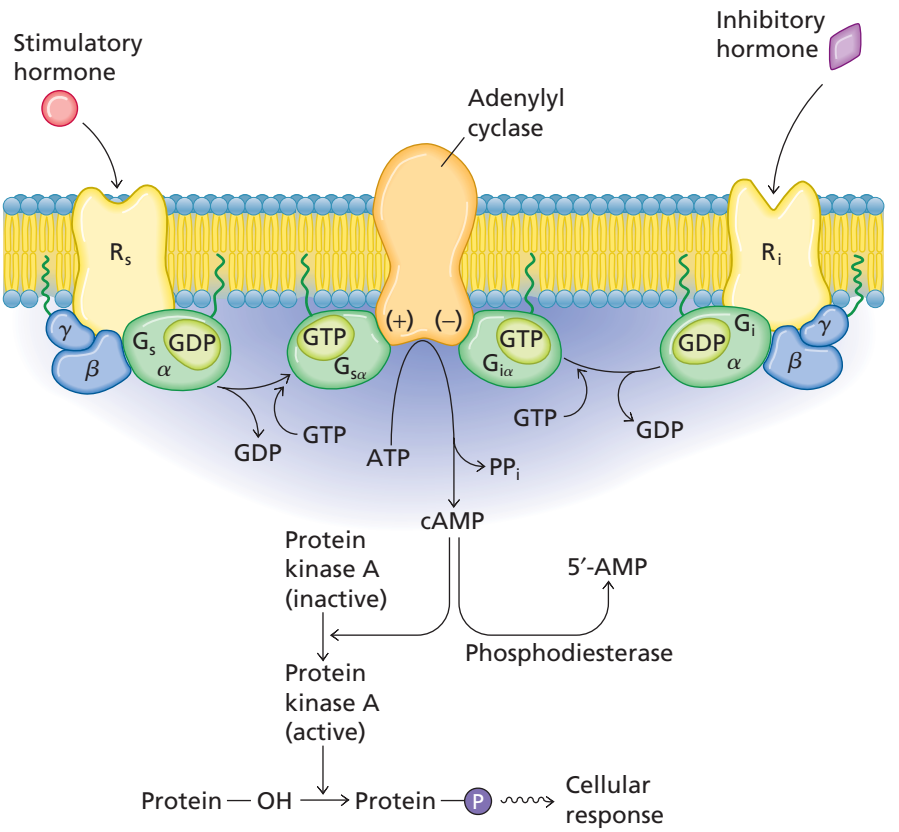
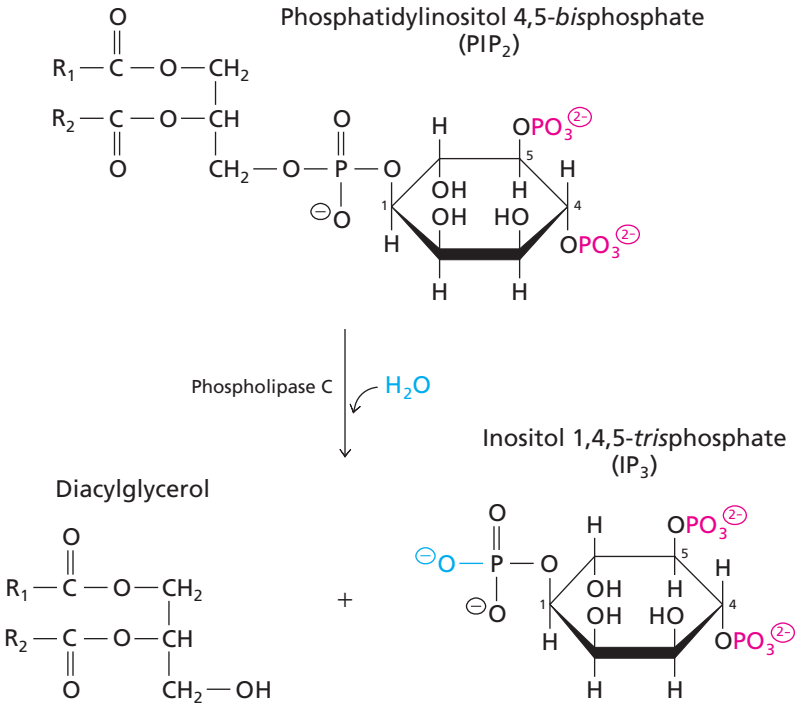
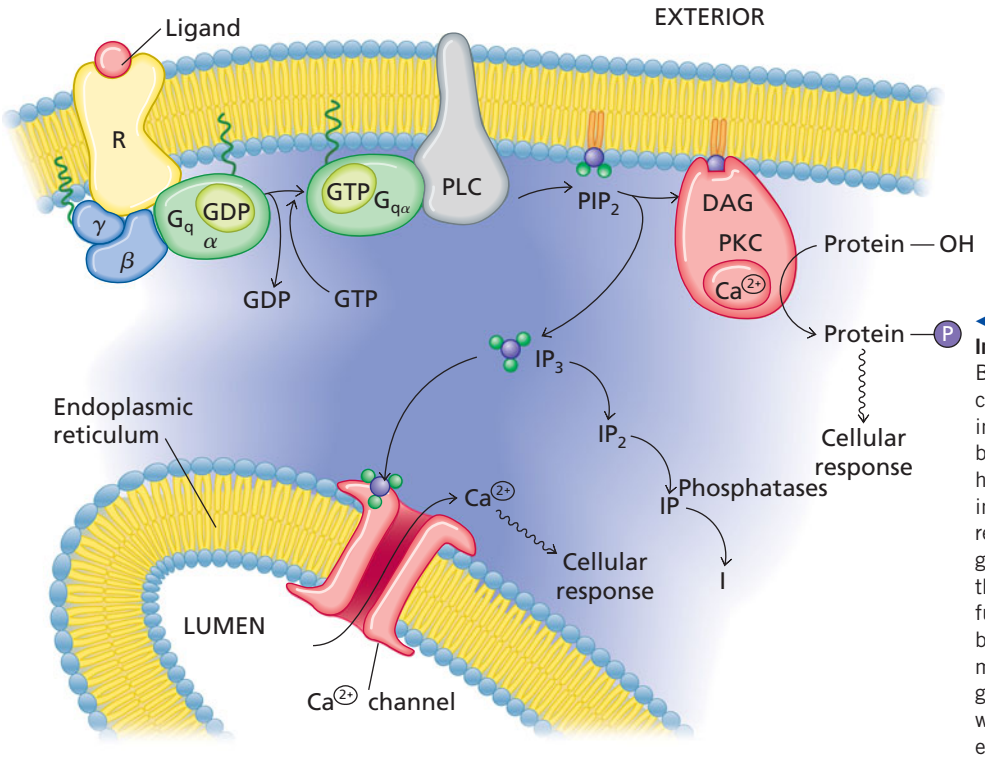


Figure 9.46 ▲
Summary of the adenylyl cyclase signaling pathway. Binding of a hormone to a stimulatory transmembrane receptor (R_s) leads to activation of the stimulatory G protein (G_s) on the inside of the membrane. Other hormones can bind to inhibitory receptors (R_i) that are coupled to adenylyl cyclase by the inhibitory G protein G_i. G_s activates the integral membrane enzyme adenylyl cyclase whereas G_i inhibits it. cAMP activates protein kinase A resulting in the phosphorylation of cellular proteins.



◀ **Figure 9.47**
Phosphatidylinositol 4,5-*bis*phosphate (PIP_2). Phosphatidylinositol 4,5-*bis*phosphate (PIP_2) produces two second messengers, inositol 1,4,5-*tris*phosphate (IP_3) and diacylglycerol. PIP_2 is synthesized by the addition of two phosphoryl groups (red) to phosphatidylinositol and hydrolyzed to IP_3 and diacylglycerol by the action of a phosphoinositide-specific phospholipase C.

D, which specifically catalyzes hydrolysis of phosphatidylcholine. The phosphatidate and the diacylglycerol formed by this hydrolysis appear to be second messengers. The full significance of the wide variety of second messengers generated from membrane lipids (each with its own specific fatty acyl groups) has not yet been determined.



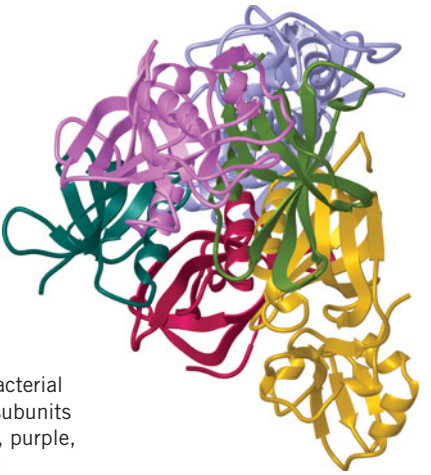
◀ **Figure 9.48**
Inositol–phospholipid signaling pathway. Binding of a ligand to its transmembrane receptor (R) activates the G protein (G_q). This in turn stimulates a specific membrane-bound phospholipase C (PLC) that catalyzes hydrolysis of the phospholipid PIP_2 in the inner leaflet of the plasma membrane. The resulting second messengers, IP_3 and diacylglycerol (DAG), are responsible for carrying the signal to the interior of the cell. IP_3 diffuses to the endoplasmic reticulum where it binds to and opens a Ca^{2+} channel in the membrane releasing stored Ca^{2+}. Diacylglycerol remains in the plasma membrane where it—along with Ca^{2+}—activates the enzyme protein kinase C (PKC).

BOX 9.7 BACTERIAL TOXINS AND G PROTEINS

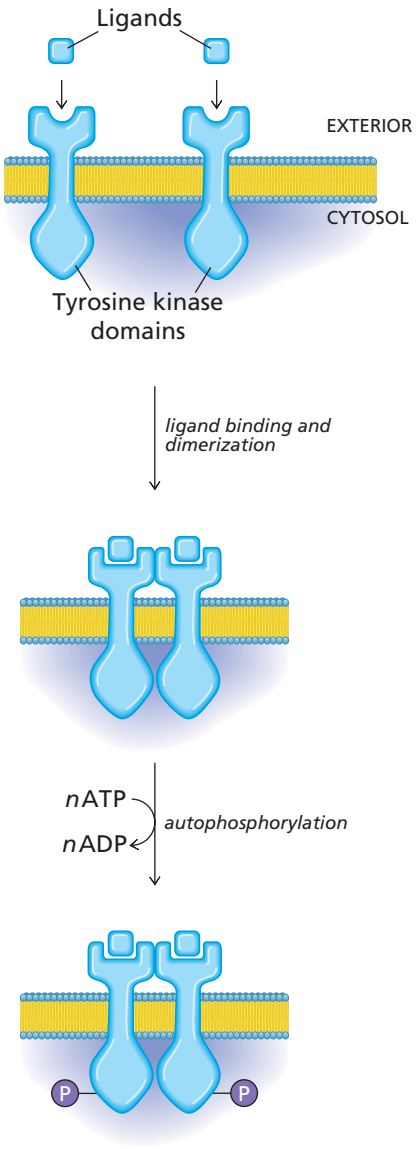
G proteins are the biological targets of cholera and pertussis (whooping cough) toxins that are secreted by the disease-producing bacteria *Vibrio cholerae* and *Bordetella pertussis*, respectively. Both diseases involve overproduction of cAMP.

Cholera toxin binds to ganglioside G_{M1} on the cell surface (Section 9.5) and a subunit of it crosses the plasma membrane and enters the cytosol. This subunit catalyzes covalent modification of the α subunit of the G protein G_s inactivating its GTPase activity. The adenylyl cyclase of these cells remains activated and cAMP levels stay high. In people infected with *V. cholerae*, cAMP stimulates certain transporters in the plasma membrane of the intestinal cells leading to a massive secretion of ions and water into the gut. The dehydration resulting from diarrhea can be fatal unless fluids are replenished.

Pertussis toxin binds to a glycolipid called lactosylceramide found on the cell surface of epithelial cells in the lung. It is taken up by endocytosis. The toxin catalyzes covalent modification of G_i. In this case, the modified G protein is unable to replace GDP with GTP and therefore adenylyl cyclase activity cannot be reduced via inhibitory receptors. The resulting increase in cAMP levels produces the symptoms of whooping cough.



► **Pertussis toxin.** The bacterial toxin has five different subunits colored red, green, blue, purple, and yellow. [PDB 1BCP]



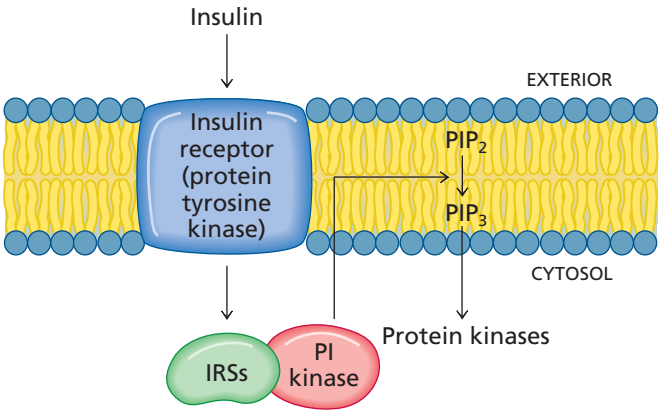
E. Receptor Tyrosine Kinases

Many growth factors operate by a signaling pathway that includes a multifunctional transmembrane protein called a receptor tyrosine kinase. As shown in Figure 9.49, the receptor, transducer, and effector functions are all found in a single membrane protein. In one type of activation, a ligand binds to the extracellular domain of the receptor, activating tyrosine kinase catalytic activity in the intracellular domain by dimerization of the receptor. When two receptor molecules associate, each tyrosine kinase domain catalyzes the phosphorylation of specific tyrosine residues of its partner, a process called *autophosphorylation*. The activated tyrosine kinase then catalyzes phosphorylation of certain cytosolic proteins, setting off a cascade of events in the cell.

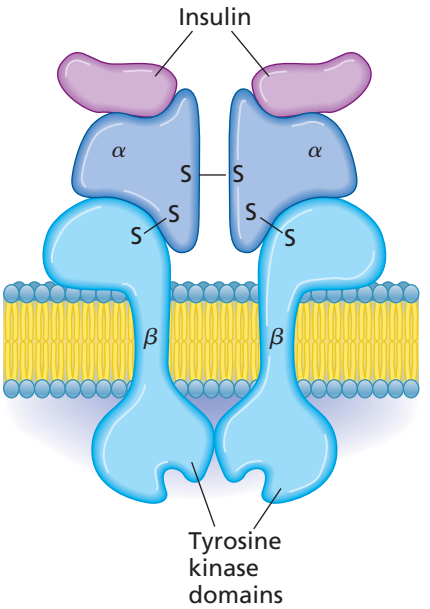
The insulin receptor is an $\alpha_2\beta_2$ tetramer (Figure 9.50). When insulin binds to the α subunit, it induces a conformational change that brings the tyrosine kinase domains of the β subunits together. Each tyrosine kinase domain in the tetramer catalyzes the phosphorylation of the other kinase domain. The activated tyrosine kinase also catalyzes the phosphorylation of tyrosine residues in other proteins that help regulate nutrient utilization.

Recent research has found that many of the signaling actions of insulin are mediated through PIP_2 (Section 9.12C and Figure 9.51). Rather than causing hydrolysis of PIP_2, insulin (via proteins called insulin receptor substrates, IRSs) activates phosphotidylinositol 3-kinase, an enzyme that catalyzes the phosphorylation of PIP_2 to phosphatidylinositol 3,4,5-*tris*phosphate (PIP_3). PIP_3 is a second messenger that transiently activates a series of target proteins, including a specific phosphoinositide-dependent protein kinase. In this way, phosphotidylinositol 3-kinase is the molecular switch that regulates several serine–threonine protein kinase cascades.

◄ **Figure 9.49**
Activation of receptor tyrosine kinases. Activation occurs as a result of ligand induced receptor dimerization. Each kinase domain catalyzes phosphorylation of its partner. The phosphorylated dimer can catalyze phosphorylation of various target proteins.



▲ **Figure 9.51**
Insulin-stimulated formation of phosphatidylinositol 3,4,5-*tris*phosphate (PIP_3). Binding of insulin to its receptor activates the protein tyrosine kinase activity of the receptor leading to the phosphorylation of insulin receptor substrates (IRSs). The phosphorylated IRSs interact with phosphotidylinositiol 3-kinase (PI kinase) at the plasma membrane where the enzyme catalyzes the phosphorylation of PIP_2 to PIP_3. PIP_3 acts as a second messenger carrying the message from extracellular insulin to certain intracellular protein kinases.



▲ **Figure 9.50**
Insulin receptor. Two extracellular α chains, each with an insulin binding site, are linked to two transmembrane β chains, each with a cytosolic tyrosine kinase domain. Following insulin binding to the α chains, the tyrosine kinase domain of each β chain catalyzes autophosphorylation of tyrosine residues in the adjacent kinase domain. The tyrosine kinase domains also catalyze the phosphorylation of proteins called insulin receptor substrates (IRSs).

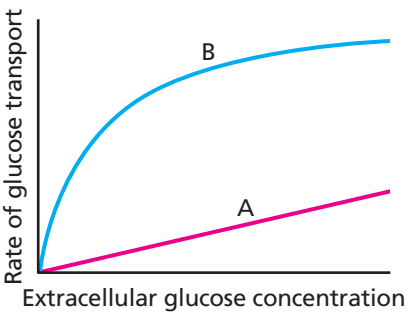
Phosphoryl groups are removed from both the growth factor receptors and their protein targets by the action of protein tyrosine phosphatases. Although only a few of these enzymes have been studied, they appear to play an important role in regulating the tyrosine kinase signaling pathway. One means of regulation appears to be the localized assembly and separation of enzyme complexes.

Summary

1. Lipids are a diverse group of water-insoluble organic compounds.
2. Fatty acids are monocarboxylic acids, usually with an even number of carbon atoms ranging from 12 to 20.
3. Fatty acids are generally stored as triacylglycerols (fats and oils), which are neutral and nonpolar.
4. Glycerophospholipids have a polar head group and nonpolar fatty acyl tails linked to a glycerol backbone.
5. Sphingolipids, which occur in plant and animal membranes, contain a sphingosine backbone. The major classes of sphingolipids are sphingomyelins, cerebrosides, and gangliosides.
6. Steroids are isoprenoids containing four fused rings.
7. Other biologically important lipids are waxes, eicosanoids, lipid vitamins, and terpenes.
8. The structural basis for all biological membranes is the lipid bilayer that includes amphipathic lipids such as glycerophospholipids, sphingolipids, and sometimes cholesterol. Lipids can diffuse rapidly within a leaflet of the bilayer.
9. A biological membrane contains proteins embedded in or associated with a lipid bilayer. The proteins can diffuse laterally within the membrane.
10. Most integral membrane proteins span the hydrophobic interior of the bilayer, but peripheral membrane proteins are more loosely associated with the membrane surface. Lipid anchored membrane proteins are covalently linked to lipids in the bilayer.
11. Some small or hydrophobic molecules can diffuse across the bilayer. Channels, pores, and passive and active transporters mediate the movement of ions and polar molecules across membranes. Macromolecules can be moved into and out of the cell by endocytosis and exocytosis, respectively.
12. Extracellular chemical stimuli transmit their signals to the cell interior by binding to receptors. A transducer passes the signal to an effector enzyme, which generates a second messenger. Signal transduction pathways often include G proteins and protein kinases. The adenylyl cyclase signaling pathway leads to activation of the cAMP-dependent protein kinase A. The inositol-phospholipid signaling pathway generates two second messengers and leads to the activation of protein kinase C and an increase in the cytosolic Ca^{2+} concentration. In receptor tyrosine kinases, the kinase is part of the receptor protein.

Problems

1. Write the molecular formulas for the following fatty acids: (a) nervonic acid (*cis*-Δ^{15}-tetracosenoate; 24 carbons); (b) vaccenic acid (*cis*-Δ^{11}-octadecenoate); and (c) EPA (all *cis*-$\Delta^{5,8,11,14,17}$-eicosapentaenoate).
2. Write the molecular formulas for the following modified fatty acids:
   (a) 10-(Propoxy) decanoate, a synthetic fatty acid with antiparasitic activity used to treat African sleeping sickness, a disease caused by the protozoan *T. brucei* (the propoxy group is $—O—CH_2CH_2CH_3$)
   (b) Phytanic acid (3,7,11,15-tetramethylhexadecanoate), found in dairy products
   (c) Lactobacillic acid (*cis*-11,12-methyleneoctadecanoate), found in various microorganisms
3. Fish ois are rich sources of omega-3 and polyunsaturated fatty acids and omega-6 fatty acids are relatively abundant in corn and sunflower oils. Classify the following fatty acids as omega-3, omega-6, or neither: (a) linolenate, (b) linoleate, (c) arachidonate, (d) oleate, (e) $\Delta^{8,11,14}$-eicosatrienoate.
4. Mammalian platelet activating factor (PAF), a messenger in signal transduction, is a glycerophospholipid with an ether linkage at C-1. PAF is a potent mediator of allergic responses, inflammation, and the toxic-shock syndrome. Draw the structure of PAF (1-alkyl-2-acetyl-phosphatidyl-choline), where the 1-alkyl group is a C_{16} chain.
5. Docosahexaenoic acid, 22:6 $\Delta^{4,7,10,13,16,19}$, is the predominate fatty acyl group in the C-2 position of glycerol-3-phosphate in phosphatidylethanolamine and phosphatidylcholine in many types of fish.
   (a) Draw the structure of docosahexaenoic acid (all double bonds are *cis*).
   (b) Classify docosahexaenoic acid as an omega-3, omega-6, or omega-9 fatty acid.
6. Many snake venoms contain phospholipase A_2 that catalyzes the degradation of glycerophospholipids into a fatty acid and a "lysolecithin." The amphipathic nature of lysolecithins allows them to act as detergents in disrupting the membrane structure of red blood cells, causing them to rupture. Draw the structures of phosphatidyl serine (PS) and the products (including a lysolecithin) that result from the reaction of PS with phospholipase A_2.
7. Draw the structures of the following membrane lipids:
   (a) 1-stearoyl-2-oleoyl-3-phosphatidylethanolamine
   (b) palmitoylsphingomyelin
   (c) myristoyl-β-D-glucocerebroside.
8. (a) The steroid cortisol participates in the control of carbohydrate, protein, and lipid metabolism. Cortisol is derived from cholesterol and possesses the same four-membered fused ring system but with: (1) a C-3 keto group, (2) C-4-C-5 double bond (instead of the C-5-C-6 as in cholesterol), (3) a C-11 hydroxyl, and (4) a hydroxyl group and a $—C(O)CH_2OH$ group at C-17. Draw the structure of cortisol.
   (b) Ouabain is a member of the cardiac glycoside family found in plants and animals. This steroid inhibits $Na^{\oplus}$–$K^{\oplus}$ ATPase and ion transport and may be involved in hypertension and high blood pressure in humans. Ouabain possesses a four-membered fused ring system similar to cholesterol but has the following structural features: (1) no double bonds in the rings, (2) hydroxy groups on C-1, C-5, C-11, and C-14, (3) $—CH_2OH$ on C-19, (4) 2-3 unsaturated five-membered lactone ring on C-17 (attached to C-3 of lactone ring), and (5) 6-deoxymannose attached β-1 to the C-3 oxygen. Draw the structure of ouabain.
9. A consistent response in many organisms to changing environmental temperatures is the restructuring of cellular membranes. In some fish, phosphatidylethanolamine (PE) in the liver microsomal lipid membrane contains predominantly docosahexaenoic acid, 22:6 $\Delta^{4,7,10,13,16,19}$ at C-2 of the glycerol-3-phosphate backbone and then either a saturated or monounsaturated fatty acyl group at C-1. The percentage of the PE containing saturated or monounsaturated fatty acyl groups was determined in fish acclimated at 10°C or 30°C. At 10°C, 61% of the PE molecules contained saturated fatty acyl groups at C-1, and 39% of the PE molecules contained monounsaturated fatty acyl groups at C-1. When fish were acclimated to 30°C, 86% of the PE lipids contained saturated fatty acyl groups at C-1, while 14% of the PE molecules had monounsaturated acyl groups at C-1 [Brooks, S., Clark, G.T., Wright, S.M., Trueman, R.J., Postle, A.D., Cossins, A.R., and Maclean, N.M. (2002). Electrospray ionisation mass spectrometric analysis of lipid restructuring in the carp (*Cyprinus carpio* L.) during cold acclimation. *J. Exp. Biol.* 205:3989–3997]. Explain the purpose of the membrane restructuring observed with the change in environmental temperature.
10. A mutant gene (*ras*) is found in as many as one-third of all human cancers including lung, colon, and pancreas, and may be partly responsible for the altered metabolism in tumor cells. The *ras* protein coded for by the *ras* gene is involved in cell signaling pathways that regulate cell growth and division. Since the *ras* protein must be converted to a lipid anchored membrane protein in order to have cell-signaling activity, the enzyme farnesyl transferase (FT) has been selected as a potential chemotherapy target for inhibition. Suggest why FT might be a reasonable target.
11. Glucose enters some cells by simple diffusion through channels or pores, but glucose enters red blood cells by passive transport. On the plot below, indicate which line represents diffusion through a channel or pore and which represents passive transport. Why do the rates of the two processes differ?



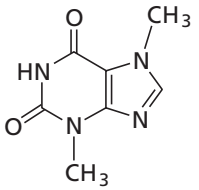
12. The pH gradient between the stomach (pH 0.8–1.0) and the gastric mucosal cells lining the stomach (pH 7.4) is maintained by an $H^{\oplus}$–$K^{\oplus}$ ATPase transport system that is similar to the ATP-driven $Na^{\oplus}$–$K^{\oplus}$ ATPase transport system (Figure 9.38). The $H^{\oplus}$–$K^{\oplus}$ ATPase antiport system uses the energy of ATP to pump $H^{\oplus}$ out of the mucosal cells (mc) into the stomach (st) in exchange for $K^{\oplus}$ ions. The $K^{\oplus}$ ions that are transported into the mucosal cells are then cotransported back into the stomach along

with $Cl^{\ominus}$ ions. The net transport is the movement of HCl into the stomach.

$$K^{\oplus}_{(mc)} + Cl^{\ominus}_{(mc)} + H^{\oplus}_{(mc)} + K^{\oplus}_{(st)} + ATP \rightleftharpoons$$
$$K^{\oplus}_{(st)} + Cl^{\ominus}_{(st)} + H^{\oplus}_{(st)} + K^{\oplus}_{(mc)} + ADP + P_i$$

Draw a diagram of this $H^{\oplus}$–$K^{\oplus}$ ATPase system.

13. Chocolate contains the compound theobromine, which is structurally related to caffeine and theophylline. Chocolate products may be toxic or lethal to dogs because these animals metabolize theobromine more slowly than humans. The heart, central nervous system, and kidneys are affected. Early signs of theobromine poisoning in dogs include nausea and vomiting, restlessness, diarrhea, muscle tremors, and increased urination or incontinence. Comment on the mechanism of toxicity of theobromine in dogs.



Theobromine

14. In the inositol signaling pathway, both IP_3 and diacylglycerol (DAG) are hormonal second messengers. If certain protein kinases in cells are activated by binding $Ca^{\oplus 2+}$, how do IP_3 and DAG act in a complementary fashion to elicit cellular responses inside cells?

15. In some forms of diabetes, a mutation in the β subunit of the insulin receptor abolishes the enzymatic activity of that subunit. How does the mutation affect the cell's response to insulin? Can additional insulin (e.g., from injections) overcome the defect?

16. The *ras* protein (described in Problem 10) is a mutated G protein that lacks GTPase activity. How does the absence of this activity affect the adenylyl cyclase signaling pathway?

17. At the momentof fertilization a female egg is about 100μm in diameter. Assuming that each lipid molecule in the plasma membrane has a suface area of 10^{-14} cm^2, how many lipid molecules are there in the egg plasma membrane if 25% of the surface is protein?

18. Each fertilized egg cell (zygote) divides 30 times to produce all the eggs that a female child will need in her lifetime. One of these eggs will be fertilized giving rise to a new generation. If lipid molecles are never degraded, how many lipid molecules have you inherited that were synthesized in your grandmother?

Selected Readings

General

Gurr, M. I., and Harwood, J. L. (1991). *Lipid Biochemistry: An Introduction*, 4th ed. (London: Chapman and Hall).

Lester, D. R., Ross, J. J., Davies, P. J., and Reid, J. B. (1997). Mendel's stem length gene (*Le*) encodes a gibberellin 3 beta-hydroxylase. *Plant Cell.* 9:1435–1443.

Vance, D. E., and Vance, J. E., eds. (2008). *Biochemistry of Lipids, Lipoproteins, and Membranes*, 5th ed. (New York: Elsevier).

Membranes

Dowhan, W. (1997). Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu. Rev. Biochem.* 66:199–232.

Jacobson, K., Sheets, E. D., and Simson, R. (1995). Revisiting the fluid mosaic model of membranes. *Science* 268:1441–1442.

Koga, Y., and Morii, H. (2007). Biosynthesis of ether-type polar lipids in Archaea and evolutionary considerations. *Microbiol. and Molec. Biol. Rev.* 71: 97–120.

Lai, E.C. (2003) Lipid rafts make for slippery platforms. *J. Cell Biol.* 162:365–370.

Lingwood, D., and Simons, K. (2010). Lipid rafts as a membrane-organizing principle. *Science.* 327:46–50.

Simons, K., and Ikonen, E. (1997). Functional rafts in cell membranes. *Nature.* 387:569–572.

Singer, S. J. (1992). The structure and function of membranes: a personal memoir. *J. Membr. Biol.* 129:3–12.

Singer, S. J. (2004) Some early history of membrane molecular biology. *Annu. Rev. Physiol.* 66:1–27.

Singer, S. J., and Nicholson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science* 175:720–731.

Membrane Proteins

Casey, P. J., and Seabra, M. C. (1996). Protein prenyltransferases. *J. Biol. Chem.* 271:5289–5292.

Bijlmakers, M-J., and Marsh, M. (2003). The on-off story of protein palmitoylation. *Trends in Cell Biol.* 13:32–42.

Elofsson, A., and von Heijne, G. (2007). Membrane protein structure: prediction versus reality. *Annu. Rev. Biochem.* 76:125–140.

Membrane Transport

Borst, P., and Elferink, R. O. (2002). Mammalian ABC transporters in health and disease. *Annu. Rev. Biochem.* 71:537–592.

Caterina, M. J., Schumacher, M. A., Tominaga, M., Rosen, T. A., Levine, J. D., and Julius, D. (1997). The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 389:816–824.

Clapham, D. (1997). Some like it hot: spicing up ion channels. *Nature* 389:783–784.

Costanzo, M. et. al. (2010). The genetic landscape of a cell. *Science* 327:425–432.

Doherty, G. J. and McMahon, H. T. (2009). Mechanisms of endocytosis. *Annu. Rev. Biochem.* 78:857–902.

Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and McKinnon, R. (1998). The structure of the potassium channel: molecular basis of $K^{\oplus}$ conduction and selectivity. *Science* 280:69–75.

Jahn, R., and Südhof, T. C. (1999). Membrane fusion and exocytosis. *Annu. Rev. Biochem.* 68:863–911.

Kaplan, J. H. (2002). Biochemistry of Na, K-AT-Pase. *Annu. Rev. Biochem.* 71:511–535.

Loo, T. W., and Clarke, D. M. (1999). Molecular dissection of the human multidrug resistance P-glycoprotein. *Biochem. Cell Biol.* 77:11–23.

Signal Transduction

Fantl, W. J., Johnson, D. E., and Williams, L. T. (1993). Signalling by receptor tyrosine kinases. *Annu. Rev. Biochem.* 62:453–481.

Hamm, H. E. (1998). The many faces of G protein signaling. *J. Biol. Chem.* 273:669–672.

Hodgkin, M. N., Pettitt, T. R., Martin, A., Michell, R. H., Pemberton, A. J., and Wakelam, M. J. O. (1998). Diacylglycerols and phosphatidates: which molecular species are intracellular messengers? *Trends Biochem. Sci.* 23:200–205.

Hurley, J. H. (1999). Structure, mechanism, and regulation of mammalian adenylyl cyclase. *J. Biol. Chem.* 274:7599–7602.

Luberto, C., and Hannun, Y. A. (1999). Sphingolipid metabolism in the regulation of bioactive molecules. *Lipids* 34 (Suppl.):S5–S11.

Prescott, S. M. (1999). A thematic series on kinases and phosphatases that regulate lipid signaling. *J. Biol. Chem.* 274:8345.

Shepherd, P. R., Withers, D. J., and Siddle, K. (1998). Phosphoinositide 3-kinase: the key switch mechanism in insulin signalling. *Biochem. J.* 333:471–490.





Introduction to Metabolism

For most metabolic sequences neither the substrate concentration nor the product concentration changes significantly, even though the flux through the pathway may change dramatically.

—Jeremy R. Knowles (1989)

In the preceding chapters, we described the structures and functions of the major components of living cells from small molecules to polymers to larger aggregates such as membranes. The next nine chapters focus on the biochemical activities that assimilate, transform, synthesize, and degrade many of the nutrients and cellular components already described. The biosynthesis of proteins and nucleic acids, which represent a significant proportion of the activity of all cells, will be described in Chapters 20–22.

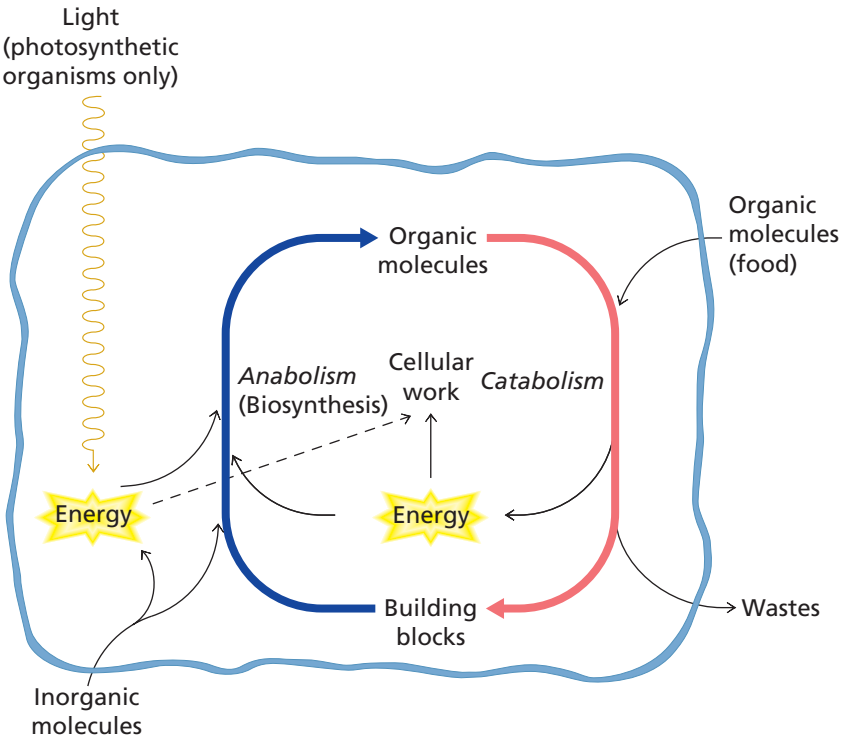
We now move from molecular structure to the dynamics of cell function. Despite the marked shift in our discussion, we will see that metabolic pathways are governed by basic chemical and physical laws. By taking a stepwise approach that builds on the foundations established in the first two parts of this book, we can describe how metabolism operates. In this chapter, we discuss some general themes of metabolism and the thermodynamic principles that underlie cellular activities.

10.1 Metabolism Is a Network of Reactions

Metabolism is the entire network of chemical reactions carried out by living cells. **Metabolites** are the small molecules that are intermediates in the degradation or biosynthesis of biopolymers. The term *intermediary metabolism* is applied to the reactions involving these low-molecular-weight molecules. It is convenient to distinguish between reactions that synthesize molecules (anabolic reactions) and reactions that degrade molecules (catabolic reactions).

Anabolic reactions are those responsible for the synthesis of all compounds needed for cell maintenance, growth, and reproduction. These biosynthesis reactions make simple metabolites such as amino acids, carbohydrates, coenzymes, nucleotides, and

Top: The fundamental principles of metabolism are the same in animals and plants and in all other organisms.



◀ **Figure 10.1**
Anabolism and catabolism. Anabolic reactions use small molecules and chemical energy in the synthesis of macromolecules and in the performance of cellular work. Solar energy is an important source of metabolic energy in photosynthetic bacteria and plants. Some molecules, including those obtained from food, are catabolized to release energy and either monomeric building blocks or waste products.

fatty acids. They also produce larger molecules such as proteins, polysaccharides, nucleic acids, and complex lipids (Figure 10.1).

In some species, all of the complex molecules that make up a cell are synthesized from inorganic precursors (carbon dioxide, ammonia, inorganic phosphates, etc.)(Section 10.3). Some species derive energy from these inorganic molecules or from the creation of membrane potential (Section 9.11). Photosynthetic organisms use light energy to drive biosynthesis reactions (Chapter 15).

Catabolic reactions degrade large molecules to liberate smaller molecules and energy. All cells carry out degradation reactions as part of their normal cell metabolism but some species rely on them as their only source of energy. Animals, for example, require organic molecules as food. The study of these energy-producing catabolic reactions in mammals is called *fuel metabolism*. The ultimate source of these fuels is a biosynthetic pathway in another species. Keep in mind that all catabolic reactions involve the breakdown of compounds that were synthesized by a living cell—either the same cell, a different cell in the same individual, or a cell in a different organism.

There is a third class of reactions called **amphibolic reactions**. They are involved in both anabolic and catabolic pathways.

Whether we observe bacteria or large multicellular organisms, we find a bewildering variety of biological adaptations. More than 10 million species may be living on Earth and several hundred million species may have come and gone throughout the course of evolution. Multicellular organisms have a striking specialization of cell types or tissues. Despite this extraordinary diversity of species and cell types the biochemistry of living cells is surprisingly similar not only in the chemical composition and structure of cellular components but also in the metabolic routes by which the components are modified. These universal pathways are the key to understanding metabolism. Once you've learned about the fundamental conserved pathways you can appreciate the additional pathways that have evolved in some species.

The complete sequences of the genomes of a number of species have been determined. For the first time we are beginning to have a complete picture of the entire metabolic network of these species based on the sequences of the genes that encode metabolic enzymes. *Escherichia coli*, for example, has about 900 genes that encode enzymes used in intermediary metabolism and these enzymes combine to create about 130 different pathways.

KEY CONCEPT

Most of the fundamental metabolic pathways are present in all species.

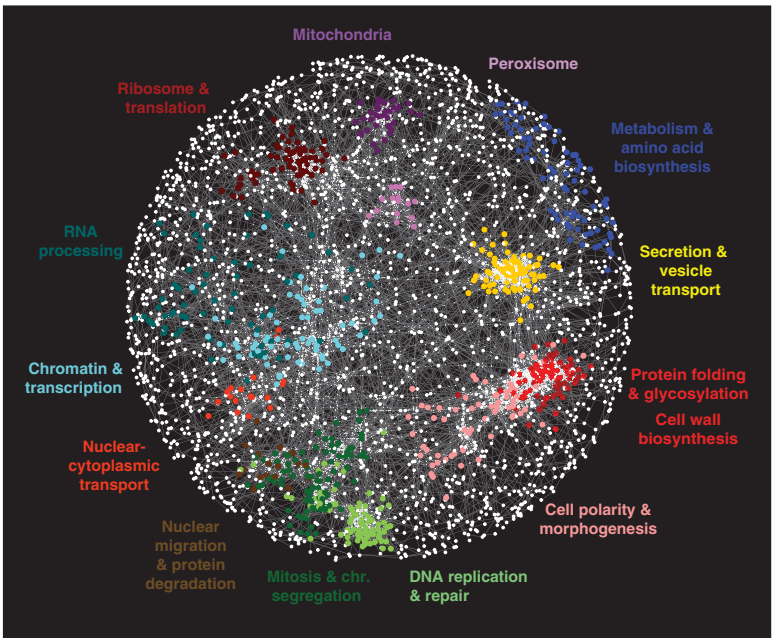


Figure 10.2 ▶
A protein interaction network for yeast (*Saccharomyces cerevisiae*). Dots represent individual proteins, colored according to function. Solid lines represent interactions between proteins. The colored clusters identify the large number of genes involved in metabolism.

These metabolic genes account for 21% of the genes in the genome. Other species of bacteria have a similar number of enzymes that carry out the basic metabolic reactions. Some species contain additional pathways. The bacterium that causes tuberculosis, *Mycobacterium tuberculosis*, has about 250 enzymes involved in fatty acid metabolism—five times as many as *E. coli*.

The yeast *Saccharomyces cerevisiae* is a single-celled member of the fungus kingdom. Its genome contains 5900 protein-encoding genes. Of these, 1200 (20%) encode enzymes involved in intermediary and energy metabolism (Figure 10.2). The nematode *Caenorhabditis elegans* is a small, multicellular animal with many of the same specialized cells and tissues found in larger animals. Its genome encodes 19,100 proteins of which 5300 (28%) are thought to be required in various pathways of intermediary metabolism. In the fruit fly, *Drosophila melanogaster*, approximately 2400 (17%) of its 14,100 genes are predicted to be involved in intermediary metabolic pathways and bioenergetics. The exact number of genes required for basic metabolism in humans is not known but it's likely that about 5000 genes are needed. (The human genome has approximately 22,000 genes.)

There are five common themes in metabolism.

1. Organisms or cells maintain specific internal concentrations of inorganic ions, metabolites, and enzymes. Cell membranes provide the physical barrier that segregates cell components from the environment.
2. Organisms extract energy from external sources to drive energy-consuming reactions. Photosynthetic organisms derive energy from the conversion of solar energy to chemical energy. Other organisms obtain energy from the ingestion and catabolism of energy-yielding compounds.
3. The metabolic pathways in each organism are specified by the genes it contains in its genome.
4. Organisms and cells interact with their environment. The activities of cells must be geared to the availability of energy, organisms grow and reproduce. When the supply of energy from the environment is plentiful. When the supply of energy from the environment is limited, energy demands can be temporarily met by using internal stores or by slowing metabolic rates as in hibernation, sporulation, or seed formation. If the shortage is prolonged, organisms die.
5. The cells of organisms are not static assemblies of mtneylecules. Many cell components are continually synthesized and degraded, that is, they undergo *turnover*, even

though their concentrations may remain virtually constant. The concentrations of other compounds change in response to changes in external or internal conditions.

The metabolism section of this book describes metabolic reactions that operate in most species. For example, enzymes of glycolysis (the degradation of sugar) and of gluconeogenesis (biosynthesis of glucose) are present in almost all species. Although most cells possess the same set of central metabolic reactions, cell and organism differentiation is possible because of additional enzymatic reactions specific to the tissue or species.

10.2 Metabolic Pathways

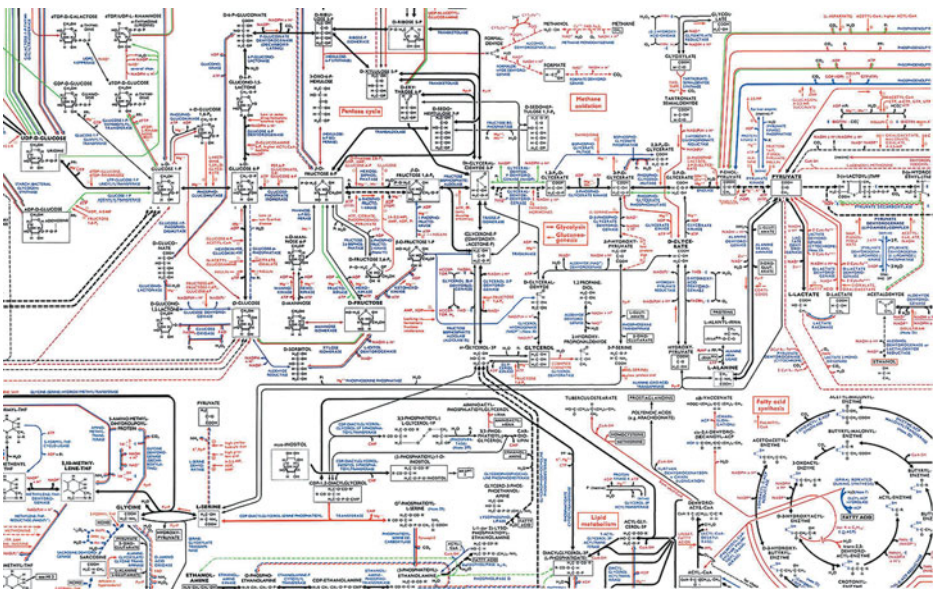
The vast majority of metabolic reactions are catalyzed by enzymes so a complete description of metabolism includes not only the reactants, intermediates, and products of cellular reactions but also the characteristics of the relevant enzymes. Most cells can perform hundreds to thousands of reactions. We can deal with this complexity by systematically subdividing metabolism into segments or branches. In the following chapters, we begin by considering separately the metabolism of the four major groups of biomolecules: carbohydrates, lipids, amino acids, and nucleotides. Within each of the four areas of metabolism, we recognize distinct sequences of metabolic reactions, called pathways.

A. Pathways Are Sequences of Reactions

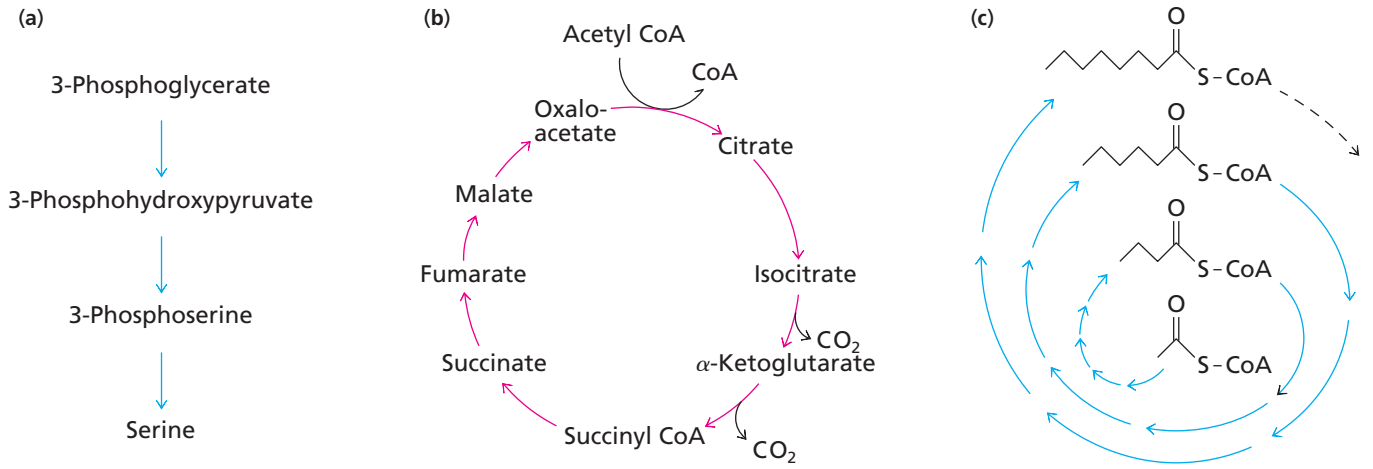
A **metabolic pathway** is the biological equivalent of a synthesis scheme in organic chemistry. A metabolic pathway is a series of reactions where the product of one reaction becomes the substrate for the next reaction. Some metabolic pathways may consist of only two steps while others may be a dozen steps in length.

It's not easy to define the limits of a metabolic pathway. In the laboratory, a chemical synthesis has an obvious beginning substrate and an obvious end product but cellular pathways are interconnected in ways that make it difficult to pick a beginning and an end. For example, in the catabolism of glucose (Chapter 11), where does glycolysis begin and end? Does it begin with polysaccharides (such as glycogen and starch), extracellular glucose, glucose 6-phosphate, or intracellular glucose? Does the pathway end with pyruvate, acetyl CoA, lactate, or ethanol? Start and end points can be assigned somewhat arbitrarily, often according to tradition or for ease of study, but keep in mind that reactions and pathways can be linked to form extended metabolic routes. This network is very obvious when you examine the large metabolic charts that are sometimes posted on the walls outside professors' offices (Figure 10.3).

Individual metabolic pathways can take different forms. A linear metabolic pathway, such as the biosynthesis of serine, is a series of independent enzyme-catalyzed reactions



◂ Figure 10.3
Part of a large metabolic chart published by Roche Applied Science.



▲ **Figure 10.4**
Forms of metabolic pathways. (a) The biosynthesis of serine is an example of a linear metabolic pathway. The product of each step is the substrate for the next step. **(b)** The sequence of reactions in a cyclic pathway forms a closed loop. In the citric acid cycle, an acetyl group is metabolized via reactions that regenerate the intermediates of the cycle. **(c)** In fatty acid biosynthesis, a spiral pathway, the same set of enzymes catalyzes a progressive lengthening of the acyl chain.

in which the product of one reaction is the substrate for the next reaction in the pathway (Figure 10.4a). A cyclic metabolic pathway, such as the citric acid cycle, is also a sequence of enzyme-catalyzed steps, but the sequence forms a closed loop, so the intermediates are regenerated with every turn of the cycle (Figure 10.4b). In a spiral metabolic pathway, such as the biosynthesis of fatty acids (Section 16.6), the same set of enzymes is used repeatedly for lengthening or shortening a given molecule (Figure 10.4c).

Each type of pathway may have branch points where metabolites enter or leave. In most cases, we don't emphasize the branching nature of pathways because we want to focus on the main routes followed by the most important metabolites. We also want to focus on the pathways that are commonly found in all species. These are the most fundamental pathways. Don't be misled by this simplification. A quick glance at any metabolic chart will show that pathways have many branch points and that initial substrates and final products are often intermediates in other pathways. The serine pathway in Figure 10.3 is a good example. Can you find it?

B. Metabolism Proceeds by Discrete Steps

Intracellular environments don't change very much. Reactions proceed at moderate temperatures and pressures, at rather low reactant concentrations, and at close to neutral pH. We often refer to this as homeostasis at the cellular level.

These conditions require a multitude of efficient enzymatic catalysts. Why are so many distinct reactions carried out in living cells? In principle, it should be possible to carry out the degradation and the synthesis of complex organic molecules with far fewer reactions.

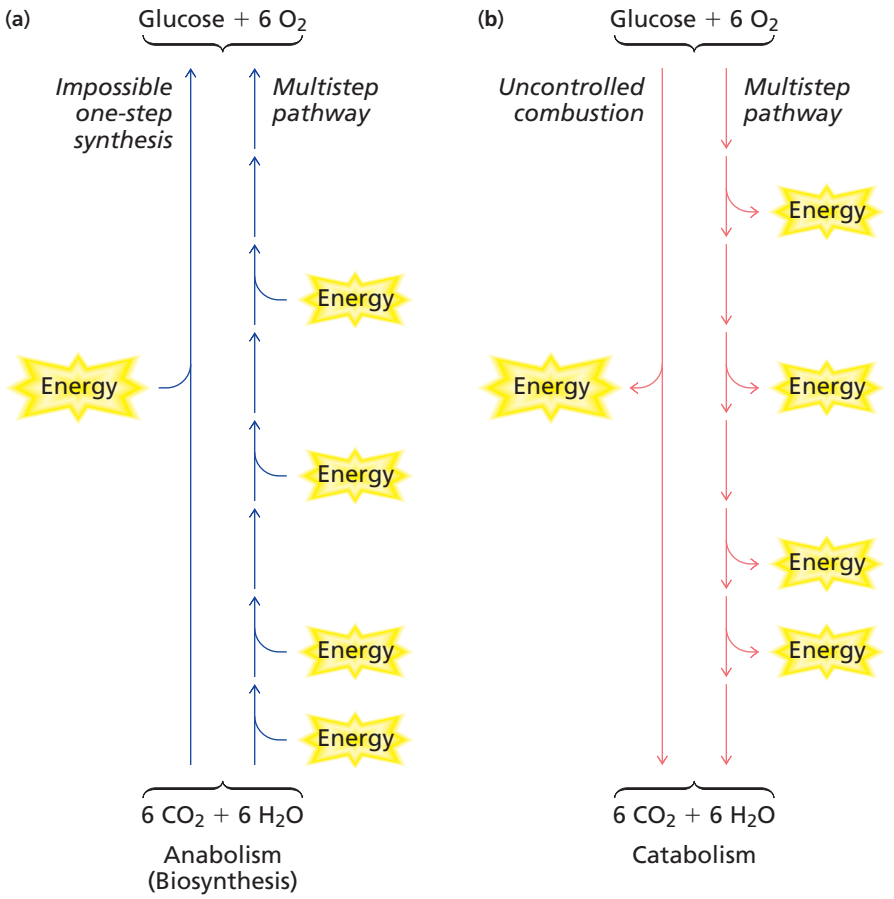
KEY CONCEPT

The limitations of chemistry and physics dictate that metabolic pathways consist of many small steps.

One reason for multistep pathways is the limited reaction specificity of enzymes. Each active site catalyzes only a single step of a pathway. The synthesis of a molecule—or its degradation—therefore follows a metabolic route defined by the availability of suitable enzymes. As a general rule, a single enzyme-catalyzed reaction can only break or form a few covalent bonds at a time. Often the reaction involves the transfer of a single chemical group. Thus, the large number of reactions and enzymes is due, in part, to the limitations of enzymes and chemistry.

Another reason for multiple steps in metabolic pathways is to control energy input and output. Energy flow is mediated by energy donors and acceptors that carry discrete quanta of energy. As we will see, the energy transferred in a single reaction seldom exceeds 60 kJ mol^{-1}. Pathways for the biosynthesis of molecules require the transfer of energy at multiple points. Each energy-requiring reaction corresponds to a single step in the reaction sequence.

The synthesis of glucose from carbon dioxide and water requires the input of ~2900 kJ mol^{-1} of energy. It is not thermodynamically possible to synthesize glucose in a single step (Figure 10.5). Similarly, much of the energy released during a catabolic process (such as the oxidation of glucose to carbon dioxide and water, which releases the same 2900 kJ mol^{-1}) is transferred to individual acceptors one step at a time rather



◀ **Figure 10.5**
Single-step versus multistep pathways. (a) The synthesis of glucose cannot be accomplished in a single step. Multistep synthesis is coupled to the input of small quanta of energy from ATP and NADH. **(b)** The uncontrolled combustion of glucose releases a large amount of energy all at once. A multistep enzyme-catalyzed pathway releases the same amount of energy but conserves much of it in a manageable form.

than being released in one grand, inefficient explosion. The efficiency of energy transfer at each step is never 100%, but a considerable percentage of the energy is conserved in manageable form. Energy carriers that accept and donate energy, such as adenine nucleotides (ATP) and nicotinamide coenzymes (NADH), are found in all life forms.

A major goal of learning about metabolism is to understand how these "quanta" of energy are used. ATP and NADH—and other coenzymes—are the "currency" of metabolism. This is why metabolism and bioenergetics are so closely linked.

C. Metabolic Pathways Are Regulated

Metabolism is highly regulated. Organisms react to changing environmental conditions such as the availability of energy or nutrients. Organisms also respond to genetically programmed instructions. For example, during embryogenesis or reproduction, the metabolism of individual cells can change dramatically.

The responses of organisms to changing conditions range from small changes to drastically reorganizing the metabolic processes that govern the synthesis or degradation of biomolecules and the generation or consumption of energy. Control processes can affect many pathways or only a few, and the response time can range from less than a second to hours or longer. The most rapid biological responses, occurring in milliseconds, include changes in the passage of small ions (e.g., $Na^{\oplus}$, $K^{\oplus}$, and $Ca^{2\oplus}$) through cell membranes. Transmission of nerve impulses and muscle contraction depend on ion movement. The most rapid responses are also the most short-lived; slower responses usually last longer.

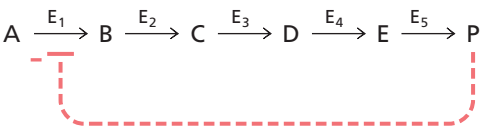
It is important to understand some basic concepts of pathways in order to see how they are regulated. Consider a simple linear pathway that begins with substrate A and ends with product P.

$$A \xrightleftharpoons{E_1} B \xrightleftharpoons{E_2} C \xrightleftharpoons{E_3} D \xrightleftharpoons{E_4} E \xrightleftharpoons{E_5} P \qquad (10.1)$$

Each of the reactions is catalyzed by an enzyme and they are all reversible. Most reactions in living cells have reached equilibrium so the concentrations of B, C, D, and E do not change very much. This is similar to the **steady state** condition we encountered in Section 5.3A. The steady state condition can be visualized by imagining a series of beakers of different sizes (Figure 10.6). Water flows into the first beaker from a tap and when it fills up the water spills over into another beaker. After filling up a series of beakers, there will be a steady flow of water from the tap onto the floor. The rate of flow is analogous to the **flux** through a metabolic pathway. The flux can vary from a trickle to a gusher but the steady state levels of water in each beaker don't change. (Unfortunately, this analogy doesn't allow us to see that in a metabolic pathway the flux could also be in the opposite direction.)

Flux through a metabolic pathway will decrease if the concentration of the initial substrate falls below a certain threshold. It will also decrease if the concentration of the final product rises. These are changes that affect all pathways. However, in addition to these normal concentration effects, there are special regulatory controls that affect the activity of particular enzymes in the pathway. It is tempting to visualize regulation of a pathway by the efficient manipulation of a single rate limiting enzymatic reaction, sometimes likened to the narrow part of an hourglass. In many cases, however, this is an oversimplification. Flux through most pathways depends on controls at several steps. These steps are special reactions in the pathways where the steady state concentrations of substrates and products are far from the equilibrium concentrations so the flux tends to go only in one direction. A regulatory enzyme contributes a particular degree of control over the overall flux of the pathway in which it participates. Because intermediates or cosubstrates from several sources can feed into or out of a pathway, the existence of multiple control points is normal; an isolated, linear, pathway is rare.

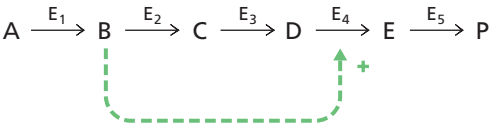
There are two common patterns of metabolic regulation: feedback inhibition and feed-forward activation. **Feedback inhibition** occurs when a product (usually the end product) of a pathway controls the rate of its own synthesis through inhibition of an early step, usually the first committed step (the first reaction that is unique to the pathway).



(10.2)

The advantage of such a regulatory pattern in a biosynthetic pathway is obvious. When the concentration of P rises above its steady state level, the effect is transmitted back through the pathway and the concentrations of each intermediate also rise. This causes flux to reverse in the pathway, leading to a net increase in the production of product A from reactant P. Flux in the normal direction is restored when P is depleted. The pathway is inhibited at an early step; otherwise, metabolic intermediates would accumulate unnecessarily. The important point in Reaction 10.2 is that the reaction catalyzed by enzyme E1 is not allowed to reach equilibrium. It is a metabolically irreversible reaction because the enzyme is regulated. Flux through this point is not allowed to go in the opposite direction.

Feed-forward activation occurs when a metabolite produced early in a pathway activates an enzyme that catalyzes a reaction further down the pathway.

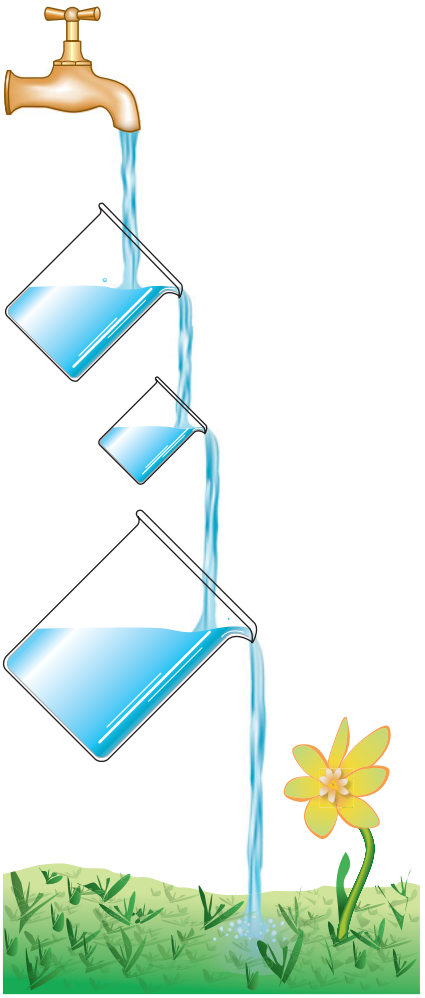


(10.3)

In this example, the activity of enzyme E_1 (which converts A to B) is coordinated with the activity of enzyme E_4 (which converts D to E). An increase in the concentration of metabolite B increases flux through the pathway by activating E4. (E4 would normally be inactive in low concentrations of B.)

In Section 5.10, we discussed the modulation of individual regulatory enzymes. Allosteric activators and inhibitors, which are usually metabolites, can rapidly alter the

The precise technical term for the condition where cellular pathways are not in a dynamic steady-state condition is . . . dead.



▲ **Figure 10.6**
Steady state and flux in a metabolic pathway. The rate of flow is equivalent to the flux in a pathway, and the constant amount of water in each beaker is analogous to the steady state concentrations of metabolites in a pathway.

activity of many of these enzymes by inducing conformational changes that affect catalytic activity. We will see many examples of allosteric modulation in the coming chapters. The allosteric modulation of regulatory enzymes is fast but not as rapid in cells as it can be with isolated enzymes.

The activity of interconvertible enzymes can also be rapidly and reversibly altered by covalent modification, commonly by the addition and removal of phosphoryl groups as described in Section 5.9D. Recall that phosphorylation, catalyzed by protein kinases at the expense of ATP, is reversed by the action of protein phosphatases, which catalyze the hydrolytic removal of phosphoryl groups. Individual enzymes differ in whether their response to phosphorylation is activation or deactivation. Interconvertible enzymes in catabolic pathways are generally activated by phosphorylation and deactivated by dephosphorylation; most interconvertible enzymes in anabolic pathways are inactivated by phosphorylation and reactivated by dephosphorylation. The activation of kinases with multiple specificities allows coordinated regulation of more than one metabolic pathway by one signal. The cascade nature of intracellular signaling pathways, described in Section 9.12, also means that the initial signal is amplified (Figure 10.7).

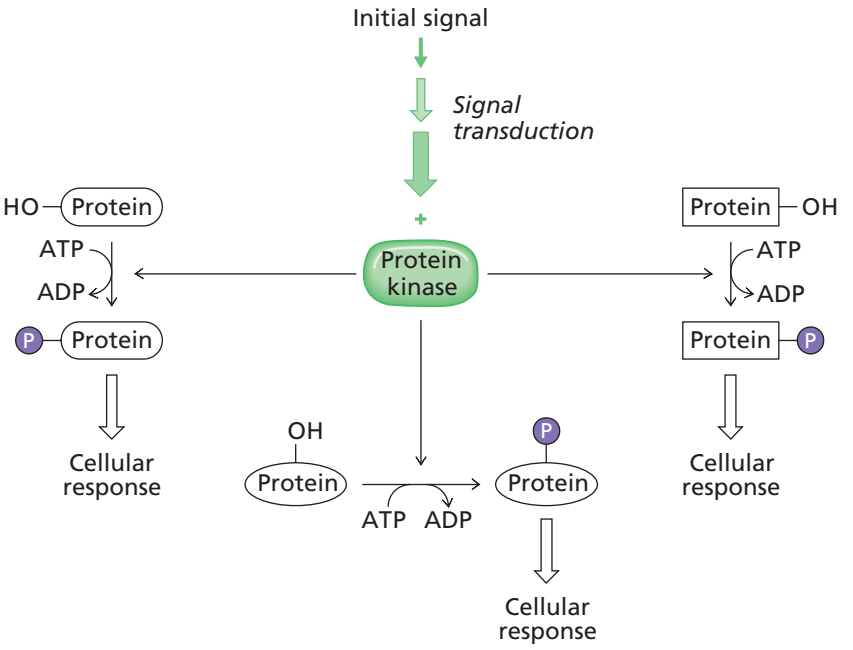
The amounts of specific enzymes can be altered by increasing the rates of specific protein synthesis or degradation. This is usually a slow process relative to allosteric or covalent activation and inhibition. However, the turnover of certain enzymes may be rapid. Keep in mind that several modes of regulation can operate simultaneously within a metabolic pathway.

In Part 4 of this book, we examine more closely the regulation of gene expression and protein synthesis.

D. Evolution of Metabolic Pathways

The evolution of metabolic pathways is an active area of biochemical research. These studies have been greatly facilitated by the publication of hundreds of complete genome sequences, especially prokaryotic genomes. Biochemists can now compare pathway enzymes in a number of species that show a diverse variety of pathways. Many of these pathways provide clues to the organization and structure of the primitive pathways that were present in the first cells.

There are many possible routes to the formation of a new metabolic pathway. The simplest case is the addition of a new terminal step to a preexisting pathway. Consider the hypothetical pathway in Equation 10.1. The original pathway might have terminated with the production of metabolite E after a four-step transformation from substrate A. The availability of substantial quantities of metabolite E might favor the evolution of a new enzyme (E_5 in this case) that could use E as a substrate to make P. The pathways



◀ **Figure 10.7**
Regulatory role of a protein kinase. The effect of the initial signal is amplified by the signaling cascade. Phosphorylation of different cellular proteins by the activated kinase results in coordinated regulation of different metabolic pathways. Some pathways may be activated, whereas others are inhibited. —P represents a protein-bound phosphate group.

leading to synthesis of asparagine and glutamine from aspartate and glutamate pathways are examples of this type of pathway evolution. This *forward evolution* is thought to be a common mechanism of evolution of new pathways.

In other cases, a new pathway can form by evolving a branch to a preexisting pathway. For example, consider the conversion of C to D in the Equation 10.1 pathway. This reaction is catalyzed by enzyme E_3. The primitive E_3 enzyme might not have been as specific as the modern enzyme. In addition to producing product D, it might have synthesized a smaller amount of another metabolite, X. The availability of product X might have conferred some selective advantage to the cell favoring a duplication of the E_3 gene. Subsequent divergence of the two copies of the gene gave rise to two related enzymes that specifically catalyzed C → D and C → X. There are many examples of *evolution by gene duplication and divergence* (e.g., lactate dehydrogenase and malate dehydrogenase, Section 4.7). (We have mostly emphasized the extreme specificity of enzyme reactions but, in fact, many enzymes can catalyze several different reactions using structurally similar substrates and products.)

Some pathways might have evolved "backwards." A primitive pathway might have utilized an abundant supply of metabolite E in the environment in order to make product P. As the supply of E became depleted over time there was selective pressure to evolve a new enzyme (E_4) that could make use of metabolite D to replenish metabolite E. When D became rate limiting, cells could gain a selective advantage by utilizing C to make more metabolite D. In this way the complete modern pathway evolved by *retroevolution*, successively adding simpler precursors and extending the pathway.

Sometimes an entire pathway can be duplicated and subsequent adaptive evolution leads to two independent pathways with homologous enzymes that catalyze related reactions. There is good evidence that the pathways leading to biosynthesis of tryptophan and histidine evolved in this manner. Enzymes can also be recruited from one pathway for use in another without necessarily duplicating an entire pathway. We'll encounter several examples of homologous enzymes that are used in different pathways.

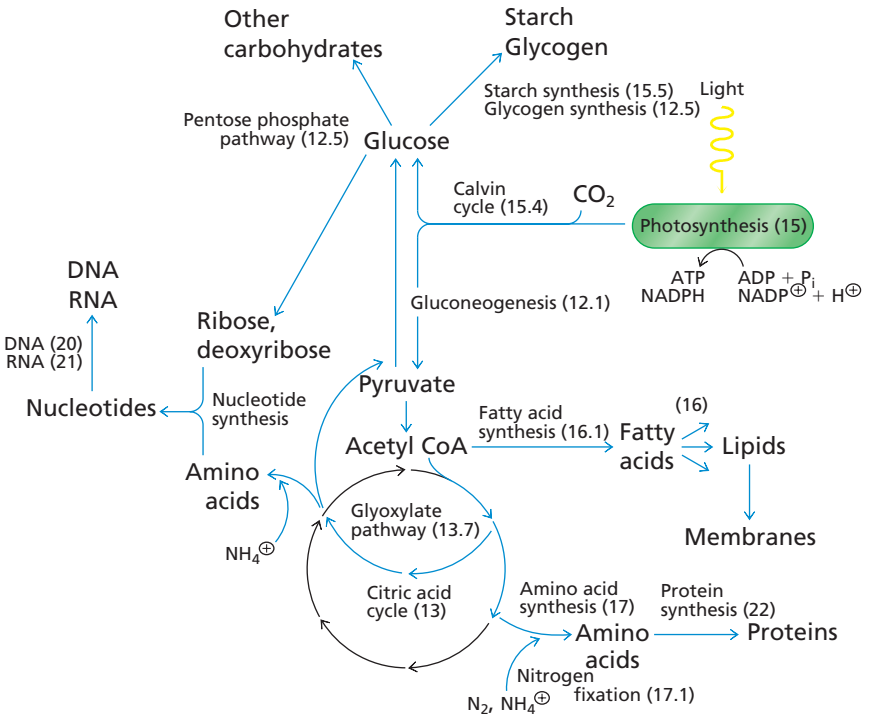
Finally, a new pathway can evolve by "reversing" an existing pathway. In most cases, there is one step in a pathway that is essentially irreversible. Let's assume that the third step in our hypothetical pathway (C → D) is unable to catalyze the conversion of D to C because the normal reaction is far from equilibrium. The evolution of a new enzyme that can catalyze D → C would allow this entire pathway to reverse direction, converting P to A. This is how the glycolysis pathway evolved from the glucose biosynthesis (gluconeogenesis) pathway. There are many other examples of *evolution by pathway reversal*.

All of these possibilities play a role in the evolution of new pathways. Sometimes a new pathway evolves by a combination of different mechanisms of adaptive evolution. The evolution of the citric acid cycle pathway, which took place several billion years ago, is an example (Section 12.9). New metabolic pathways are evolving all the time in response to pesticides, herbicides, antibiotics, and industrial waste. Organisms that can metabolize these compounds, thus escaping their toxic effects, have evolved new pathways and enzymes by modifying existing ones.

10.3 Major Pathways in Cells

This section provides an overview of the organization and function of some central metabolic pathways that are discussed in subsequent chapters. We begin with the anabolic, or biosynthetic, pathways since these pathways are the most important for growth and reproduction. A general outline of biosynthetic pathways is shown in Figure 10.8. All cells require an external source of carbon, hydrogen, oxygen, nitrogen, phosphorus, and sulfur plus additional inorganic ions (Section 1.2). Some species, notably bacteria and plants, can grow and reproduce by utilizing inorganic sources of these essential elements. These species are called **autotrophs**. There are two distinct categories of autotrophic species. **Heterotrophs**, such as animals, need an organic carbon source (e.g., glucose).

Biosynthetic pathways require energy. The most complex organisms (from a biochemical perspective!) can generate useful metabolic energy from sunlight or by oxidizing inorganic molecules such as $NH_4^{\oplus}$, H_2, or H_2S. The energy from these reactions is



◀ **Figure 10.8**
Overview of anabolic pathways. Large molecules are synthesized from smaller ones by adding carbon (usually in the form of CO_2) and nitrogen (usually as NH_4^+). The main pathways include the citric acid cycle, which supplies the intermediates in amino acid biosynthesis, and gluconeogenesis, which results in the production of glucose. The energy for biosynthetic pathways is supplied by light in photosynthetic organisms or by the breakdown of inorganic molecules in other autotrophs. (Numbers in parentheses refer to the chapters and sections of this book.)

used to synthesize the energy-rich compound ATP and the reducing power of NADH. These cofactors transfer their energy to biosynthetic reactions.

There are two types of autotrophic species. **Photoautotrophs** obtain most of their energy by photosynthesis and their main source of carbon is CO_2. This category includes photosynthetic bacteria, algae, and plants. **Chemoautotrophs** obtain their energy by oxidizing inorganic molecules and utilizing CO_2 as a carbon source. Some bacterial species are chemoautotrophs but there are no eukaryotic examples.

Heterotrophs can be split into two categories. **Photoheterotrophs** are photosynthetic organisms that require an organic compound as a carbon source. There are several groups of bacteria that are capable of capturing light energy but must rely on some organic molecules as a carbon source. **Chemoheterotrophs** are nonphotosynthetic organisms that require organic molecules as carbon sources. Their metabolic energy is usually derived from the breakdown of the imported organic molecules. We are chemoheterotrophs, as are all animals, most protists, all fungi, and many bacteria.

The main catabolic pathways are shown in Figure 10.9. As a general rule, these degradative pathways are not simply the reverse of biosynthesis pathways. Note that the citric acid cycle is a major pathway in both anabolic and catabolic metabolism. The main roles of catabolism are to eliminate unwanted molecules and to generate energy for use in other processes.

We will examine metabolism in the next few chapters. Our discussion of metabolic pathways begins in Chapter 11 with glycolysis, a ubiquitous pathway for glucose catabolism. There is a long-standing tradition in biochemistry of introducing students to glycolysis before any other pathways are encountered. We know a great deal about the reactions in this pathway and they will illustrate many of the fundamental principles of biochemistry. In glycolysis, the hexose is split into two three-carbon metabolites. This pathway can generate ATP in a process called *substrate level phosphorylation*. Often, the product of glycolysis is pyruvate, which can be converted to acetyl CoA for further oxidation.

Chapter 12 describes the synthesis of glucose, or gluconeogenesis. This chapter also covers starch and glycogen metabolism and outlines the pathway by which glucose is oxidized to produce NADPH for biosynthetic pathways and ribose for the synthesis of nucleotides.

The citric acid cycle (Chapter 13) facilitates complete oxidation of the acetate carbons of acetyl CoA to carbon dioxide. The energy released from this oxidation is conserved in



▲ **Chemoautotrophs in Yellowstone National Park.** There are many species of *Thiobacillus* that derive their energy from the oxidation of iron or sulfur. They do not require any organic molecules. The orange and yellow colors surrounding this hot spring in Yellowstone National Park are due to the presence of *Thiobacillus*. See Chapter 14 for an explanation of how such organisms generate energy from inorganic molecules.

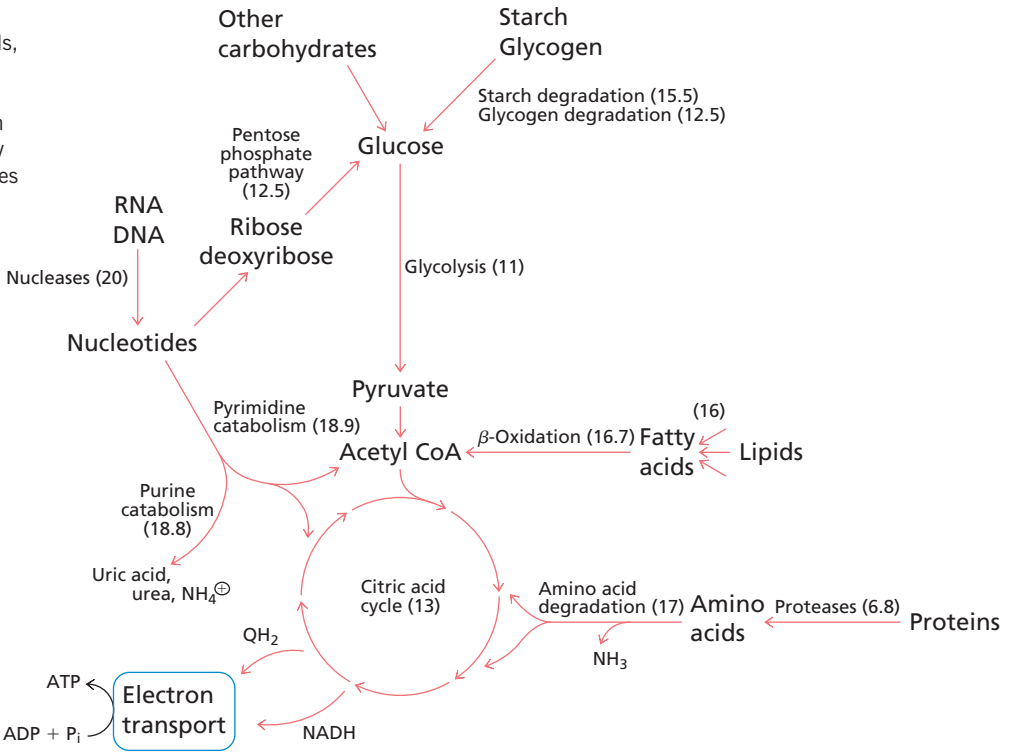


Figure 10.9 ▶
Overview of catabolic pathways. Amino acids, nucleotides, monosaccharides, and fatty acids are formed by enzymatic hydrolysis of their respective polymers. They are then degraded in oxidative reactions and energy is conserved in ATP and reduced coenzymes (mostly NADH). (Numbers in parentheses refer to the chapters and sections of this book.)

the formation of NADH and ATP. As mentioned above, the citric acid cycle is an essential part of both anabolic and catabolic metabolism.

The production of ATP is one of the most important reactions in metabolism. The synthesis of most ATP is coupled to membrane-associated electron transport (Chapter 14). In electron transport, the energy of reduced coenzymes such as NADH is used to generate an electrochemical gradient of protons across a cell membrane. The potential energy of this gradient is harnessed to drive the phosphorylation of ADP to ATP.

$$ADP + P_i \longrightarrow ATP + H_2O \qquad \textbf{(10.4)}$$

We will see that the reactions of membrane-associated electron transport and coupled ATP synthesis are similar in many ways to the reactions that capture light energy during photosynthesis (Chapter 15).

Three additional chapters examine the anabolism and catabolism of lipids, amino acids, and nucleotides. Chapter 16 discusses the storage of nutrient material as triacylglycerols and the subsequent oxidation of fatty acids. This chapter also describes the synthesis of phospholipids and isoprenoid compounds. Amino acid metabolism is discussed in Chapter 17. Although amino acids were introduced as the building blocks of proteins, some also play important roles as metabolic fuels and biosynthetic precursors. Nucleotide biosynthesis and degradation are considered in Chapter 18. Unlike the other three classes of biomolecules, nucleotides are catabolized primarily for excretion rather than for energy production. The incorporation of nucleotides into nucleic acids and of amino acids into proteins are major anabolic pathways. Chapters 20 to 22 describe these biosynthetic reactions.

10.4 Compartmentation and Interorgan Metabolism

Some metabolic pathways are localized to particular regions within a cell. For example, the pathway of membrane-associated electron transport coupled to ATP synthesis takes place within the membrane. In bacteria this pathway is located in the plasma membrane and in eukaryotes it is found in the mitochondrial membrane. Photosynthesis is another example of a membrane-associated pathway in bacteria and eukaryotes.

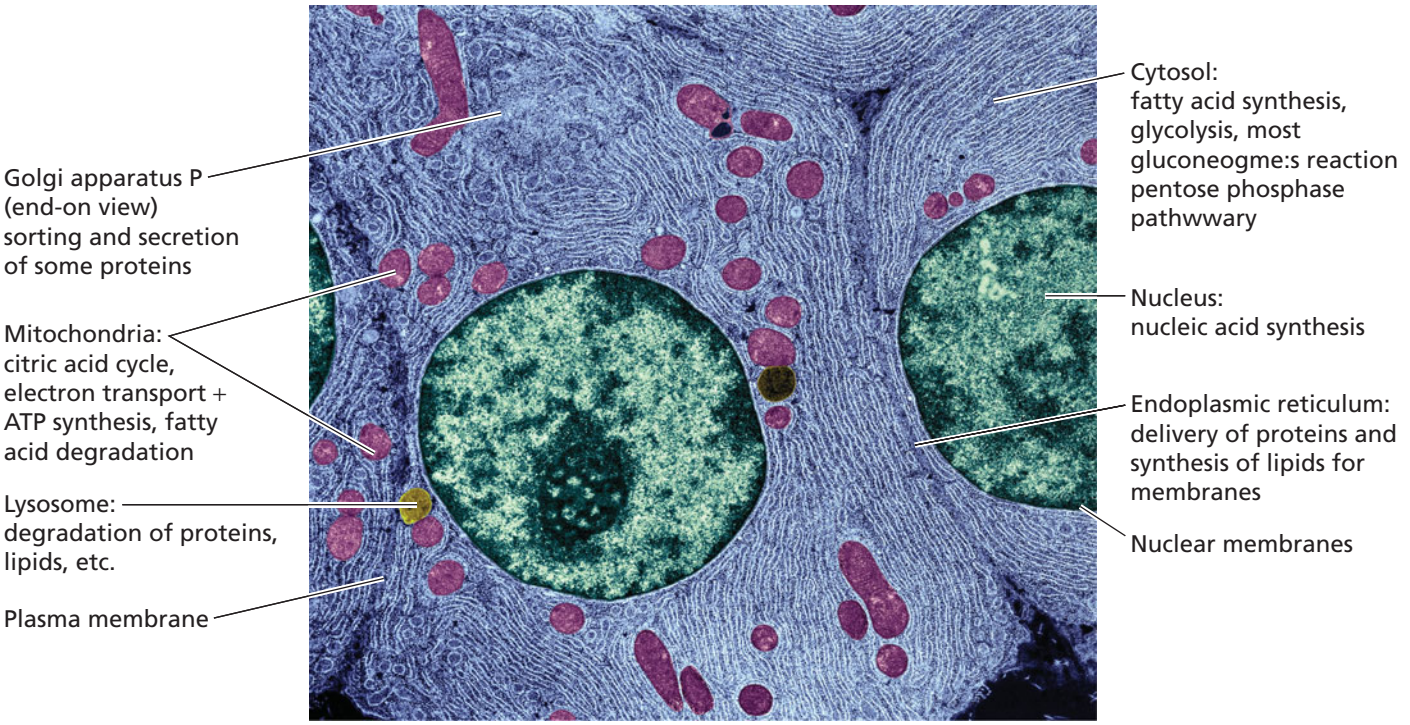


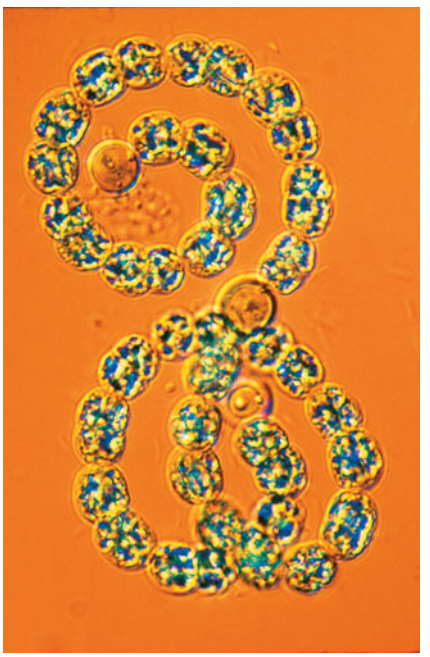
Figure 10.10 ▲
Compartmentation of metabolic processes within a eukaryotic cell. This is a colored electron micrograph of a cell showing the nucleus (green), mitochondria (purple), lysosomes (brown), and extensive endoplasmic reticulum (blue). (Not all pathways and organelles are shown.)

In eukaryotes, metabolic pathways are localized within several membrane-bound compartments (Figure 10.10). For example, the enzymes that catalyze fatty acid synthesis are located in the cytosol, whereas the enzymes that catalyze fatty acid breakdown are located inside mitochondria. One consequence of compartmentation is that separate pools of metabolites can be found within a cell. This arrangement permits the simultaneous operation of opposing metabolic pathways. Compartmentation can also offer the advantage of high local concentrations of metabolites and coordinated regulation of enzymes. Some of the enzymes that catalyze reactions in mitochondria (which have evolved from a symbiotic prokaryote) are encoded by mitochondrial genes; this origin explains their compartmentation.

There is also compartmentation at the molecular level. Enzymes that catalyze some pathways are physically organized into multienzyme complexes (Section 5.11). With these complexes, channeling of metabolites prevents their dilution by diffusion. Some enzymes catalyzing adjacent reactions in pathways are bound to membranes and can diffuse rapidly in the membrane for interaction.

Individual cells of multicellular organisms maintain different concentrations of metabolites, depending in part on the presence of specific transporters that facilitate the entry and exit of metabolites. In addition, depending on the cell-surface receptors and signal-transduction mechanisms present, individual cells respond differently to external signals.

In multicellular organisms, compartmentation can also take the form of specialization of tissues. The division of labor among tissues allows site-specific regulation of metabolic processes. Cells from different tissues are distinguished by their complement of enzymes. We are very familiar with the specialized role of muscle tissue, red blood cells, and brain cells but cell compartmentation is a common feature even in simple species. In cyanobacteria, for example, the pathway for nitrogen fixation is sequestered in special cells called heterocysts (Figure 10.11). This separation is necessary because nitrogenase is inactivated by oxygen and the cells that carry out photosynthesis produce lots of oxygen.



▲ Figure 10.11
Anabaena spherica. Many species of cyanobacteria form long, multicellular filaments. Some specialized cells have adapted to carry out nitrogen fixation. These heterocysts have become rounded and are surrounded by a thickened cell wall. The heterocysts are connected to adjacent cells by internal pores. The formation of heterocysts is an example of compartmentation of metabolic pathways.

10.5 Actual Gibbs Free Energy Change, Not Standard Free Energy Change, Determines the Direction of Metabolic Reactions

The Gibbs free energy change is a measure of the energy available from a reaction (Section 1.4B). The *standard* Gibbs free energy change for any given reaction ($\Delta G^{\circ\prime}_{\text{reaction}}$) is the change under standard conditions of pressure (1 atm), temperature (25°C = 298 K), and hydrogen ion concentration (pH = 7.0). The concentration of every reactant and product is 1 M under standard conditions. For biochemical reactions, the concentration of water is assumed to be 55 M.

The standard Gibbs free energy change in a reaction can be determined by using tables that list the Gibbs free energies of formation ($\Delta_f G^{\circ\prime}$) of important biochemical molecules.

$$\Delta G^{\circ\prime}_{\text{reaction}} = \Delta_f G^{\circ\prime}_{\text{products}} - \Delta_f G^{\circ\prime}_{\text{reactants}} \qquad \textbf{(10.5)}$$

Keep in mind that Equation 10.5 only applies to the free energy change under standard conditions where the concentrations of products and reactants are 1 M. It's also important to use tables that apply to biochemical reactions. These tables correct for pH and ionic strength. The Gibbs free energies of formation under cellular conditions are often quite different from the ones used in chemistry and physics.

The *actual* Gibbs free energy change (ΔG) for a reaction depends on the real concentrations of reactants and products, as described in Section 1.4B. The relationship between the standard free energy change and the actual free energy change is given by

$$\Delta G_{\text{reaction}} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln \frac{[\text{products}]}{[\text{reactants}]} \qquad \textbf{(10.6)}$$

For a chemical or physical process, the free energy change is expressed in terms of the changes in enthalpy (heat content) and entropy (randomness) as the reactants are converted to products at constant pressure and volume.

$$\Delta G = \Delta H - T\Delta S \qquad \textbf{(10.7)}$$

ΔH is the change in enthalpy, ΔS is the change in entropy, and T is the temperature in degrees Kelvin.

When ΔG for a reaction is negative, the reaction will proceed in the direction it is written. When ΔG is positive, the reaction will proceed in the reverse direction—there will be a net conversion of products to reactants. For such a reaction to proceed in the direction written, enough energy must be supplied from outside the system to make the free energy change negative. When ΔG is zero, the reaction is at equilibrium and there is no net synthesis of product.

Because changes in both enthalpy and entropy contribute to ΔG, the sum of these contributions at a given temperature (as indicated in Equation 10.7) must be negative for a reaction to proceed. Thus, even if ΔS for a particular process is negative (i.e., the products are more ordered than the reactants), a sufficiently negative ΔH can overcome the decrease in entropy, resulting in a ΔG that is less than zero. Similarly, even if ΔH is positive (i.e., the products have a higher heat content than the reactants), a sufficiently positive ΔS can overcome the increase in enthalpy, resulting in a negative ΔG. Reactions that proceed because of a large positive ΔS are said to be entropy driven. Examples of entropy-driven processes include protein folding (Section 4.10) and the formation of lipid bilayers (Section 9.8A), both of which depend on the hydrophobic effect (Section 2.5D). The processes of protein folding and lipid-bilayer formation result in states of decreased entropy for the protein molecule and bilayer components, respectively. However, the decrease in entropy is offset by a large increase in the entropy of surrounding water molecules.

For any enzymatic reaction within a living organism, the actual free energy change (the free energy change under cellular conditions) must be less than zero in order for

the reaction to occur in the direction it is written. Many metabolic reactions have *standard* Gibbs free energy changes ($\Delta G^{\circ\prime}_{\text{reaction}}$) that are positive. The difference between ΔG and $\Delta G^{\circ\prime}$ depends on cellular conditions. The most important condition affecting free energy change in cells is the concentrations of substrates and products of a reaction. Consider the reaction

$$A + B \rightleftharpoons C + D \qquad (10.8)$$

At equilibrium, the ratio of substrates and products is by definition the equilibrium constant (K_{eq}) and the Gibbs free energy change under these conditions is zero.

$$\text{(at equilibrium)} \qquad K_{eq} = \frac{[C][D]}{[A][B]} \quad \Delta G = 0 \qquad (10.9)$$

When this reaction is not at equilibrium, a different ratio of products to substrates is observed and the Gibbs free energy change is derived using Equation 10.6.

$$\Delta G_{\text{reaction}} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln \frac{[C][D]}{[A][B]} = \Delta G^{\circ\prime}_{\text{reaction}} + RT \ln Q$$
$$\left(\text{where } Q = \frac{[C][D]}{[A][B]}\right) \qquad (10.10)$$

Q is the **mass action ratio**. The difference between this ratio and the ratio of products to substrates at equilibrium determines the actual Gibbs free energy change for a reaction. In other words, the free energy change is a measure of how far from equilibrium the reacting system is operating. Consequently, ΔG, not $\Delta G^{\circ\prime}$, is the criterion for assessing the direction of a reaction in a biological system.

We can divide metabolic reactions into two types. Let Q represent the steady-state ratio of product and reactant concentrations in a living cell. Reactions for which Q is close to K_{eq} are called **near-equilibrium reactions**. The free energy changes associated with near-equilibrium reactions are small, so these reactions are readily reversible. Reactions for which Q is far from K_{eq} are called **metabolically irreversible reactions**. These reactions are greatly displaced from equilibrium, with Q usually differing from K_{eq} by two or more orders of magnitude. Thus, ΔG is a large negative number for metabolically irreversible reactions.

When flux through a pathway changes by a large amount, there may be short-term perturbations of metabolite concentrations in the pathway. The intracellular concentrations of metabolites vary, but usually over a range of not more than two- or threefold and equilibrium is quickly restored. As mentioned above, this is called the steady state condition and it's typical of most of the reactions in a pathway. Most enzymes in a pathway catalyze near-equilibrium reactions and have sufficient activity to quickly restore concentrations of substrates and products to near-equilibrium conditions. They can accommodate flux in either direction. The Gibbs free energy change for these reactions is effectively zero.

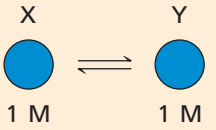
In contrast, the activities of enzymes that catalyze metabolically irreversible reactions are usually insufficient to achieve near-equilibrium status for the reactions. Metabolically irreversible reactions are generally the control points of pathways, and the enzymes that catalyze these reactions are usually regulated in some way. In fact, the regulation maintains metabolic irreversibility by preventing the reaction from reaching equilibrium. Metabolically irreversible reactions can act as bottlenecks in metabolic traffic, helping control the flux through reactions further along the pathway.

Near-equilibrium reactions are not usually suitable control points. Flux through a near-equilibrium step cannot be significantly increased since it is already operating under conditions where the concentrations of products and reactants are close to the equilibrium values. The direction of near-equilibrium reactions can be controlled by changes in substrate and product concentrations. In contrast, flux through metabolically irreversible reactions is relatively unaffected by changes in metabolite concentration; flux through these reactions must be controlled by modulating the activity of the enzyme.

KEY CONCEPT

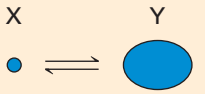
Metabolically irreversible reactions are catalyzed by enzymes whose activity is regulated in order to prevent the reaction from reaching equilibrium.

Consider a sample reaction X = Y under standard conditions of pressure, temperature, and concentration. Assume that $\Delta G^{\circ\prime}$ is negative.



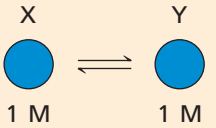
$\Delta G^{\circ\prime}$ negative

Inside the cell, the reaction will likely be at equilibrium and $\Delta G = 0$



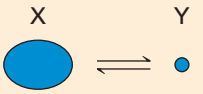
$\Delta G = 0$
($\Delta G^{\circ\prime}$ negative)

For a reaction in which $\Delta G^{\circ\prime}$ is positive,



$\Delta G^{\circ\prime}$ positive

at equilibrium, the concentration of reactant will be higher than that of the product.



$\Delta G = 0$
($\Delta G^{\circ\prime}$ positive)

The standard Gibbs free energy change does not predict whether a reaction will proceed in one direction or another. Instead, it predicts the steady state concentrations of reactants and products in near-equilibrium reactions.

Because so many metabolic reactions are near-equilibrium reactions, we have chosen not to emphasize $\Delta G°'$ values in our discussions of most reactions. Those values are not relevant except when they are used to calculate steady state concentrations.

SAMPLE CALCULATION 10.1 Calculating Standard Gibbs Free Energy Change from Energies of Formation

For any reaction, the standard Gibbs free energy change for the reaction is given by

$$\Delta G°'_{\text{reaction}} = \Delta_f G°'_{\text{products}} - \Delta_f G°'_{\text{reactants}}$$

For the oxidation of glucose,

$$(CH_2O)_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

you obtain the standard Gibbs free energies of formation from biochemical tables.

$\Delta_f G°'(\text{glucose}) = -426 \text{ kJ mol}^{-1}$
$\Delta_f G°'(O_2) = 0$
$\Delta_f G°'(CO_2) = -394 \text{ kJ mol}^{-1}$
$\Delta_f G°'(H_2O) = -156 \text{ kJ mol}^{-1}$
$\Delta G°'_{\text{reaction}} = 6(-394) + 6(-156) - (-426)$
$= -2874 \text{ kJ mol}^{-1}$

Glucose is an energy-rich organic molecule and its oxidation releases a great deal of energy. Nevertheless, all living cells routinely synthesize glucose from simple precursors. In many cases, the precursors are CO_2 and H_2O in the reverse of the reaction shown here. How do they do it?

Section 7.2 A described the structure and functions of nucleoside triphosphates.

Another example of the role of pyrophosphate is discussed in Section 10.7C. Hydrolysis of pyrophosphate is often counted as one ATP equivalent in terms of energy currency.

10.6 The Free Energy of ATP Hydrolysis

ATP contains one phosphate ester formed by linkage of the α-phosphoryl group to the 5′-oxygen of ribose and two phosphoanhydrides formed by the α,β and β,γ linkages between phosphoryl groups (Figure 10.12). ATP is a donor of several metabolic groups, usually a phosphoryl group, leaving ADP, or an AMP group, leaving inorganic pyrophosphate (PPi). Both reactions require the cleavage of a phosphoanhydride linkage. Although the various groups of ATP are not transferred directly to water, hydrolytic reactions provide useful estimates of the Gibbs free energy changes involved. Table 10.1 lists the free energies of formation of the various reactants and products under standard conditions, 1 mM Mg^{2+}, and an ionic strength of 0.25 M. Table 10.2 lists the standard Gibbs free energies of hydrolysis ($\Delta G°'_{\text{hydrolysis}}$) for ATP and AMP, and Figure 10.9 shows the hydrolytic cleavage of each of the phosphoanhydrides of ATP. Note from Table 10.2 that cleavage of the ester releases only 13 kJ mol^{-1} under standard conditions but cleavage of either of the phosphoanhydrides releases at least 30 kJ mol^{-1} under standard conditions.

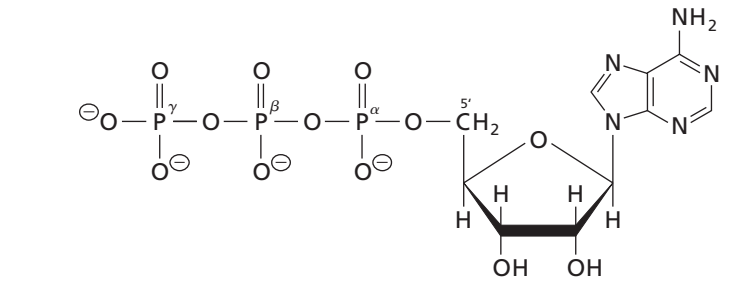
Table 10.1 Free Energies of Formation ($\Delta_f G°'$)

	kJ mol^{-1}
ATP	−2102
ADP	−1231
AMP	−360
P_i	−1059
H_2O	−156

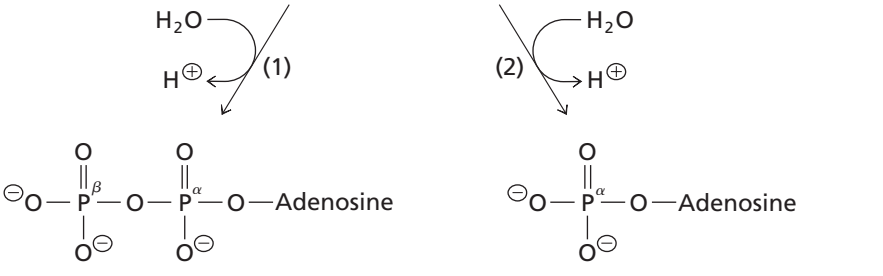
(1 mM Mg^{2+}, ionic strength of 0.25 M)

Table 10.2 also gives the standard Gibbs free energy change for hydrolysis of pyrophosphate. All cells contain an enzyme called pyrophosphatase that catalyzes this reaction. The cellular concentration of pyrophosphate is maintained at a very low concentration as a consequence of this highly favorable reaction. This means that the hydrolysis of ATP to AMP + pyrophosphate will always be associated with a negative Gibbs free energy change even when the AMP concentration is significant.

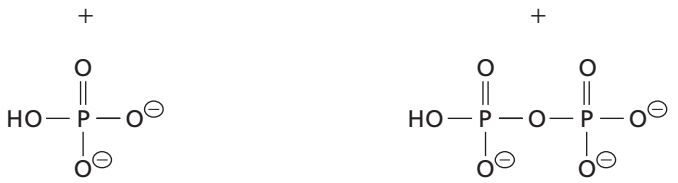
Nucleoside diphosphates and triphosphates in both aqueous solution and at the active sites of enzymes are usually present as complexes with magnesium (or sometimes manganese) ions. These cations coordinate with oxygen atoms of the phosphate groups, forming six-membered rings. A magnesium ion can form several different complexes with ATP; the complexes involving the α and β and the β and γ phosphate groups are shown in Figure 10.13. Formation of the β,γ complex is favored in aqueous solutions. We will see later that nucleic acids are also usually complexed with counterions such as



Adenosine 5′ -triphosphate (ATP^{4-})



Adenosine 5′-diphosphate (ADP^{3-}) Adenosine 5′-monophosphate (AMP^{2-})



Inorganic phosphate (P_i) Inorganic pyrophosphate (PP_i)

◂ **Figure 10.12**
Hydrolysis of ATP to (1) ADP and inorganic phosphate (P_i) and (2) AMP and inorganic pyrophosphate (PP_i).

The release of a free proton in these reactions depends on the conditions since the pKa values of the various components are close to the value inside cells (see Figure 2.19).

Mg^{2+} or cationic proteins. For convenience, we usually refer to the nucleoside triphosphates as adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), and uridine triphosphate (UTP), but remember that these molecules actually exist as complexes with Mg^{2+} in cells.

Several factors contribute to the large amount of energy released during hydrolysis of the phosphoanhydride linkages of ATP.

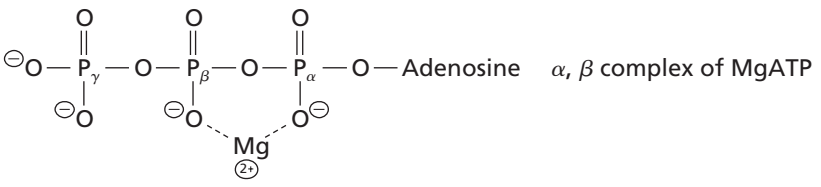
1. **Electrostatic repulsion.** Electrostatic repulsion among the negatively charged oxygen atoms of the phosphoanhydride groups of ATP is less after hydrolysis. [In cells, $\Delta G^{\circ\prime}_{\text{hydrolysis}}$ is actually increased (made more positive) by the presence of Mg^{2+}, which partially neutralizes the charges on the oxygen atoms of ATP and diminishes electrostatic repulsion.]
2. **Solvation effects.** The products of hydrolysis, ADP and inorganic phosphate, or AMP and inorganic pyrophosphate, are better solvated than ATP itself. When ions

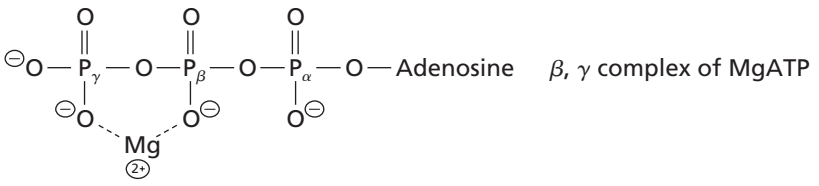
Table 10.2 Standard Gibbs free energies of hydrolysis for ATP, AMP, and pyrophosphate

Reactants and products	$\Delta G^{\circ\prime}_{\text{hydrolysis}}$ (kJ mol^{-1})
$ATP + H_2O \rightarrow ADP + P_i + H^{\oplus}$	−32
$ATP + H_2O \rightarrow AMP + PP_i + H^{\oplus}$	−45
$AMP + H_2O \rightarrow Adenosine + P_i + H^{\oplus}$	−13
$PP_i + H_2O \rightarrow 2P_i$	−29

P_i(inorganic phosphate) = HPO_4^{2-}

PP_i(pyrophosphate) = $HP_2O_7^{3-}$





◂ **Figure 10.13**
Complexes between ATP and Mg^{2+}.

are solvated, they are electrically shielded from each other. Solvation effects are probably the most important factor contributing to the energy of hydrolysis.

3. **Resonance stabilization.** The products of hydrolysis are more stable than ATP. The electrons on terminal oxygen atoms are more delocalized than those on bridging oxygen atoms. Hydrolysis of ATP replaces one bridging oxygen atom with two new terminal oxygen atoms.

A quantitative definition of a "high energy" compound is presented in Section 10.7A.

Because of the free energy change associated with the cleavage of their phosphoanhydrides, ATP and the other nucleoside triphosphates (UTP, GTP, and CTP) are often referred to as **energy-rich compounds**, but keep in mind that it's the system, not the molecule, that contributes free energy to biochemical reactions. ATP, by itself, is not really a high energy compound. It can only work if the system (reactants and products) is far from equilibrium. The ATP currency becomes worthless if the reaction reaches equilibrium and $\Delta G = 0$. We will find it useful to refer to "energy-rich" or "**high energy**" molecules in the jargon of biochemistry but we will put the terms in quotation marks to remind you that it is jargon.

KEY CONCEPT

The large free energy change associated with hydrolysis of ATP is only possible if the system is far from equilibrium.

All the phosphoanhydrides of nucleoside triphosphates have nearly equal standard Gibbs free energies of hydrolysis. We occasionally express the consumption or formation of the phosphoanhydride linkages of nucleoside triphosphates in terms of ATP equivalents.

ATP is usually the phosphoryl group donor when nucleoside monophosphates and diphosphates are phosphorylated. Of course, the intracellular concentrations of individual nucleoside mono-, di-, and triphosphates differ, depending on metabolic needs. For example, the intracellular levels of ATP are far greater than deoxythymidine triphosphate (dTTP) levels. ATP is involved in many reactions, whereas dTTP has fewer functions and is primarily a substrate for DNA synthesis.

A series of kinases (phosphotransferases) catalyze interconversions of nucleoside mono-, di-, and triphosphates. Phosphoryl group transfers between nucleoside phosphates have equilibrium constants close to 1.0. Nucleoside monophosphate kinases are a group of enzymes that catalyze the conversion of nucleoside monophosphates to nucleoside diphosphates. For example, guanosine monophosphate (GMP) is converted to guanosine diphosphate (GDP) by the action of guanylate kinase. GMP or its deoxy analog dGMP is the phosphoryl group acceptor in the reaction, and ATP or dATP is the phosphoryl group donor.

$$\text{GMP} + \text{ATP} \rightleftharpoons \text{GDP} + \text{ADP} \qquad (10.11)$$

Nucleoside diphosphate kinase acts in the conversion of nucleoside diphosphates to nucleoside triphosphates. This enzyme, present in both the cytosol and mitochondria of eukaryotes, is much less specific than nucleoside monophosphate kinases. All nucleoside diphosphates, regardless of the purine or pyrimidine base, are substrates for nucleoside diphosphate kinase. Nucleoside monophosphates are not substrates. Because of its relative abundance, ATP is usually the phosphoryl-group donor in cells:

$$\text{GDP} + \text{ATP} \rightleftharpoons \text{GTP} + \text{ADP} \qquad (10.12)$$

Although the concentration of ATP varies among cell types, the intracellular ATP concentration fluctuates very little within a particular cell, and the sum of the concentrations of the adenine nucleotides remains nearly constant. Intracellular ATP concentrations are maintained in part by the action of adenylate kinase that catalyzes the following near-equilibrium reaction:

$$\text{AMP} + \text{ATP} \rightleftharpoons 2\ \text{ADP} \qquad (10.13)$$

When the concentration of AMP increases, AMP can react with ATP to form two molecules of ADP. These ADP molecules can be converted to two molecules of ATP. The overall process is

$$\text{AMP} + \text{ATP} + 2\ P_i \rightleftharpoons 2\ \text{ATP} + 2\ H_2O \qquad (10.14)$$

ATP concentrations in cells are greater than ADP or AMP concentrations, and relatively minor changes in the concentration of ATP can result in large changes in the concentrations of the di- and monophosphates. Table 10.3 shows the theoretical increases in

Table 10.3 Theoretical changes in concentrations of adenine nucleotides

ATP (mM)	ADP (mM)	AMP (mM)
4.8	0.2	0.004
4.5	0.5	0.02
3.9	1.0	0.11
3.2	1.5	0.31

[Adapted from Newsholme. E. A., and Leech, A. R. (1986). *Biochemistry for the Medical Science* (New York: John Wiley & Sons), p. 315.]

[ADP] and [AMP] under conditions in which ATP is consumed, assuming that the total adenine nucleotide concentration remains 5.0 mM. Note that when the ATP concentration decreases from 4.8 mM to 4.5 mM (a decrease of about 6%), the ADP concentration increases 2.5-fold and the AMP concentration increases 5-fold. In fact, when cells are well supplied with oxidizable fuels and oxygen, they maintain a balance of adenine nucleotides in which ATP is present at a steady concentration of 2 to 10 mM, [ADP] is less than 1 mM, and [AMP] is even lower. As we will see, ADP and AMP are often effective allosteric modulators of some energy-yielding metabolic processes. ATP, whose concentration is relatively constant, is generally not an important modulator under physiological conditions.

One important consequence of the concentrations of ATP and its hydrolysis products *in vivo* is that the free energy change for ATP hydrolysis is actually greater than the standard value of −32 kJ mol^{-1}. This is illustrated in Sample Calculation 10.2 using measured concentrations of ATP, ADP, and P_i from rat liver cells. The calculated Gibbs free energy change is close to the value determined in many other types of cells.

As mentioned above, ATP hydrolysis is an example of a metabolically irreversible reaction. The activities of various enzymes are regulated so they become inactive as ATP concentrations fall below a minimal threshold. Thus, the reverse of the hydrolysis reaction, leading to ATP synthesis, does not occur except under special circumstances (Chapter 14). We will see in Chapter 14 that ATP is synthesized by another pathway.

The importance of maintaining a high concentraion of ATP cannot be overemphasized. It is required in order to get a large free energy change from ATP hydrolysis. Cells will die if the reactants and products reach equilibrium.

10.7 The Metabolic Roles of ATP

The energy produced by one biological reaction or process, such as the synthesis of X—Y in Reaction 10.15, is often coupled to a second reaction, such as the hydrolysis of ATP. The first reaction would not otherwise occur spontaneously.

$$\begin{aligned} X + Y &\rightleftharpoons X{-}Y \\ ATP + H_2O &\rightleftharpoons ADP + P_i + H^{\oplus} \end{aligned} \qquad \textbf{(10.15)}$$

SAMPLE CALCULATION 10.2 Gibbs Free Energy Change

Q: In a rat hepatocyte, the concentrations of ATP, ADP, and P_i are 3.4 mM, 1.3 mM, and 4.8 mM, respectively. Calculate the Gibbs free energy change for hydrolysis of ATP in this cell. How does this compare to the standard free energy change?

A: The actual Gibbs free energy change is calculated according to Equation 10.10.

$$\Delta G_{reaction} = \Delta G^{\circ\prime}_{reaction} + RT \ln \frac{[ADP][P_i]}{[ATP]} = \Delta G^{\circ}_{reaction} + 2.303\ RT \log \frac{[ADP][P_i]}{[ATP]}$$

When known values and constants are substituted (with concentrations expressed as molar values), assuming pH7.0 and 25°C.

$$\Delta G = -32000\ J\ mol^{-1} + (8.31\ JK^{-1}mol^{-1})(298\ K)\left[2.303 \log \frac{(1.3 \times 10^{-3})(4.8 \times 10^{-3})}{(3.4 \times 10^{-3})}\right]$$

$$\Delta G = -32000\ J\ mol^{-1} + (2480\ J\ mol^{-1})[2.303\ \log(1.8 \times 10^{-3})]$$

$$\Delta G = -32000\ J\ mol^{-1} - 16000\ J\ mol^{-1}$$

$$\Delta G = -48000\ J\ mol^{-1} = -48\ kJ\ mol^{-1}$$

The actual free energy change is about $1^1/_2$ times the standard free energy change.

The sum of the Gibbs free energy changes for the coupled reactions must be negative for the reactions to proceed. This does not mean that both of the individual reactions have to be favored in isolation ($\Delta G < 0$). The advantage of coupled reactions is that the energy released from one of them can be used to drive the other even when the second reaction is unfavorable by itself ($\Delta G > 0$). (Recall that the ability to couple reactions is one of the key properties of enzymes.)

Energy flow in metabolism depends on many coupled reactions involving ATP. In many cases, the coupled reactions are linked by a shared intermediate such as a phosphorylated derivative of reactant X.

$$\begin{aligned} X + ATP &\rightleftharpoons X{-}P + ADP \\ X{-}P + Y + H_2O &\rightleftharpoons X{-}Y + P_i + H^{\oplus} \end{aligned} \qquad \textbf{(10.16)}$$

Transfer of either a phosphoryl group or a nucleotidyl group to a substrate activates that substrate (i.e., prepares it for a reaction that has a large negative Gibbs free energy change). The activated compound (X—P), can be either a metabolite or the side chain of an amino acid residue in the active site of an enzyme. The intermediate then reacts with a second substrate to complete the reaction.

A. Phosphoryl Group Transfer

The synthesis of glutamine from glutamate and ammonia illustrates how the "high energy" compound ATP drives a biosynthetic reaction. This reaction, catalyzed by glutamine synthetase, allows organisms to incorporate inorganic nitrogen into biomolecules as carbon-bound nitrogen. In this synthesis of an amide bond, the γ-carboxyl group of the substrate is activated by synthesis of an anhydride intermediate.

Glutamine synthetase catalyzes the nucleophilic displacement of the γ-phosphoryl group of ATP by the γ-carboxylate of glutamate. ADP is released, producing enzyme-bound γ-glutamyl phosphate as an intermediate (Figure 10.14). γ-Glutamyl phosphate is unstable in aqueous solution but is protected from water in the active site of glutamine synthetase. In the second step of the mechanism, ammonia acts as a nucleophile, displacing the phosphate (a good leaving group) from the carbonyl carbon of γ-glutamyl phosphate to generate the product, glutamine. Overall, one molecule of ATP is converted to ADP + P_i for every molecule of glutamine formed from glutamate and ammonia.

BOX 10.1 THE SQUIGGLE

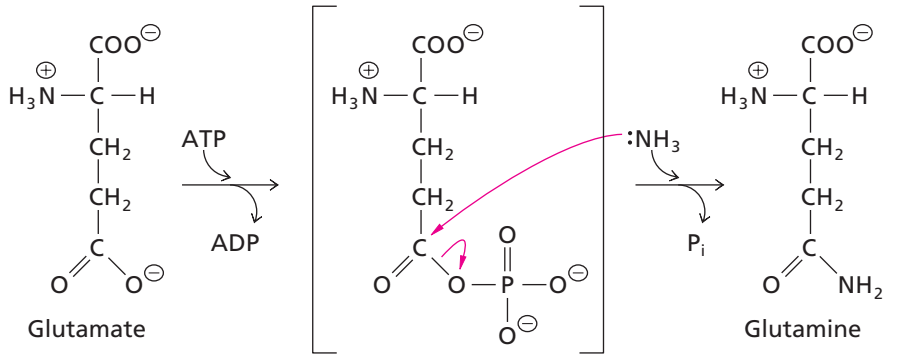
Fritz Lipmann (1899–1986) won the Nobel Prize in Physiology and Medicine in 1953 for discovering coenzyme A. He also made important contributions to our understanding of ATP as an energy currency. In 1941 he introduced the idea of a high energy bond in ATP by drawing it as a squiggle (~). For the next several decades, biochemistry textbooks often depicted ATP with two high energy bonds.

AMP~P~P

We know now that this depiction is misleading since there's nothing special about the covalent bonds in phosphoanhydride linkages. It's the overall system of reactants and products that makes the ATP currency so valuable and not the energy of individual bonds. However, it's true that the three main explanations for the high energy of ATP (electrostatic repulsion, solvation effects, and resonance stabilization) are due mostly to the phosphoanhydride linkages so the focus on that particular linkage isn't entirely wrong. The squiggle used to be very common in the older scientific literature and in textbooks but it's much less common today.



Source: Lipmann, F. (1941) Metabolic generation and utilization of phosphate bond energy. *Advances in Enzymology* 1:99–162.



(10.17)

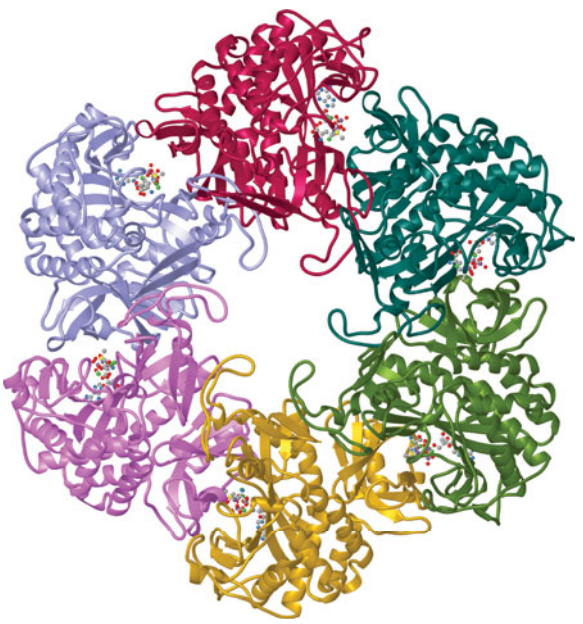
We can calculate the predicted standard Gibbs free energy change for the reaction that is not coupled to ATP hydrolysis.

$$\text{Glutamate} + NH_4^{\oplus} \rightleftharpoons \text{glutamine} + H_2O \qquad \textbf{(10.18)}$$

$$\Delta G^{\circ\prime}_{\text{reaction}} = +14 \text{ kJ mol}^{-1}$$

This is a standard free energy change so it doesn't necessarily reflect the actual Gibbs free energy change given cellular concentrations of glutamate, glutamine, and ammonia. The hypothetical Reaction 10.18 might be associated with a negative free energy change inside the cell if the concentrations of glutamate and ammonia were high relative to the concentration of glutamine. But this is not the case. The steady-state concentrations of glutamate and glutamine must be kept nearly equivalent in order to support protein synthesis and other metabolic pathways. This means that the Gibbs free energy change for the hypothetical Reaction 10.18 cannot be negative. Furthermore, the concentration of ammonia is very low relative to glutamate and glutamine. In both bacteria and eukaryotes, ammonia must be efficiently incorporated into glutamine even when the concentration of free ammonia is very low. Thus Reaction 10.18 is not possible in living cells due to the requirement for a high steady-state concentration of glutamine and due to a limiting supply of ammonia. Glutamine synthesis must be coupled to hydrolysis of ATP in order to drive it in the right direction.

Glutamine synthetase catalyzes a phosphoryl group transfer reaction in which the phosphorylated compound is a transient intermediate (Reaction 10.17). There are other reactions that produce a stable phosphorylated product. As we have seen, kinases catalyze



◂ **Figure 10.14**
Glutamine synthetase bound to ADP and a transition state analog. Glutamine synthetase from *Mycobacterium tuberculosis* is a complex enzyme consisting of two hexameric rings on top of each other. Only one ring is shown in this figure. The active site is occupied by ADP and a transition state analog (L-methionine-*S*-sulfoximine phosphate) that resembles γ-glutamyl phosphate. [PDB 2BVC]

Table 10.4 Standard Gibbs free energies of hydrolysis for common metabolites

Metabolite	$\Delta G^{\circ\prime}_{hydrolysis}$ (kJ mol^{-1})
Phosphoenolpyruvate	−62
1, 3-*Bis*phosphoglycerate	−49
ATP to AMP + PP_i	−45
Phosphocreatine	−43
Phosphoarginine	−32
Acetyl CoA	−32
Acyl CoA	−31
ATP to ADP + P_i	−32
Pyrophosphate	−29
Glucose 1-phosphate	−21
Glucose 6-phosphate	−14
Glycerol 3-phosphate	−9

KEY CONCEPT

Many phosphorylated metabolites have group transfer potentials similar to that of ATP.

transfer of the γ-phosphoryl group from ATP (or, less frequently, from another nucleoside triphosphate) to another substrate. Kinases typically catalyze metabolically irreversible reactions. A few kinase reactions, however, such as those catalyzed by adenylate kinase (Reaction 10.13) and creatine kinase (Section 10.7B), are near equilibrium reactions. Although the reactions they catalyze are sometimes described as phosphate group transfer reactions, kinases actually transfer a phosphoryl group (—$PO_3^{(2-)}$—)to their acceptors.

The ability of a phosphorylated compound to transfer its phosphoryl group(s) is termed its **phosphoryl group transfer potential**, or simply group transfer potential. Some compounds, such as phosphoanhydrides, are excellent phosphoryl group donors. They may have a group transfer potential equal to or greater than that of ATP. Other compounds, such as phosphoesters, are poor phosphoryl group donors. They have a group transfer potential less than that of ATP. Under standard conditions, group transfer potentials have the same values as the standard free energies of hydrolysis but are opposite in sign. Thus, the group transfer potential is a measure of the free energy required for formation of the phosphorylated compound. In Table 10.4 we list the standard Gibbs free energy of hydrolysis for a number of phosphorylated compounds.

B. Production of ATP by Phosphoryl Group Transfer

Often, one kinase catalyzes transfer of a phosphoryl group from an excellent donor to ADP to form ATP, which then acts as a donor for a different kinase reaction. Phosphoenolpyruvate and 1,3-bisphosphoglycerate are two examples of common metabolites that have higher energy than ATP even under conditions found inside the cell ($\Delta G <$ −50 kJ mol^{-1}). Some of these compounds are intermediates in catabolic pathways; others are energy storage compounds.

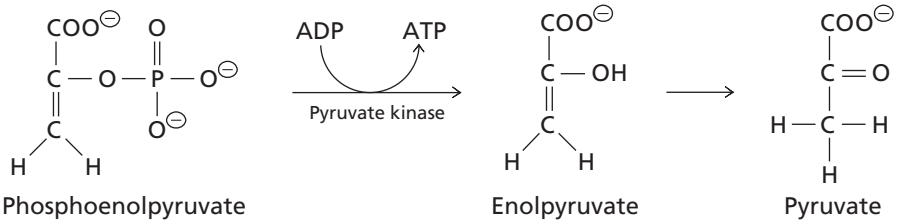
Phosphoenolpyruvate, an intermediate in the glycolytic pathway, has the highest phosphoryl group transfer potential known. The standard free energy of phosphoenolpyruvate hydrolysis is −62 kJ mol^{-1} and the actual Gibbs free energy change is comparable to that of ATP. The free energy of hydrolysis for phosphoenolpyruvate can be understood by picturing the molecule as an enol whose structure is locked by attachment of the phosphoryl group. When the phosphoryl group is removed, the molecule spontaneously forms the much more stable keto tautomer (Figure 10.15). Transfer of the phosphoryl group from phosphoenolpyruvate to ADP is catalyzed by the enzyme pyruvate kinase. Because the $\Delta G^{\circ\prime}$ for the reaction is about −30 kJ mol^{-1}, the equilibrium for this reaction under standard conditions lies far in the direction of transfer of the phosphoryl group from phosphoenolpyruvate to ADP. In cells, this metabolically irreversible reaction is an important source of ATP.

Phosphagens, including phosphocreatine and phosphoarginine, are "high energy" phosphate storage molecules found in animal muscle cells. Phosphagens are phosphoamides (rather than phosphoanhydrides) and have higher group transfer potentials than ATP. In the muscles of vertebrates, large amounts of phosphocreatine are formed during times of ample ATP supply. In resting muscle, the concentration of phosphocreatine is about fivefold higher than that of ATP. When ATP levels fall, creatine kinase catalyzes rapid replenishment of ATP through transfer of the activated phosphoryl group from phosphocreatine to ADP.

$$\text{Phosphocreatine} + \text{ADP} \overset{\text{Creatine kinase}}{\rightleftharpoons} \text{creatine} + \text{ATP} \qquad (10.19)$$

The supply of phosphocreatine is adequate for 3- to 4-second bursts of activity, long enough for other metabolic processes to begin restoring the ATP supply. Under cellular conditions, the creatine kinase reaction is a near-equilibrium reaction. In many invertebrates—notably mollusks and arthropods—phosphoarginine is the source of the activated phosphoryl group.

Because ATP has an intermediate phosphoryl group transfer potential, it is thermodynamically suited as a carrier of phosphoryl groups. (Figure 10.15) ATP is also kinetically stable under physiological conditions until acted on by an enzyme so it can carry chemical potential energy from one enzyme to another without being hydrolyzed. Not surprisingly, ATP mediates most chemical energy transfers in all organisms.

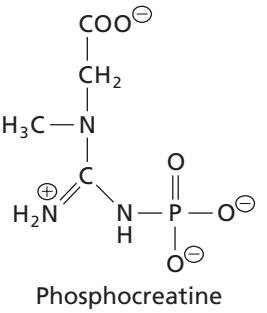


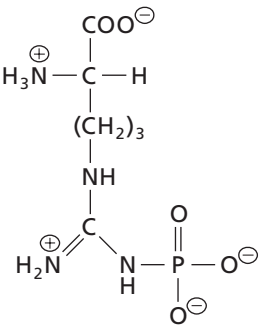
◀ **Figure 10.15**
Transfer of the phosphoryl group from phosphoenolpyruvate to ADP.

C. Nucleotidyl Group Transfer

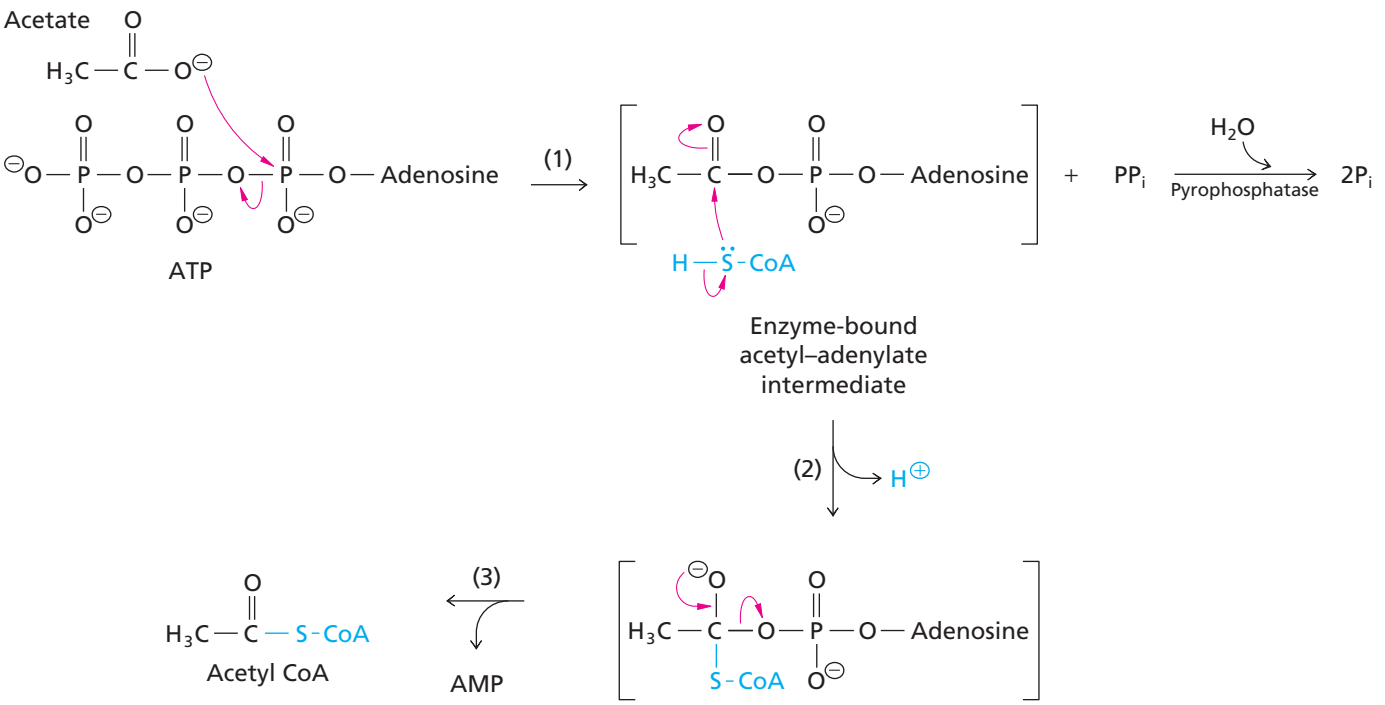
The other common group transfer reaction involving ATP is transfer of the nucleotidyl group. An example is the synthesis of acetyl CoA, catalyzed by acetyl-CoA synthetase. In this reaction, the AMP moiety of ATP is transferred to the nucleophilic carboxylate group of acetate to form an acetyl–adenylate intermediate (Figure 10.16). Note that pyrophosphate (PP_i) is released in this step. Like the glutamyl–phosphate intermediate in Reaction 10.17, the reactive intermediate is shielded from nonenzymatic hydrolysis by tight binding within the active site of the enzyme. The reaction is completed by transfer of the acetyl group to the nucleophilic sulfur atom of coenzyme A, leading to the formation of acetyl CoA and AMP.

The synthesis of acetyl CoA also illustrates how the removal of a product can cause a metabolic reaction to approach completion, just as the formation of a precipitate or a gas can drive an inorganic reaction toward completion. The standard Gibbs free energy for the formation of acetyl CoA from acetate and CoA is about -13 kJ mol^{-1} ($\Delta G^{\circ\prime}_{\text{hydrolysis}}$ of acetyl CoA = -32 kJ mol^{-1}). But note that the product PP_i is hydrolyzed to two molecules of P_i by the action of pyrophosphatase (Section 10.6). Almost all cells have high levels of activity of this enzyme, so the concentration of PP_i in cells is generally very low (less than 10^{-6} M). Cleavage of PP_i contributes to the negative value of the standard Gibbs free energy change for the overall reaction. The additional hydrolytic reaction adds the energy cost of one phosphoanhydride linkage to the overall synthetic process. In reactions such as this, we say that the cost is two ATP equivalents in order to emphasize that two "high energy" compounds are hydrolyzed. Hydrolysis of pyrophosphate accompanies many synthetic reactions in metabolism.





▲ **Structures of phosphocreatine and phosphoarginine.**



▲ **Figure 10.16**
Synthesis of acetyl CoA from acetate, catalyzed by acetyl-CoA synthetase.

10.8 Thioesters Have High Free Energies of Hydrolysis

Thioesters are another class of "high energy" compounds forming part of the currency of metabolism. Acetyl CoA is one example. It occupies a central position in metabolism (Figures 10.8 and 10.9). The high energy of thioester reactions can be used in generating ATP equivalents or in transferring the acyl groups to acceptor molecules. Recall that acyl groups are attached to coenzyme A (or acyl carrier protein) via a thioester linkage (Section 7.6 and Figure 7.13).

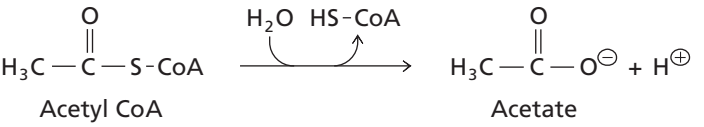


(10.20)

Unlike oxygen esters of carboxylic acids, thioesters resemble carboxylic acid anhydrides in reactivity. Sulfur is in the same group of the periodic table as oxygen but thioesters are less stable than typical esters because the unshared electrons of the sulfur atom are not as effectively delocalized in a thioester as the unshared electrons in an oxygen ester. The energy associated with hydrolyzing the thioester linkage is similar to the energy of hydrolysis of the phosphoanhydride linkages in ATP. The standard Gibbs free energy change for hydrolysis of acetyl CoA is -31 kJ mol^{-1}, and the actual change may somewhat smaller (more negative) under conditions inside the cell.

KEY CONCEPT

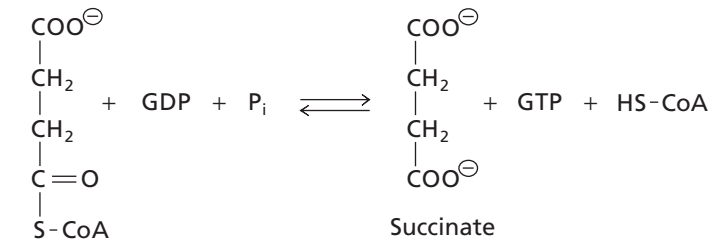
Reactions involving thioesters, such as acetyl CoA, release amounts of energy comparable to that of ATP hydrolysis.



(10.21)

Despite its high free energy of hydrolysis, a CoA thioester resists nonenzymatic hydrolysis at neutral pH values. In other words, it is kinetically stable in the absence of appropriate catalysts.

The high energy of hydrolysis of a CoA thioester is used in the fifth step of the citric acid cycle, when the thioester succinyl CoA reacts with GDP (or sometimes ADP) and P_i to form GTP (or ATP).



(10.22)

We discuss succinyl CoA synthetase in Section 13.4, part 5, and fatty acid synthesis in Section 16.5.

This substrate-level phosphorylation conserves energy used in the formation of succinyl CoA as ATP equivalents. The energy of thioesters also drives the synthesis of fatty acids.

10.9 Reduced Coenzymes Conserve Energy from Biological Oxidations

In Section 14.11 we will learn that NADH is equivalent to 2.5 ATPs and QH_2 is equivalent to 1.5 ATPs.

Many reduced coenzymes are "high energy" compounds in the sense we described earlier (i.e., part of a system). Their high energy (or reducing power) can be donated in oxidation-reduction reactions. The energy of reduced coenzymes may be represented as ATP equivalents since their oxidation can be coupled to the synthesis of ATP.

As described in Section 6.1C, the oxidation of one molecule must be coupled with the reduction of another molecule. A molecule that accepts electrons and is reduced is an oxidizing agent. A molecule that loses electrons and is oxidized is a reducing agent. The net oxidation–reduction reaction is

$$A_{red} + B_{ox} \rightleftharpoons A_{ox} + B_{red} \qquad (10.23)$$

The electrons released in biological oxidation reactions are transferred enzymatically to oxidizing agents, usually a pyridine nucleotide ($NAD^{\oplus}$ or sometimes $NADP^{\oplus}$), a flavin coenzyme (FMN or FAD), or ubiquinone (Q). When $NAD^{\oplus}$ and $NADP^{\oplus}$ are reduced, their nicotinamide rings accept a hydride ion (Figure 7.8). One electron is lost when a hydrogen atom (composed of one proton and one electron) is removed and two electrons are lost when a hydride ion (composed of one proton and two electrons) is removed. (Remember that oxidation is loss of electrons.)

NADH and NADPH, along with QH_2, supply reducing power. $FMNH_2$ and $FADH_2$ are reduced enzyme-bound intermediates in some oxidation reactions.

The structures and functions of $NAD^{\oplus}$ and $NADP^{\oplus}$ are discussed in Section 7.4, of FMN and FAD in Section 7.5, and of ubiquinone in Section 7.14.

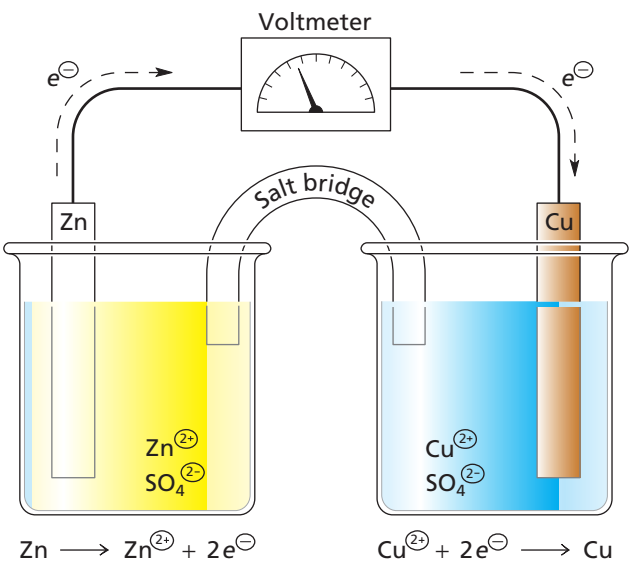
A. Gibbs Free Energy Change Is Related to Reduction Potential

The **reduction potential** of a reducing agent is a measure of its thermodynamic reactivity. Reduction potential can be measured in electrochemical cells. An example of a simple inorganic oxidation–reduction reaction is the transfer of a pair of electrons from a zinc atom (Zn) to a copper ion (Cu^{2+}).

$$Zn + Cu^{2+} \rightleftharpoons Zn^{2+} + Cu \qquad (10.24)$$

This reaction can be carried out in two separate solutions that divide the overall reaction into two half-reactions (Figure 10.17). At the zinc electrode, two electrons are given up by each zinc atom that reacts (the reducing agent). The electrons flow through a wire to the copper electrode, where they reduce Cu^{2+} (the oxidizing agent) to metallic copper. A salt bridge, consisting of a tube with a porous partition filled with electrolyte, preserves electrical neutrality by providing an aqueous path for the flow of nonreactive counterions between the two solutions. The flow of ions and the flow of electons are separated in such an electrochemical cell and electron flow through the wire (i.e., electric energy) can be measured using a voltmeter.

The direction of the current through the circuit in Figure 10.17 indicates that Zn is more easily oxidized than Cu (i.e., Zn is a stronger reducing agent than Cu). The reading on the voltmeter represents a potential difference, the difference between the reduction potential of the reaction on the left and that on the right. The measured potential difference is the **electromotive force**.



◂ **Figure 10.17**
Diagram of an electrochemical cell. Electrons flow through the external circuit from the zinc electrode to the copper electrode. The salt bridge permits the flow of counterions (sulfate ions in this example) without extensive mixing of the two solutions. The electromotive force is measured by the voltmeter connected across the two electrodes. (Two other kinds of salt bridges are shown in Section 2.5A.)

KEY CONCEPT

All standard reduction potentials are measured relative to the reduction of $H^{\oplus}$ under standard conditions.

KEY CONCEPT

ΔE must be positive for an oxidation reduction reaction to proceed in the direction written.

It is useful to have a reference standard for measurements of reduction potentials just as in measurements of Gibbs free energy changes. For reduction potentials, the reference is not simply a set of reaction conditions, but a reference half-reaction to which all other half-reactions can be compared. The reference half-reaction is the reduction of $H^{\oplus}$ to hydrogen gas (H_2). The reduction potential of this half-reaction under standard conditions ($E°$) is arbitrarily set at 0.0 V. The standard reduction potential of any other half-reaction is measured with an oxidation–reduction coupled reaction in which the reference half-cell contains a solution of 1 M $H^{\oplus}$ and 1 atm H_2 (gaseous), and the sample half-cell contains 1 M each of the oxidized and reduced species of the substance whose reduction potential is to be determined. Under standard conditions for biological measurements, the hydrogen ion concentration in the sample half-cell is (10^{-7}M). The voltmeter across the oxidation–reduction couple measures the electromotive force, or the difference in the reduction potential, between the reference and sample half-reactions. Since the standard reduction potential of the reference half-reaction is 0.0 V, the measured potential is that of the sample half-reaction.

Table 10.5 gives the standard reduction potentials at pH 7.0 ($E°'$) of some important biological half-reactions. Electrons flow spontaneously from the more readily oxidized

Table 10.5 Standard reduction potentials of some important biological half-reactions

Reduction half-reaction	$E°'$ (V)
Acetyl CoA + CO_2 + $H^{\oplus}$ + $2e^{\ominus}$ → Pyruvate + CoA	−0.48
Ferredoxin (spinach). Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	−0.43
2 $H^{\oplus}$ + $2e^{\ominus}$ → H_2 (at pH 7.0)	−0.42
α-Ketoglutarate + CO_2 + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Isocitrate	−0.38
Lipoyl dehydrogenase (FAD) + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Lipoyl dehydrogenase ($FADH_2$)	−0.34
$NADP^{\oplus}$ + $H^{\oplus}$ + $2e^{\ominus}$ → NADPH	−0.32
$NAD^{\oplus}$ + $H^{\oplus}$ + $2e^{\ominus}$ → NADH	−0.32
Lipoic acid + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Dihydrolipoic acid	−0.29
Thioredoxin (oxidized) + $2H^{\oplus}$ + $2e$ → Thioredoxin (reduced)	−0.28
Glutathione (oxidized) + 2 $H^{\oplus}$ + $2e^{\ominus}$ → 2 Glutathione (reduced)	−0.23
FAD + 2 $H^{\oplus}$ + $2e^{\ominus}$ → $FADH_2$	−0.22
FMN + 2 $H^{\oplus}$ + $2e^{\ominus}$ → $FMNH_2$	−0.22
Acetaldehyde + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Ethanol	−0.20
Pyruvate + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Lactate	−0.18
Oxaloacetate + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Malate	−0.17
Cytochrome b_5 (microsomal). Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.02
Fumarate + 2 $H^{\oplus}$ + $2e^{\ominus}$ → Succinate	0.03
Ubiquinone (Q) + 2 $H^{\oplus}$ + $2e^{\ominus}$ → QH_2	0.04
Cytochrome b (mitochondrial), Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.08
Cytochrome c_1, Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.22
Cytochrome c, Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.23
Cytochrome a, Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.29
Cytochrome f, Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.36
Plastocyanin, Cu^{2+} + $e^{\ominus}$ → Cu^{+}	0.37
$NO_3^{\ominus}$ + 2 $H^{\oplus}$ + $2e^{\ominus}$ → $NO_2^{\ominus}$ + H_2O	0.42
Photosystem I (P700)	0.43
Fe^{3+} + $e^{\ominus}$ → Fe^{2+}	0.77
$^1/_2O_2$ + 2 $H^{\oplus}$ + $2e^{\ominus}$ → H_2O	0.82
Photosystem II (P680)	1.1

substance (the one with the more negative reduction potential) to the more readily reduced substance (the one with the more positive reduction potential). Therefore, more negative potentials are assigned to reaction systems that have a greater tendency to donate electrons (i.e., systems that tend to oxidize more easily).

The standard reduction potential for the transfer of electrons from one molecular species to another is related to the standard free energy change for the oxidation–reduction reaction by the equation

$$\Delta G^{\circ\prime} = -nF\Delta E^{\circ\prime} \tag{10.25}$$

where n is the number of electrons transferred and F is Faraday's constant (96.48 kJ V^{-1} mol^{-1}). Note that Equation 10.25 resembles Equation 9.5 except that here we are dealing with reduction potential and not membrane potential. $\Delta E^{\circ\prime}$ is defined as the difference in volts between the standard reduction potential of the electron-acceptor system and that of the electron donor system.

$$\Delta E^{\circ\prime} = E^{\circ\prime}_{\text{electron acceptor}} - E^{\circ\prime}_{\text{electron donor}} \tag{10.26}$$

Recall from Equation 10.6 that $\Delta G^{\circ\prime} = -RT \ln K_{eq}$. Combining this equation with Equation 10.25, we have

$$\Delta E^{\circ\prime} = \frac{RT}{nF} \ln K_{eq} \tag{10.27}$$

Under biological conditions, the reactants in a system are not present at standard concentrations of 1 M. Just as the actual Gibbs free energy change for a reaction is related to the standard Gibbs free energy change by Equation 10.6, an observed difference in reduction potentials (ΔE) is related to the difference in the standard reduction potentials ($\Delta E^{\circ\prime}$) by the Nernst equation. For Reaction 10.23, the Nernst equation is

$$\Delta E = \Delta E^{\circ\prime} - \frac{RT}{nF} \ln \frac{[A_{ox}][B_{red}]}{[A_{red}][B_{ox}]} \tag{10.28}$$

At 298 K, Equation 10.28 reduces to

$$\Delta E = \Delta E^{\circ\prime} - \frac{0.026}{n} \ln Q \tag{10.29}$$

where Q represents the actual concentrations of reduced and oxidized species. To calculate the electromotive force of a reaction under nonstandard conditions, use the Nernst equation and substitute the actual concentrations of reactants and products. Keep in mind that a *positive* ΔE value indicates that an oxidation–reduction reaction will have a *negative* standard Gibbs free energy change.

B. Electron Transfer from NADH Provides Free Energy

$NAD^{\oplus}$ is reduced to NADH in coupled reactions where electrons are transferred from a metabolite to $NAD^{\oplus}$. The reduced form of the coenzyme (NADH) becomes a source of electrons in other oxidation–reduction reactions. The Gibbs free energy changes associated with the overall oxidation–reduction reaction under standard conditions can be calculated from the standard reduction potentials of the two half-reactions using Equation 10.25. As an example, let's consider the reaction where NADH is oxidized and molecular oxygen is reduced. This represents the available free energy change during membrane-associated electron transport. This free energy is recovered in the form of ATP synthesis (Chapter 14).

The two half reactions from Table 10.5 are

$$NAD^{\oplus} + H^{\oplus} + 2\,e^{\ominus} \longrightarrow NADH \qquad E^{\circ\prime} = -0.32\ V \tag{10.30}$$

and

$$^1/_2\,O_2 + 2\,H^{\oplus} + 2\,e^{\ominus} \longrightarrow H_2O \qquad E^{\circ\prime} = 0.82\ V \tag{10.31}$$

Since the $NAD^{\oplus}$ half-reaction has the more negative standard reduction potential, NADH is the electron donor and oxygen is the electron acceptor. Note that the values in Table 10.5 are for half-reactions written as reductions (gain of electrons). That's because $E^{\circ\prime}$ is a reduction potential. In an oxidation–reduction reaction, two of these half-reactions are combined. One of them will be an oxidation reaction, so the equation in Table 10.5 must be reversed. The reduction potentials tell you which way the electrons will flow. They flow from the half-reaction near the top of the table (more negative $E^{\circ\prime}$) to the one nearer the bottom of the table (less negative $E^{\circ\prime}$) (Figure 10.18). What this means is that the overall $\Delta E^{\circ\prime}$ for the complete reaction will be positive according to Equation 10.26. (This is the American convention. The European convention uses a different way of arriving at the same answer.)

The net oxidation–reduction reaction is Reaction 10.31 plus the reverse of Reaction 10.30.

$$NADH + {}^{1}\!/_{2}\, O_2 + H^{\oplus} \longrightarrow NAD^{\oplus} + H_2O \qquad \textbf{(10.32)}$$

and $\Delta E^{\circ\prime}$ for the reaction is

$$\Delta E^{\circ\prime} = E^{\circ\prime}_{O_2} - E^{\circ\prime}_{NADH} = 0.82\ V - (-0.32\ V) = 1.14\ V \qquad \textbf{(10.33)}$$

Using Equation 10.25,

$$\Delta G^{\circ\prime} = -(2)(96.48\ kJ\ V^{-1}\ mol^{-1})(1.14\ V) = -220\ kJ\ mol^{-1} \qquad \textbf{(10.34)}$$

The standard Gibbs free energy change for the formation of ATP from ADP + P_i is +32 kJ mol^{-1} (the actual free energy change is about +48 kJ mol^{-1} under the conditions of the living cell, as noted earlier). The energy released during the oxidation of NADH under cellular conditions is sufficient to drive the formation of several molecules of ATP. We will learn in Chapter 14 that the actual energy yield of an NADH molecule is about 2.5 ATP equivalents (Section 14.11).

KEY CONCEPT

The standard Gibbs free energy change of an oxidation–reduction reaction is calculated from the reduction potentials of the two half-reactions.

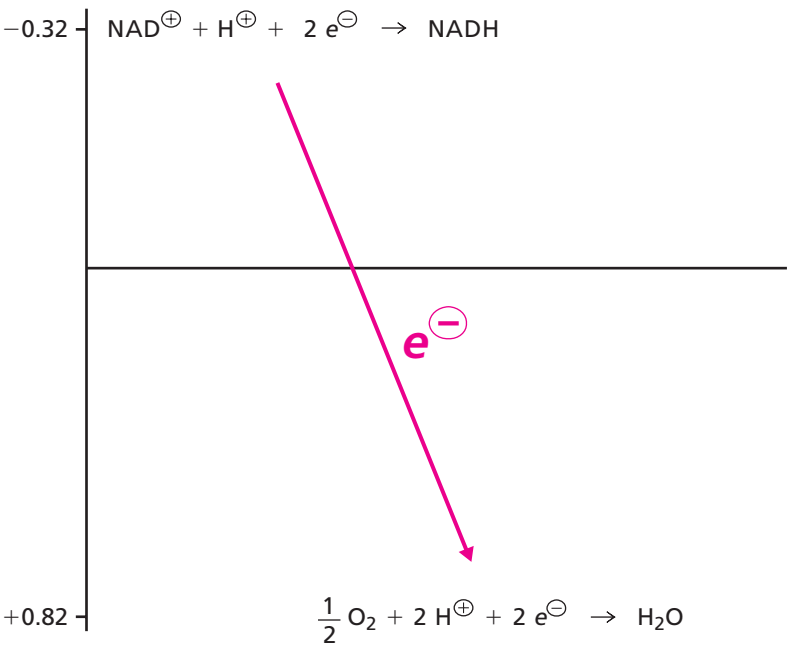
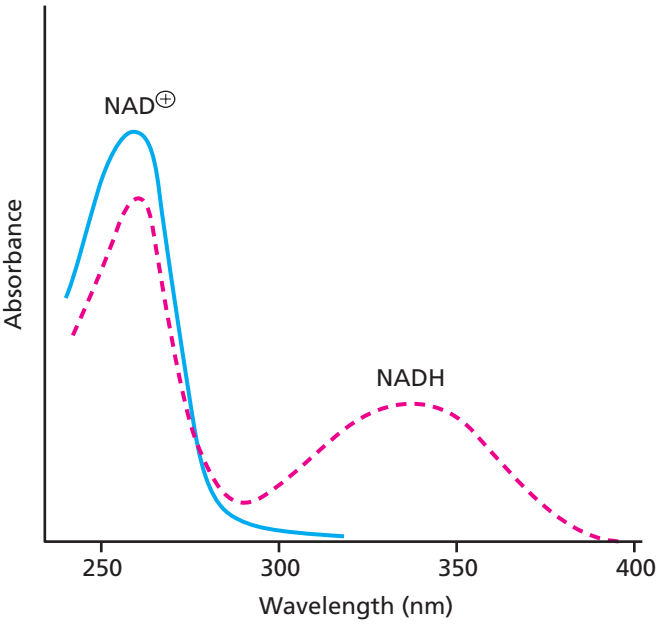


Figure 10.18 ▸
Electron flow in oxidation–reduction reactions. Half-reactions can be plotted on a chart where the standard reduction potentials are on the x axis, arranged so that the most negative values are at the top of the chart. Using this convention, electrons flow from the half-reaction at the top of the chart to the one nearer the bottom of the chart.

BOX 10.2 NAD⊕ AND NADH DIFFER IN THEIR ULTRAVIOLET ABSORPTION SPECTRA

The differing absorption spectra of NAD⊕ and NADH are useful in experimental work. NAD⊕ (and NADP⊕) absorbs maximally at 260 nm. This absorption is due to both the adenine and nicotinamide moieties. When NAD⊕ is reduced to NADH (or NADP⊕ to NADPH), the absorbance at 260 nm decreases and an absorption band centered at 340 nm appears (adjacent figure). The 340-nm band comes from the formation of the reduced nicotinamide ring. The spectra of NAD⊕ and NADH do not change in the pH range 2 to 10 in which most enzymes are active. In addition, few other biological molecules undergo changes in light absorption near 340 nm.

In a suitably prepared enzyme assay, one can determine the rate of formation of NADH by measuring an increase in the absorbance at 340 nm. Similarly, in a reaction proceeding in the opposite direction, the rate of NADH oxidation is indicated by the rate of decrease in absorbance at 340 nm. Many dehydrogenases can be directly assayed by this procedure. In addition, the concentrations of a product formed in a nonoxidative reaction can often be determined by oxidizing the product in a dehydrogenase– NAD⊕ system. Such a measurement of concentrations of NAD⊕ or NADH by their absorption of ultraviolet light is used not only in the research laboratory but also in many clinical analyses.



▲ Ultraviolet absorption spectra of NAD⊕ and NADH.

10.10 Experimental Methods for Studying Metabolism

The complexity of many metabolic pathways makes them difficult to study. Reaction conditions used with isolated reactants in the test tube (*in vitro*) are often very different from the reaction conditions in the intact cell (*in vivo*). The study of the chemical events of metabolism is one of the oldest branches of biochemistry, and many approaches have been developed to characterize the enzymes, intermediates, flux, and regulation of metabolic pathways.

A classical approach to unraveling metabolic pathways is to add a substrate to preparations of tissues, cells, or subcellular fractions and then follow the emergence of intermediates and end products. The fate of a substrate is easier to trace when the substrate has been specifically labeled. Since the advent of nuclear chemistry, isotopic tracers have been used to map the transformations of metabolites. For example, compounds containing atoms of radioactive isotopes such as ^{3}H or ^{14}C can be added to cells or other preparations, and the radioactive compounds produced by anabolic or catabolic reactions can be purified and identified. Nuclear magnetic resonance (NMR) spectroscopy can trace the reactions of certain isotopes. It can also be employed to study the metabolism of whole animals (including humans) and is being used for clinical analysis.

Verification of the steps of a particular pathway can be accomplished by reproducing the separate reactions *in vitro* using isolated substrates and enzymes. Individual enzymes have been isolated for almost all known metabolic steps. By determining the substrate specificity and kinetic properties of a purified enzyme, it is possible to draw some conclusions regarding the regulatory role of that enzyme. This reductionist approach has led to many of the key concepts in this book. It's the approach that allows us to understand the relationship between structure and function. However, a complete assessment of the regulation of a pathway requires analysis of metabolite concentrations in the intact cell or organism under various conditions.

Valuable information can be acquired by studying mutations in single genes associated with the production of inactive or defective individual enzyme forms. Whereas some mutations are lethal and not transmitted to subsequent generations, others can be tolerated by the descendants. The investigation of mutant organisms has helped identify enzymes and intermediates of numerous metabolic pathways. Typically, a defective enzyme results in a deficiency of its product and the accumulation of its substrate or a product derived from the substrate by a branch pathway. This approach has been extremely successful in identifying metabolic pathways in simple organisms such as bacteria, yeast, and *Neurospora* (Box 7.4). In humans, enzyme defects are manifested in metabolic diseases. Hundreds of single-gene diseases are known. Some are extremely rare, and others are fairly common; some are tragically severe. In cases where a metabolic disorder produces only mild symptoms, it appears that the network of metabolic reactions contains enough overlap and redundancy to allow near-normal development of the organism.

In instances where natural mutations are not available, mutant organisms can be generated by treatment with radiation or chemical mutagens (agents that cause mutation). Biochemists have characterized entire pathways by producing a series of mutants, isolating them, and examining their nutritional requirements and accumulated metabolites. More recently, site-directed mutagenesis (Box 6.1) has proved valuable in defining the roles of enzymes. Bacterial and yeast systems have been the most widely used for introducing mutations because large numbers of these organisms can be grown in a short period of time. It is possible to produce animal models—particularly insects and nematodes—in which certain genes are not expressed. It is also possible to delete certain genes in vertebrates. "Gene knockout" mice, for instance, provide an experimental system for investigating the complexities of mammalian metabolism.

In a similar fashion, investigating the actions of metabolic inhibitors has helped identify individual steps in metabolic pathways. The inhibition of one step of a pathway affects the entire pathway. Because the substrate of the inhibited enzyme accumulates, it can be isolated and characterized more easily. Intermediates formed in steps preceding the site of inhibition also accumulate. The use of inhibitory drugs not only helps in the study of metabolism but also determines the mechanism of action of the drug, often leading to improved drug variations.

Summary

1. The chemical reactions carried out by living cells are collectively called metabolism. Sequences of reactions are called pathways. Degradative (catabolic) and synthetic (anabolic) pathways proceed in discrete steps.
2. Metabolic pathways are regulated to allow an organism to respond to changing demands. Individual enzymes are commonly regulated by allosteric modulation or reversible covalent modification.
3. The major catabolic pathways convert macromolecules to smaller, energy-yielding metabolites. The energy released in catabolic reactions is conserved in the form of ATP, GTP, and reduced coenzymes.
4. Within a cell or within a multicellular organism, metabolic processes are sequestered.
5. Metabolic reactions are in a steady state. If the steady state concentration of reactants and products is close to the equilibrium ratio the reaction is said to be a near-equilibrium reaction. If the steady state concentrations are far from equilibrium the reaction is said to be metabolically irreversible.
6. The actual free energy change (ΔG) of a reaction inside a cell differs from the standard free energy change ($\Delta G^{\circ\prime}$).
7. Hydrolytic cleavage of the phosphoanhydride groups of ATP releases large amounts of free energy.
8. The energy of ATP is made available when a terminal phosphoryl group or a nucleotidyl group is transferred. Some metabolites with high phosphoryl group transfer potentials can transfer their phosphoryl groups to ADP to produce ATP. Such metabolites are called energy-rich compounds.
9. Thioesters, such as acyl coenzyme A, can donate acyl groups and can sometimes also generate ATP equivalents.
10. The free energy of biological oxidation reactions can be captured in the form of reduced coenzymes. This form of energy is measured as the difference in reduction potentials.
11. Metabolic pathways are studied by characterizing their enzymes, intermediates, flux, and regulation.

Problems

1. A biosynthetic pathway proceeds from compound A to compound E in four steps and then branches. One branch is a two-step pathway to G, and the other is a three-step pathway to J. Substrate A is a feed-forward activator of the enzyme that catalyzes the synthesis of E. Products G and J are feedback inhibitors of the initial enzyme in the common pathway, and they also inhibit the first enzyme after the branch point in their own pathways.

(a) Draw a diagram showing the regulation of this metabolic pathway.

(b) Why is it advantageous for each of the two products to inhibit two enzymes in the pathway?

2. Glucose degradation can be accomplished by a combination of the glycolytic and citric acid pathways. The enzymes for glycolysis are located in the cytosol, while the enzymes for the citric acid cycle are located in the mitochondria. What are two advantages in separating the enzymes for these major carbohydrate degradation pathways in different cellular compartments?

3. In bacteria, the glycolytic and citric acid cycle pathways are both cytosolic. Why don't the "advantages" in Question 2 apply to bacteria?

4. In multistep metabolic pathways, enzymes for successive steps may be associated with each other in multienzyme complexes or be bound in close proximity on membranes. Explain the major advantage of having enzymes organized in either of these associations.

5. (a) Calculate the K_{eq} at 25°C and pH 7.0 for the following reaction using the data in Table 10.4.

$$\text{Glycerol 3-phosphate} + H_2O \rightarrow \text{glycerol} + P_i.$$

(b) The final step in the pathway for the synthesis of glucose from lactate (gluconeogenesis) is:

$$\text{Glucose 6-P} + H_2O \rightarrow \text{glucose} + P_i.$$

When glucose 6-P is incubated with the proper enzyme and the reaction runs until equilibrium has been reached, the final concentrations are found to be: glucose 6-P (0.035 mM), glucose (100 mM), and P_i (100 mM). Calculate $\Delta G^{o\prime}$ at 25°C and pH 7.0.

6. $\Delta G^{o'}$ for the hydrolysis of phosphoarginine is −32 kJ mol^{-1}.

(a) What is the actual free energy change for the reaction at 25°C and pH 7.0 in resting lobster muscle, where the concentrations of phosphoarginine, arginine, and P_i are 6.8 mM, 2.6 mM, and 5 mM, respectively?

(b) Why does this value differ from $\Delta G^{o\prime}$?

(c) High-energy compounds have large negative free energies of hydrolysis, indicating that their reactions with water proceed almost to completion. How can millimolar concentrations of acetyl CoA, whose $\Delta G^{o\prime}_{\text{hydrolysis}}$ is −32 kJ mol^{-1}, exist in cells?

7. Glycogen is synthesized from glucose-1-phosphate. Glucose-1-phosphate is activated by a reaction with UTP, forming UDP-glucose and pyrophosphate (PP_i).

$$\text{Glucose-1-phosphate} + \text{UTP} \rightarrow \text{UDP-glucose} + PP_i$$

UDP-glucose is the substrate for the enzyme glycogen synthase which adds glucose molecules to the growing carbohydrate chain. The $\Delta G^{o\prime}$ value for the condensation of UTP with glucose-1-phosphate to form UDP-glucose is approximately 0 kJ mol^{-1}. The PP_i that is released is rapidly hydrolyzed by inorganic pyrophosphatase. Determine the overall $\Delta G^{o\prime}$ value if the formation of UDP-glucose is coupled to the hydrolysis of PP_i.

8. (a) A molecule of ATP is usually consumed within a minute after synthesis, and the average human adult requires about 65 kg of ATP per day. Since the human body contains only about 50 grams of ATP and ADP combined, how it is possible that so much ATP can be utilized?

(b) Does ATP have a role in energy storage?

9. Phosphocreatine is produced from ATP and creatine in mammalian muscle cells at rest. What ATP/ADP ratio is necessary to maintain a phosphocreatine/creatine ratio of 20:1? (To maintain the coupled reaction at equilibrium, the actual free energy change must be zero.)

10. Amino acids must be covalently attached to a ribose hydroxyl group on the correct tRNA (transfer RNA) prior to recognition and insertion into a growing polypeptide chain. The overall reaction carried out by the amino acyl tRNA synthetase enzymes is:

$$\text{Amino acid} + \text{HO-tRNA} + \text{ATP} \longrightarrow \text{amino acyl-O-tRNA} + \text{AMP} + 2P_i$$

Assuming this reaction proceeds through an acyl adenylate intermediate, write all the steps involved in this enzyme-catalyzed reaction.

11. When a mixture of glucose 6-phosphate and fructose 6-phosphate is incubated with the enzyme phosphohexose isomerase, the final mixture contains twice as much glucose 6-phosphate as fructose 6-phosphate. Calculate the value of $\Delta G^{o\prime}$.

$$\text{Glucose 6-phosphate} \longleftrightarrow \text{fructose 6-phosphate}$$

12. Coupling ATP hydrolysis to a thermodynamically unfavorable reaction can markedly shift the equilibrium of the reaction.

(a) Calculate K_{eq} for the energetically unfavorable biosynthetic reaction A → B when $\Delta G^{o'} = +25$ kJ mol^{-1} at 25°C.

(b) Calculate K_{eq} for the reaction A → B when it is coupled to the hydrolysis of ATP. Compare this value to the value in Part (a).

(c) Many cells maintain [ATP]/[ADP] ratios of 400 or more. Calculate the ratio of [B] to [A] when [ATP]: [ADP] is 400:1 and [P_i] is constant at standard conditions. How does this ratio compare to the ratio of [B] to [A] in the uncoupled reaction?

13. Using data from Table 10.5, write the coupled reaction that would occur spontaneously for the following pairs of molecules under standard conditions:

(a) Cytochrome *f* and cytochrome b_5

(b) Fumarate/succinate and ubiquinone/ubiquinol (Q/QH_2)

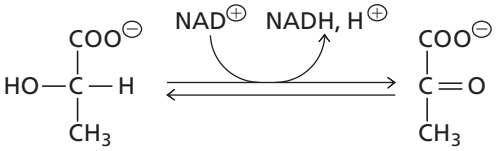
(c) α-ketoglutarate/isocitrate and $NAD^{\oplus}$/NADH

14. Using data from Table 10.5, calculate the standard reduction potential and the standard free energy change for each of the following oxidation–reduction reactions:

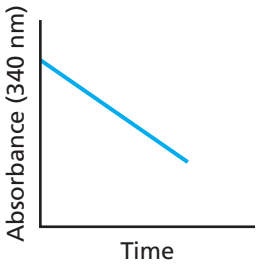
(a) $\text{Ubiquinol } (QH_2) + 2 \text{ cytochrome c } (Fe^{3+}) \rightleftharpoons \text{ubiquinone } (Q) + 2 \text{ cytochrome c } (Fe^{2+}) + 2\,H^{\oplus}$

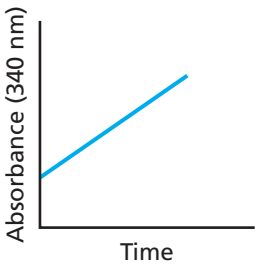
(b) $\text{Succinate} + {}^1\!/_2 O_2 \rightleftharpoons \text{fumarate} + H_2O$

15. Lactate dehydrogenase is an NAD-dependent enzyme that catalyzes the reversible oxidation of lactate.



Initial reaction rates are followed spectrophotometrically at 340 nm after addition of lactate, NAD⊕, lactate dehydrogenase, and buffer to the reaction vessel. When the change in absorbance at 340 nm is monitored over time, which graph is representative of the expected results? Explain.





16. Using the standard reduction potentials for Q and FAD in Table 10.5, show that the oxidation of $FADH_2$ by Q liberates enough energy to drive the synthesis of ATP from ADP and P_i under cellular conditions where $[FADH_2] = 5$ mM, $[FAD] = 0.2$ mM, $[Q] = 0.1$ mM, and $[QH_2] = 0.05$ mM. Assume that ΔG for ATP synthesis from ADP and P_i is +30 kJ mol^{-1}.

Selected Readings

Alberty, R. A. (1996). Recommendations for nomenclature and tables in biochemical thermodynamics. *Eur. J. Biochem.* 240:1–14.

Alberty, R. A. (2000). Calculating apparent equilibrium constants of enzyme-catalyzed reactions at pH 7. *Biochem. Educ.* 28:12–17.

Burbaum, J. J., Raines, R. T., Albery, W. J., and Knowles, J. R. (1989). Evolutionary optimization of the catalytic effectiveness of an enzyme. *Biochem.* 28:9293–9305.

Edwards, R. A. (2001). The free energies of metabolic reactions (ΔG) are not positive. *Biochem. Mol. Bio. Educ.* 29:101–103.

Hayes, D. M., Kenyon, G. L., and Kollman, P. A. (1978). Theoretical calculations of the hydrolysis energies of some "high-energy" molecules. 2. A survey of some biologically important hydrolytic reactions. *J. Am. Chem. Soc.* 100:4331–4340.

Schmidt. S., Sunyaev, S., Bork. P., and Dandekar, T. (2003). Metabolites: a helping hand for pathway evolution? *Trends Biochem. Sci.* 28:336–341.

Silverstein, T. (2005). Redox redox: a response to Feinman's "Oxidation–reduction calculations in the biochemistry course." *Biochem. Mol. Bio. Educ.* 33:252–253.

Tohge, T., Nunes-Nesi, A., and Fernie, A. R. (2009). Finding the paths: metabolomics and approaches to metabolic flux analysis. *The Biochem. Soc.* (June 2009):8–12.

Yus, E., et al. (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science* 326:1263–1272.





Glycolysis

The first three metabolic pathways we examine are central to both carbohydrate metabolism and energy generation. Gluconeogenesis is the main pathway for synthesis of hexoses from three carbon precursors. As the name of the pathway indicates, glucose is the primary end product of gluconeogenesis. This biosynthetic pathway will be described in the next chapter. Glucose, and other hexoses, can be the precursors for synthesis of many complex carbohydrates. Glucose can also be degraded in a catabolic glycolytic pathway with recovery of the energy used in its synthesis. In glycolysis, the subject of this chapter, glucose is converted to the three-carbon acid pyruvate. Pyruvate has several possible fates, one of which is oxidative decarboxylation to form acetyl CoA. The third pathway is the citric acid cycle, described in Chapter 13. This is the route by which the acetyl group of acetyl CoA is oxidized to carbon dioxide and water. One of the important intermediates in the citric acid cycle, oxaloacetate, is also an intermediate in the synthesis of glucose from pyruvate. Figure 11.1 shows the relationship among the three pathways. All three pathways play a role in the formation and degradation of noncarbohydrate molecules such as amino acids and lipids.

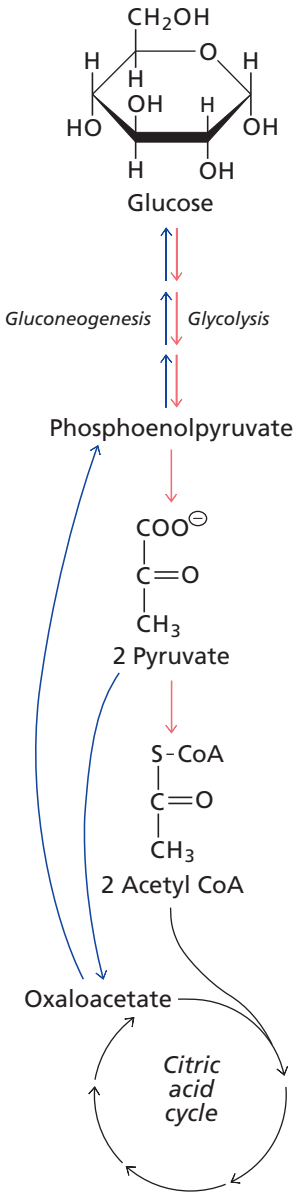
We present the reactions of glycolysis, gluconeogenesis, and the citric acid cycle in more detail than those of other metabolic pathways in this book but the same principles apply to all pathways. We introduce many biomolecules and enzymes, some of which appear in more than one pathway. Keep in mind that the chemical structures of the metabolites prompt the enzyme names and that the names of the enzymes reflect the substrate specificity and the type of reaction catalyzed. A confident grasp of terminology will prepare you to enjoy the chemical elegance of metabolism. However, do not lose sight of the major concepts and general strategies of metabolism while memorizing the details. The names of particular enzymes might fade over time but we hope you will retain an understanding of the patterns and purposes behind the interconversion of metabolites in cells.

In this book we follow the tradition of presenting glycolysis as our first metabolic pathway. The catabolism of glucose is a major source of energy in animals. The details of the various reactions, and their regulation, are well known.

The glycolytic sequence of reactions is perhaps the best understood and most studied multi-enzyme system of the cell. The pattern of interplay between enzymes and substrates in this relatively simple multi-enzyme system applies to all the multi-enzyme systems of the cell, especially the very complex systems involved in respiration and photosynthesis.

—Albert Lehninger (1965), *Bioenergetics*, p. 75

Top: Wine, beer and bread. For centuries, wineries, breweries, and bakeries have exploited the basic biochemical pathway of glycolysis where glucose is converted to ethanol and CO_2.



▲ **Figure 11.1**
Gluconeogenesis, glycolysis, and the citric acid cycle. Glucose is synthesized from pyruvate via oxaloacetate and phosphoenolpyruvate. In glycolysis, glucose is degraded to pyruvate. Many (but not all) of the steps in glycolysis are the reverse of the gluconeogenesis reactions. The acetyl group of pyruvate is transferred to coenzyme A (CoA) and oxidized to carbon dioxide by the citric acid cycle. Energy in the form of ATP equivalents is required for the synthesis of glucose. Some of this energy is recovered in glycolysis but much more is recovered as a result of the citric acid cycle.

KEY CONCEPT

The main energy gain in glycolysis is due to production of NADH molecules.

11.1 The Enzymatic Reactions of Glycolysis

Glycolysis is a sequence of ten enzyme-catalyzed reactions by which glucose is converted to pyruvate (Figure 11.2 on page 328). The conversion of one molecule of glucose to two molecules of pyruvate is accompanied by the net conversion of two molecules of ADP to two molecules of ATP and the reduction of two molecules of $NADH^{\oplus}$ to two molecules NADH. The enzymes of this pathway are found in most living species and are located in the cytosol. The glycolytic pathway is active in all differentiated cell types in multicellular organisms. In some mammalian cells (such as those in the retina and some brain cells), it is the only ATP-producing pathway.

The net reaction of glycolysis is shown in Reaction 11.1.

$$\text{Glucose} + 2\ \text{ADP} + 2\ \text{NAD}^{\oplus} + 2\ \text{P}_i \rightarrow 2\ \text{Pyruvate} + 2\ \text{ATP} + 2\ \text{NADH} + 2\ \text{H}^{\oplus} + 2\ \text{H}_2\text{O} \qquad (11.1)$$

The ten reactions of glycolysis are listed in Table 11.1. They can be divided into two stages: the hexose stage and the triose stage. The left page of Figure 11.2 shows the hexose stage. At Step 4, the C-3—C-4 bond of the hexose is cleaved to produce two trioses. From that point on the intermediates of the pathway are triose phosphates. Two triose phosphates are formed from fructose 1,6-*bis*phosphate. Dihydroxyacetone phosphate is converted to glyceraldehyde 3-phosphate in Step 5 and glyceraldehyde 3-phosphate continues through the pathway. All subsequent steps of the triose stage of glycolysis (right page of Figure 11.2) are traversed by two molecules for each molecule of glucose metabolized.

Two molecules of ATP are converted to ADP in the hexose stage of glycolysis. In the triose stage, four molecules of ATP are formed from ADP for each molecule of glucose metabolized. Thus, glycolysis has a net yield of two molecules of ATP per molecule of glucose.

ATP consumed per glucose:	2 (hexose stage)	
ATP produced per glucose:	4 (triose stage)	(11.2)
Net ATP production per glucose:	2	

The first and third reactions of glycolysis are coupled to the utilization of ATP. These priming reactions help drive the pathway in the direction of glycolysis since the reverse reactions are thermodynamically favored in the absence of ATP. Two later intermediates of glycolysis have sufficient group transfer potentials to allow the transfer of a phosphoryl group to ADP producing ATP (Steps 7 and 10). Step 6 is coupled to the synthesis of reducing equivalents in the form of NADH. Each molecule of NADH is equivalent to several molecules of ATP (Section 10.9) so the net energy gain in glycolysis is mostly due to production of NADH.

11.2 The Ten Steps of Glycolysis

Now we examine the chemistry and enzymes of each glycolytic reaction. As you read, pay attention to the chemical logic and economy of the pathway. Consider how each chemical reaction prepares a substrate for the next step in the process. Note, for example, that a cleavage reaction converts a hexose to two trioses, not to a two-carbon compound and a tetrose. The two trioses rapidly interconvert allowing both products of the cleavage reaction to be further metabolized by the action of one set of enzymes, not two. Finally, be aware of how ATP is both consumed and produced in glycolysis. We have already seen a number of examples of the transfer of the chemical potential energy of ATP (e.g., in Section 10.7) but the reactions in this chapter are our first detailed examples of how the energy released by oxidation reactions is captured for use in other biochemical pathways.

1. Hexokinase

In the first reaction of glycolysis, the γ-phosphoryl group of ATP is transferred to the oxygen atom at C-6 of glucose producing glucose 6-phosphate and ADP (Figure 11.3 on

Table 11.1 The reactions and enzymes of glycolysis

1. Glucose + ATP ⟶ Glucose 6-phosphate + ADP + $H^{\oplus}$	Hexokinase, glucokinase
2. Glucose 6-phosphate ⇌ Fructose 6-phosphate	Glucose-6-phosphate isomerase
3. Fructose 6-phosphate + ATP ⟶ Fructose 1,6-*bis*phosphate + ADP + $H^{\oplus}$	Phosphofructokinase-1
4. Fructose 1,6-*bis*phosphate ⇌ Dihydroxyacetone phosphate + Glyceraldehyde 3-phosphate	Aldolase
5. Dihydroxyacetone phosphate ⇌ Glyceraldehyde 3-phosphate	Triose phosphate isomerase
6. Glyceraldehyde 3-phosphate + $NAD^{\oplus}$ + P_i ⇌ 1,3-*Bis*phosphoglycerate + NADH + $H^{\oplus}$	Glyceraldehyde 3-phosphate dehydrogenase
7. 1,3-*Bis*phosphoglycerate + ADP ⇌ 3-Phosphoglycerate + ATP	Phosphoglycerate kinase
8. 3-Phosphoglycerate ⇌ 2-Phosphoglycerate	Phosphoglycerate mutase
9. 2-Phosphoglycerate ⇌ Phosphoenolpyruvate + H_2O	Enolase
10. Phosphoenolpyruvate + ADP + $H^{\oplus}$ ⟶ Pyruvate + ATP	Pyruvate kinase

page 330). This phosphoryl group transfer reaction is catalyzed by hexokinase. Kinases catalyze four reactions in the glycolytic pathway—Steps 1, 3, 7, and 10.

The hexokinase reaction is regulated making it a metabolically irreversible reaction. Cells need to maintain a relatively high concentration of glucose 6-phosphate and a low internal concentration of glucose. As we'll see in Section 11.5B, the reverse reaction is inhibited by glucose 6-phosphate. Hexokinases from yeast and mammalian tissues have been thoroughly studied. These enzymes have a broad substrate specificity; they catalyze the phosphorylation of glucose and mannose, and of fructose when it is present at high concentrations.

Multiple forms, or **isozymes**, of hexokinase occur in many eukaryotic cells. (Isozymes are different proteins from one species that catalyze the same chemical reaction.) Four hexokinase isozymes have been isolated from mammalian liver. All four are found in varying proportions in other mammalian tissues. These isozymes catalyze the same reaction but have different K_m values for glucose. Hexokinases I, II, and III have K_m values of about 10^{-6} to 10^{-4} M, whereas hexokinase IV, also called glucokinase, has a much higher K_m value for glucose (about 10^{-2} M). In eukaryotes, glucose is taken up and secreted by passive transport using various glucose transporters (GLUT). The concentration of glucose in the blood and the cell cytoplasm is usually below the K_m of glucokinase for glucose. At these low concentrations the other hexokinase isozymes catalyze the phosphorylation of glucose. With high glucose levels, glucokinase is active. Because glucokinase is never saturated with glucose, the liver can respond to large increases in blood glucose by phosphorylating it for entry into glycolysis or the glycogen synthesis pathway.

In most bacteria, the uptake of glucose is coupled to the phosphorylation of glucose to glucose 6-phosphate via the phosphoenolpyruvate sugar transport system (Section 21.7B). The phosphoryl group is donated by phosphoenolpyruvate. Hexokinases and glucokinases can be found in bacteria but they play a minor role in glycolysis because, unlike the situation in eukaryotic cells, the bacterial enzymes rarely encounter free glucose in their cytoplasm.

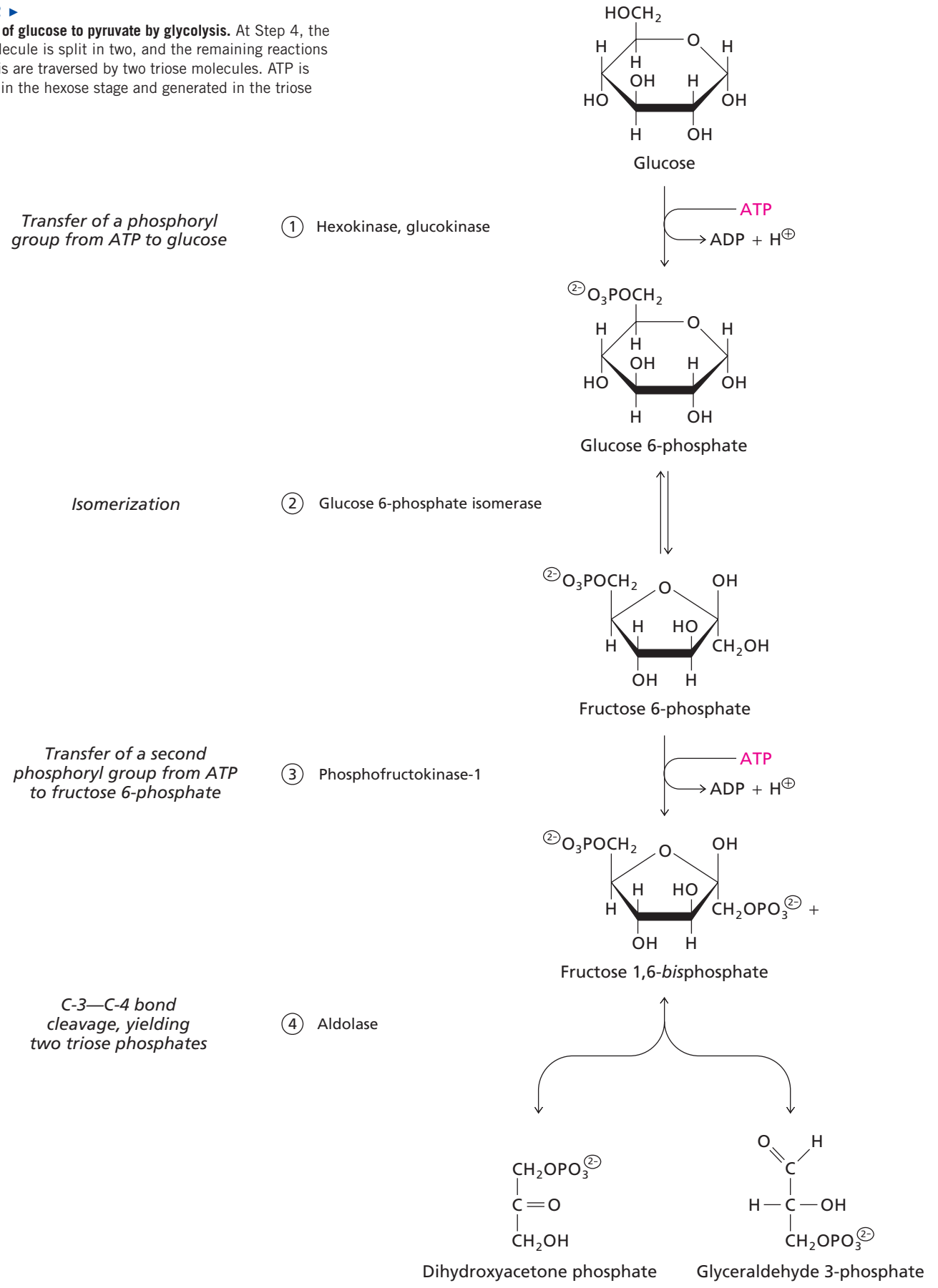
2. Glucose 6-Phosphate Isomerase

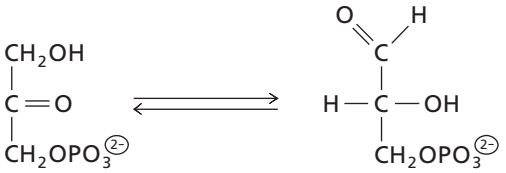
In the second step of glycolysis, glucose 6-phosphate isomerase catalyzes the conversion of glucose 6-phosphate (an aldose) to fructose 6-phosphate (a ketose), as shown in Figure 11.4. The enzyme is also known as phosphoglucose isomerase (PGI). Isomerases interconvert aldoses and ketoses that have identical configurations at all other chiral atoms.

The α anomer of glucose 6-phosphate (α-D-glucopyranose 6-phosphate) preferentially binds to glucose 6-phosphate isomerase. The open-chain form of glucose 6-phosphate is then generated within the active site of the enzyme, and an aldose-to-ketose conversion occurs. The open-chain form of fructose 6-phosphate cyclizes to form α-D-fructofuranose 6-phosphate. The mechanism of glucose 6-phosphate isomerase is similar to the mechanism of triose phosphate isomerase (Section 6.4A).

Glucose 6-phosphate isomerase is highly stereospecific. For example, in the reverse reaction catalyzed by this enzyme fructose 6-phosphate (in which C-2 is not chiral) is

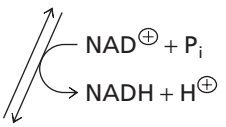
Figure 11.2 ►
Conversion of glucose to pyruvate by glycolysis. At Step 4, the hexose molecule is split in two, and the remaining reactions of glycolysis are traversed by two triose molecules. ATP is consumed in the hexose stage and generated in the triose stage.

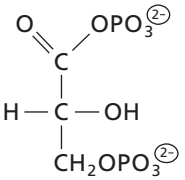




Dihydroxyacetone phosphate ⇌ Glyceraldehyde 3-phosphate

Triose phosphate isomerase (5) — *Rapid interconversion of triose phosphates*

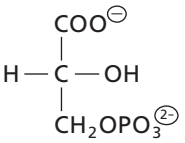




1,3-*Bis*phosphoglycerate

Glyceraldehyde 3-phosphate dehydrogenase (6) — *Oxidation and phosphorylation, yielding a high-energy mixed-acid anhydride*



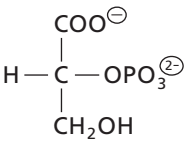


3-Phosphoglycerate

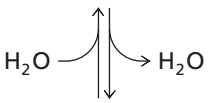
Phosphoglycerate kinase (7) — *Transfer of a high-energy phosphoryl group to ADP, yielding ATP*



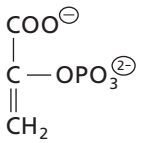
Phosphoglycerate mutase (8) — *Intramolecular phosphoryl-group transfer*



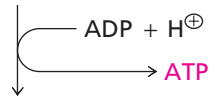
2-Phosphoglycerate



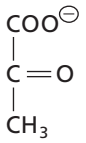
Enolase (9) — *Dehydration to an energy-rich enol ester*



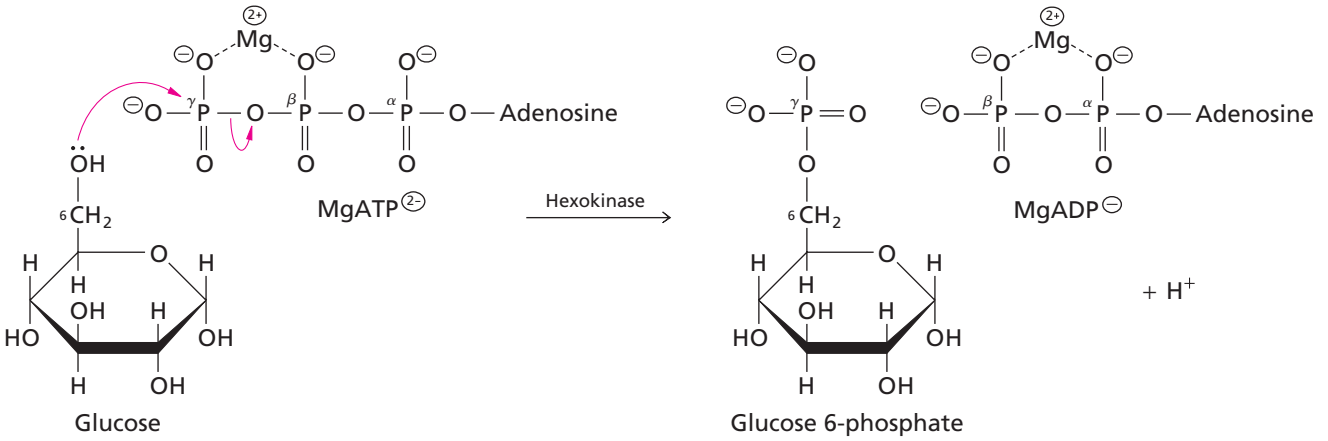
Phosphoenolpyruvate



Pyruvate kinase (10) — *Transfer of a high-energy phosphoryl group to ADP, yielding ATP*



Pyruvate



▲ **Figure 11.3**
Phosphoryl group transfer reaction catalyzed by hexokinase. This reaction occurs by attack of the C-6 hydroxyl oxygen of glucose on the γ-phosphorus of MgATP^{2-}. MgADP^{-} is displaced, and glucose 6-phosphate is generated. All four kinases in glycolysis catalyze direct nucleophilic attack of a hydroxyl group on the terminal phosphoryl group of ATP (and/or its reverse under cellular conditions). (Mg^{2+}, shown explicitly here, is also required in the other kinase reactions in this chapter, although it is not shown for those reactions.)

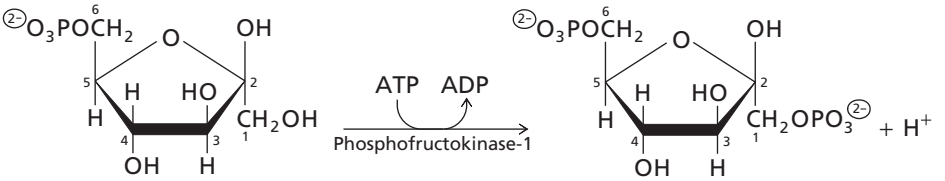
The hexokinase mechanism is a classic example of induced fit (Section 6.5C).

converted almost exclusively to glucose 6-phosphate. Only traces of mannose 6-phosphate, the C-2 epimer of glucose 6-phosphate, are formed.

The glucose 6-phosphate isomerase reaction is a near-equilibrium reaction. The reverse reaction is part of the pathway for the biosynthesis of glucose.

3. Phosphofructokinase-1

Phosphofructokinase-1 (PFK-1) catalyzes the transfer of a phosphoryl group from ATP to the C-1 hydroxyl group of fructose 6-phosphate producing fructose 1,6-*bis*phosphate. The "*bis*" in *bis*phosphate indicates that the two phosphoryl groups are attached to different carbon atoms (cf. diphosphate).



(11.3)

Note that the reaction catalyzed by glucose 6-phosphate isomerase produces α-D-fructose 6-phosphate. However, it is the β-D anomer that is the substrate for the next step in glycolysis—the one catalyzed by phosphofructokinase-1. The α and β anomers of fructose 6-phosphate equilibrate spontaneously (Section 8.2). This interconversion is extremely rapid in aqueous solution and has no effect on the overall rate of glycolysis.

We discuss the regulation of glycolysis in detail in Section 11.5.

The reaction catalyzed by PFK-1 is metabolically irreversible indicating that the activity of the enzyme is regulated. In fact, this step is a critical control point for the regulation of glycolysis in most species. The PFK-1 catalyzed reaction is the first *committed* step of glycolysis because some substrates other than glucose can enter the glycolytic pathway by direct conversion to fructose 6-phosphate, thus bypassing the steps catalyzed by hexokinase and glucose 6-phosphate isomerase (Section 11.6C). (The metabolically irreversible reaction catalyzed by hexokinase is *not* the first committed step.) Another reason for regulating PFK-1 activity has to do with the competing glycolysis and gluconeogenesis pathways (Figure 11.1). PFK-1 activity must be inhibited when glucose is being synthesized.

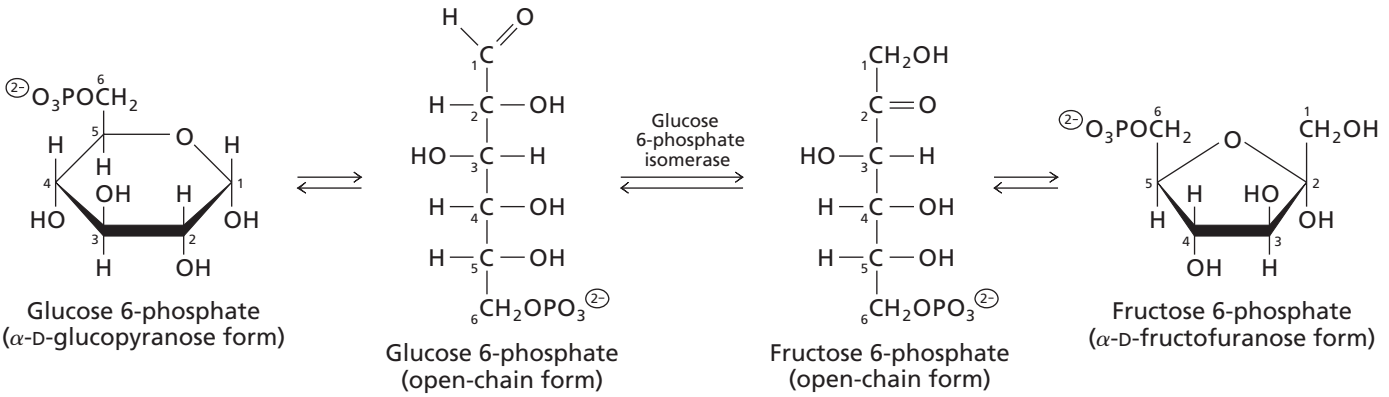
PFK-1 is one of the classic allosteric enzymes. Recall that the bacterial enzyme is activated by ADP and allosterically inhibited by phosphoenolpyruvate (Section 5.10A). The activity of the mammalian enzyme is regulated by AMP and citrate (Section 11.6C).

PFK-1 has the suffix "1" because there is a second phosphofructokinase that catalyzes the synthesis of fructose 2,6-*bis*phosphate instead of fructose 1,6-*bis*phosphate. This second enzyme, which we will encounter later in this chapter, is known as PFK-2.

We explore glycogen synthesis in Section 12.5.

4. Aldolase

The first three steps of glycolysis prepare the hexose for cleavage into two triose phosphates, glyceraldehyde 3-phosphate and dihydroxyacetone phosphate.



▲ Figure 11.4
Conversion of glucose 6-phosphate to fructose 6-phosphate. This aldose-ketose isomerization is catalyzed by glucose 6-phosphate isomerase.

Dihydroxyacetone phosphate (DHAP) is derived from C-1 to C-3 of fructose 1,6-*bis*phosphate, and glyceraldehyde 3-phosphate (GAP) is derived from C-4 to C-6. The enzyme that catalyzes the cleavage reaction is fructose 1,6-*bis*phosphate aldolase, commonly shortened to aldolase. Aldol cleavage is a common mechanism for cleaving C—C bonds in biological systems and for C—C bond formation in the reverse direction.

BOX 11.1 A BRIEF HISTORY OF THE GLYCOLYTIC PATHWAY

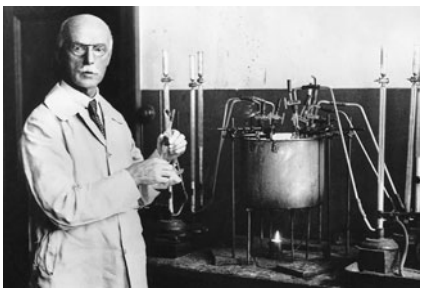
Glycolysis was one of the first metabolic pathways to be elucidated. It played an important role in the development of biochemistry. In 1897, Eduard Buchner (Section 1.1) discovered that bubbles of carbon dioxide were released from a mixture of sucrose and a cell-free yeast extract. He concluded that fermentation was occurring in his cell-free extract. More than 20 years earlier, Louis Pasteur had shown that yeast cells ferment sugar to alcohol (i.e., produce ethanol and CO_2) but Buchner showed that intact cells were not required. Buchner named the fermenting activity zymase. Today, we recognize that the zymase of yeast extracts is not a single enzyme but a mixture of enzymes that together catalyze the reactions of glycolysis.

The steps of the glycolytic pathway were gradually discovered by analyzing the reactions catalyzed by extracts of yeast or muscle. In 1905, Arthur Harden and William John Young found that when the rate of glucose fermentation by yeast extract decreased it could be restored by adding inorganic phosphate. Harden and Young assumed that phosphate derivatives of glucose were being formed. They succeeded in isolating fructose 1,6-*bis*phosphate and showed that it is an intermediate in the fermentation of glucose because it too is fermented by cell-free yeast extracts. Harden was awarded the Nobel Prize in Chemistry in 1929 for his work on glycolysis.

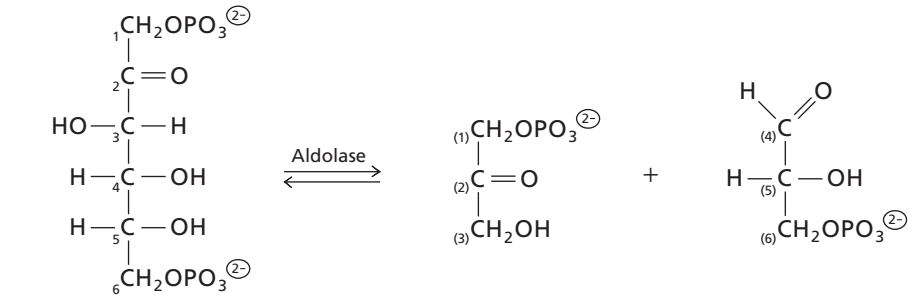
By the 1940s, the complete glycolytic pathway in eukaryotes—including its enzymes, intermediates, and coenzymes—was known. The further characterization of individual enzymes and studies of the regulation of glycolysis and its integration with other pathways have taken many more years. In bacteria, the classic glycolytic pathway is called the Embden–Meyerhof–Parnas pathway after Gustav Embden (1874–1933), Otto Meyerhof (1884–1951), and Jacob Parnas (1884–1949). The bacterial pathway differs in some minor ways from the eukaryotic pathway. In 1922 Meyerhof was awarded the Nobel Prize in Physiology or Medicine for his work on the production of lactic acid in muscle cells.



▲ Louis Pasteur (1822–1895).



▲ Arthur Harden (1865–1940).



(11.4)

There are two distinct classes of aldolases. class I enzymes are found in plants and animals; class II enzymes are more common in bacteria, fungi, and protists. Many species have both types of enzyme. Class I and class II aldolases are unrelated. The enzymes have very different structures and sequences in spite of the fact that they catalyze the same reaction. This is an example of convergent evolution.

The two classes of aldolase have slightly different mechanisms. Class I aldolases involve formation of a covalent Schiff base between lysine and pyruvate derivatives (Section 6.3) and class II aldolases use a metal ion cofactor (Figures 11.5 and 11.6).

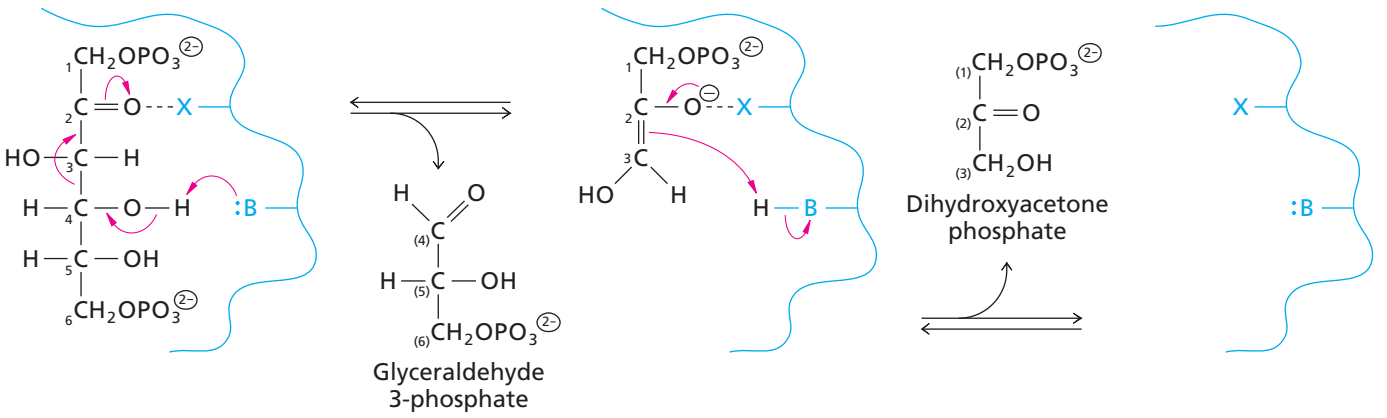
The standard Gibbs free energy change for this reaction is strongly positive ($\Delta G^{\circ\prime} = +28$ kJ mol^{-1}). Nevertheless, the aldolase reaction is a near-equilibrium reaction (actual $\Delta G \cong 0$) in cells where glycolysis is an important catabolic pathway. This means that the concentration of fructose 1,6-*bis*phosphate is very high relative to the two trioses. (But see Problem 10).

The key to understanding the strategy of glycolysis lies in appreciating the significance of the aldolase reaction. It's best to think of this as a near-equilibrium biosynthesis reaction and not a degradation reaction. Aldolases evolved originally as enzymes that could catalyze the synthesis of fructose 1,6-*bis*phosphate. This reaction occurred at the end of a biosynthesis pathway leading from pyruvate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate.

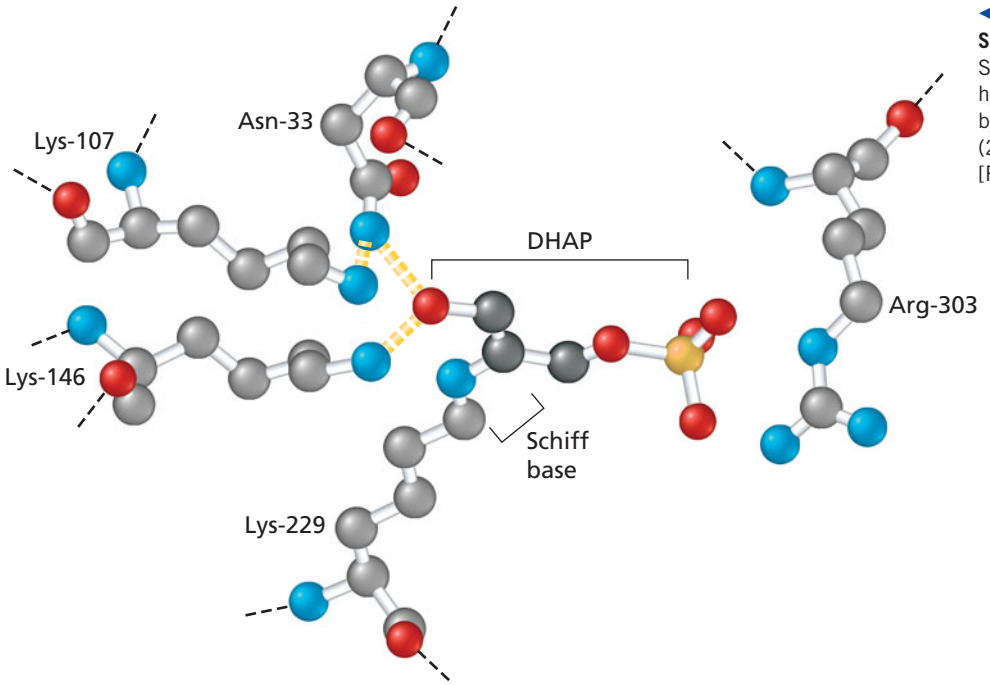
During glycolysis, flux in the triose stage is in the opposite direction—toward pyruvate synthesis. The first steps of glycolysis—the hexose stage—are directed toward formation of fructose 1,6-*bis*phosphate so that it can serve as substrate for the reversal of the pathway leading to its synthesis. Keep in mind that the glucose biosynthesis pathway (gluconeogenesis) evolved first. It was only after glucose became readily available that pathways for its degradation evolved.

5. Triose Phosphate Isomerase

Of the two molecules produced by the splitting of fructose 1,6-*bis*phosphate, only glyceraldehyde 3-phosphate is a substrate for the next reaction in the glycolytic pathway.

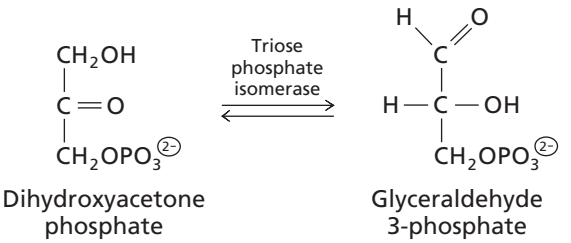


▼ Figure 11.5
Mechanism of aldol cleavage catalyzed by aldolases. Fructose 1,6-*bis*phosphate is the aldol substrate. Aldolases have an electron-withdrawing group (—X) that polarizes the C-2 carbonyl group of the substrate. Class I aldolases use the amino group of a lysine residue at the active site, and the other class II aldolases use Zn^{2+} for this purpose. A basic residue (designated —B:) removes a proton from the C-4 hydroxyl group of the substrate.



◀ **Figure 11.6**
Schiff base in the active site of aldolase. A Schiff base forms between Lys-229 and dihydroxyacetone during the reaction catalyzed by aldolase. Modified after St-Jean et al. (2009). (Hydrogen atoms not shown.) [PEB 3DFO]

The other product, dihydroxyacetone phosphate, is converted to glyceraldehyde 3-phosphate in a near-equilibrium reaction catalyzed by triose phosphate isomerase.



(11.5)

As glyceraldehyde 3-phosphate is consumed in Step 6, its steady state concentration is maintained by flux from dihydroxyacetone phosphate. In this way, two molecules of glyceraldehyde 3-phosphate are supplied to glycolysis for each molecule of fructose 1,6-*bis*phosphate split. Triose phosphate isomerase catalyzes a stereospecific reaction so that only the D isomer of glyceraldehyde 3-phosphate is formed.

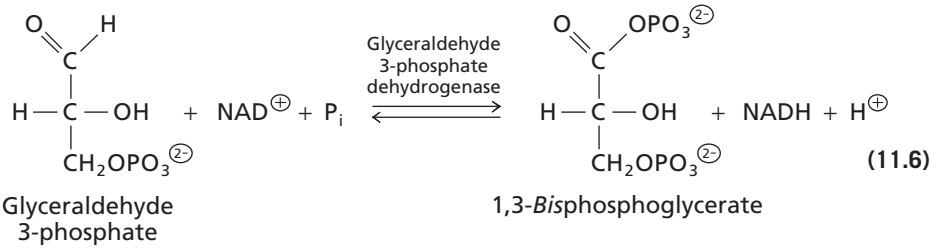
Triose phosphate isomerase, like glucose 6-phosphate isomerase, catalyzes an aldose-to-ketose conversion. The mechanism of the triose phosphate isomerase reaction is described in Section 6.4A. The catalytic mechanisms of aldose–ketose isomerases have been studied extensively, and the formation of an enzyme-bound enediolate intermediate appears to be a common feature.

The rate of the triose phosphate isomerase reaction is close to the theoretical limit for a diffusion controlled reaction.

The fate of the individual carbon atoms of a molecule of glucose is shown in Figure 11.7. This distribution has been confirmed by radioisotopic tracer studies in a variety of organisms. Note that carbons 1, 2, and 3 of one molecule of glyceraldehyde 3-phosphate are derived from carbons 4, 5, and 6 of glucose, whereas carbons 1, 2, and 3 of the second molecule of glyceraldehyde 3-phosphate (converted from dihydroxyacetone phosphate) originate as carbons 3, 2, and 1 of glucose. When these molecules of glyceraldehyde 3-phosphate mix to form a single pool of metabolites, a carbon atom from C-1 of glucose can no longer be distinguished from a carbon atom from C-6 of glucose.

6. Glyceraldehyde 3-Phosphate Dehydrogenase

The recovery of energy from triose phosphates begins with the reaction catalyzed by glyceraldehyde 3-phosphate dehydrogenase. In this step, glyceraldehyde 3-phosphate is oxidized and phosphorylated to produce 1,3-*bis*phosphoglycerate.



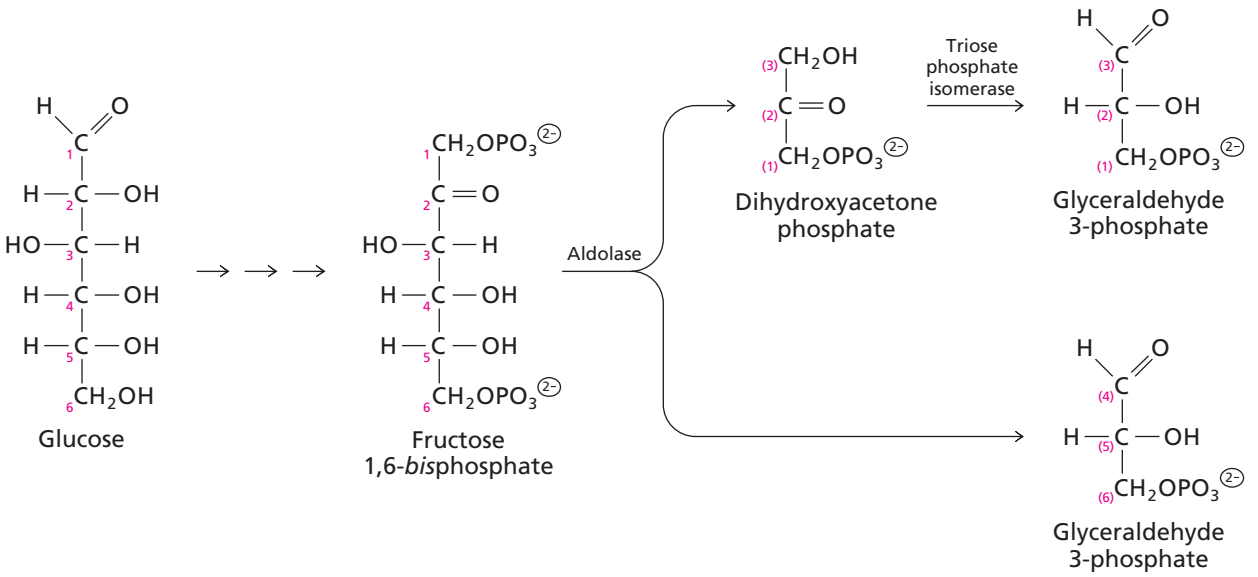
This is an oxidation–reduction reaction; the oxidation of glyceraldehyde 3-phosphate is coupled to the reduction of $NAD^{\oplus}$ to NADH. In some species the coenzyme is $NADP^{\oplus}$.

The oxidation of the aldehyde group of glyceraldehyde 3-phosphate proceeds with a large negative Gibbs standard free energy change, and some of this energy is conserved in the acid–anhydride linkage of 1,3-*bis*phosphoglycerate. In the next step of glycolysis, the C-1 phosphoryl group of 1,3-*bis*phosphoglycerate is transferred to ADP to form ATP. The remaining energy is conserved in the form of reducing equivalents (NADH). As we saw in the previous chapter, each molecule of NADH is equivalent to several molecules of ATP. Thus, this step of glycolysis is the main energy-producing step in the entire pathway.

The overall standard Gibbs free energy change (oxidation of the aldehyde and reduction of $NAD^{\oplus}$) for this reaction is positive ($\Delta G^{\circ\prime} = +6.7$ kJ mol^{-1}), which means that the 1,3-*bis*phosphate concentration should be much lower than that of glyceraldehyde 3-phosphate at the near-equilibium conditions that exist inside the cell. However, glyceraldehyde 3-phosphate dehydrogenase associates with the next enzyme in the pathway (phosphoglycerate kinase), to form a complex. The product of the first reaction, 1,3-*bis*phosphoglycerate, appears to be channeled directly into the active site of phosphoglycerate kinase. In this way the two reactions are effectively linked to form a single reaction and the effective concentration of 1,3-*bis*phosphoglycerate is close to zero.

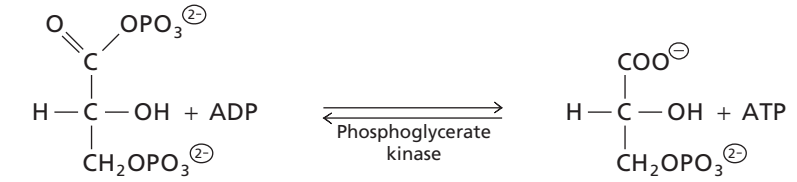
The NADH formed in the glyceraldehyde 3-phosphate dehydrogenase reaction is reoxidized, either by the membrane-associated electron transport chain (Chapter 14) or in other reactions where NADH serves as a reducing agent, such as the reduction of acetaldehyde to ethanol or of pyruvate to lactate (Section 11.3B). The concentration of $NAD^{\oplus}$ in most cells is low. Thus, it is essential to replenish it by reoxidizing NADH or glycolysis will stop at this step. We will see in Section 11.3 that there are several different ways of accomplishing this goal.

▼ Figure 11.7
Fate of carbon atoms from the hexose stage to the triose stage of glycolysis. All numbers refer to the carbon atoms in the original glucose molecule.



7. Phosphoglycerate Kinase

Phosphoglycerate kinase catalyzes phosphoryl group transfer from the "high-energy" mixed anhydride 1,3-*bis*phosphoglycerate to ADP, generating ATP and 3-phosphoglycerate. The enzyme is called a kinase because of the reverse reaction in which 3-phosphoglycerate is phosphorylated.



1,3-*Bis*phosphoglycerate — 3-Phosphoglycerate

(11.7)

Steps 6 and 7 together couple the oxidation of an aldehyde to a carboxylic acid with the phosphorylation of ADP to ATP and the formation of a reducing equivalent.

$$\text{Glyceraldehyde 3-phosphate} + NAD^{\oplus} \longrightarrow \text{1,3-}\textit{Bis}\text{phosphoglycerate} + NADH + H^{\oplus}$$

$$\text{1,3-}\textit{Bis}\text{phosphoglycerate} + ADP \longrightarrow \text{3-Phosphoglycerate} + ATP$$

$$\text{Glyceraldehyde 3-phosphate} + NAD^{\oplus} + P_i + ADP \longrightarrow \text{3-Phosphoglycerate} + NADH + H^{\oplus} + ATP$$

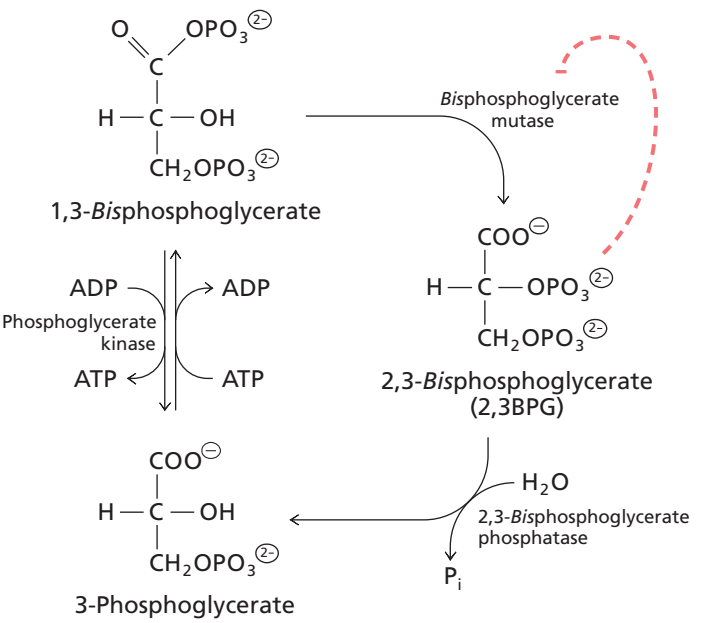
(11.8)

BOX 11.2 FORMATION OF 2,3-*BIS*PHOSPHOGLYCERATE IN RED BLOOD CELLS

An important function of glycolysis in red blood cells is the production of 2,3-*bis*phosphoglycerate, an allosteric inhibitor of the oxygenation of hemoglobin (Section 4.13C). This metabolite is a reaction intermediate and cofactor in Step 8 of glycolysis.

Erythrocytes contain *bis*phosphoglycerate mutase. This enzyme catalyzes the transfer of a phosphoryl group from C-1 to C-2 of 1,3-*bis*phosphoglycerate, to form 2,3-*bis*phosphoglycerate. As shown in the reaction scheme, 2,3-*bis*phosphoglycerate phosphatase catalyzes the hydrolysis of excess 2,3BPG to 3-phosphoglycerate, which can reenter glycolysis and be converted to pyruvate.

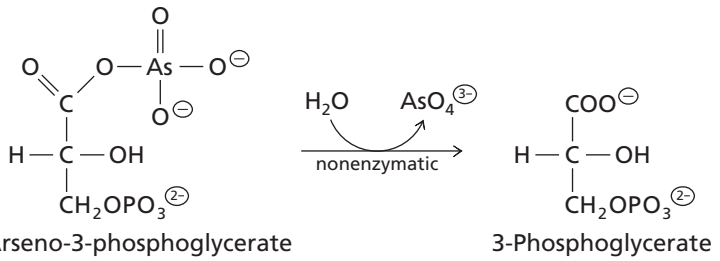
The shunting of 1,3-*bis*phosphoglycerate through these two enzymes bypasses phosphoglycerate kinase, which catalyzes Step 7 of glycolysis, one of the two ATP-generating steps. However, only a portion of the glycolytic flux in red blood cells—about 20%—is diverted through the mutase and phosphatase. Accumulation of free 2,3BPG (i.e., 2,3BPG not bound to hemoglobin) inhibits *bis*phosphoglycerate mutase. In exchange for diminished ATP generation, this bypass provides a regulated supply of 2,3BPG, which is necessary for the efficient release of O_2 from oxyhemoglobin.



▲ **Formation of 2,3-*bis*phosphoglycerate (2,3BPG) in red blood cells.**

BOX 11.3 ARSENATE POISONING

Arsenic, like phosphorus, is in Group V of the periodic table. Arsenate (AsO_4^{3-}) therefore, is an analog of inorganic phosphate. Arsenate competes with phosphate for its binding site in glyceraldehyde 3-phosphate dehydrogenase. Like phosphate, arsenate cleaves the energy-rich thioacyl–enzyme intermediate. However, arsenate produces an unstable analog of 1,3-*bis*phosphoglycerate, called 1-arseno-3-phosphoglycerate, which is rapidly hydrolyzed on contact with water. This nonenzymatic hydrolysis produces 3-phosphoglycerate and regenerates inorganic arsenate, which can again react with a thioacyl–enzyme intermediate. In the presence of arsenate, glycolysis can proceed from 3-phosphoglycerate, but the ATP-producing reaction involving 1,3-*bis*phosphoglycerate is bypassed. As a result, there is no net formation of ATP from glycolysis. Arsenate is a poison because it can replace phosphate in many phosphoryl transfer reactions.



1-Arseno-3-phosphoglycerate — 3-Phosphoglycerate

▲ Spontaneous hydrolysis of 1-arseno-3-phosphoglycerate. Inorganic arsenate can replace inorganic phosphate as a substrate for glyceraldehyde 3-phosphate dehydrogenase, forming the unstable 1-arseno analog of 1,3-*bis*phosphoglycerate.

Arsenite, (AsO_2^-) is much more toxic than arsenate. Arsenite poisons by an entirely different mechanism than arsenate. The arsenic atom of arsenite binds tightly to the two sulfur atoms of lipoamide (Section 7.12), thereby inhibiting the enzymes that require this coenzyme.

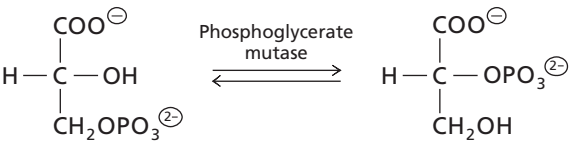


▲ **Cary Grant learned about the effects of arsenic in a popular 1944 movie.**

The formation of ATP by the transfer of a phosphoryl group from a "high energy" compound (such as 1,3-*bis*phosphoglycerate) to ADP is termed **substrate level phosphorylation**. This reaction is the first ATP-generating step of glycolysis. It operates at substrate and product concentrations that are close to the equilibrium concentrations. This is not surprising since the reverse reaction is important in gluconeogenesis, where ATP is utilized. Flux can proceed easily in either direction.

8. Phosphoglycerate Mutase

Phosphoglycerate mutase catalyzes the near-equilibrium interconversion of 3-phosphoglycerate and 2-phosphoglycerate.



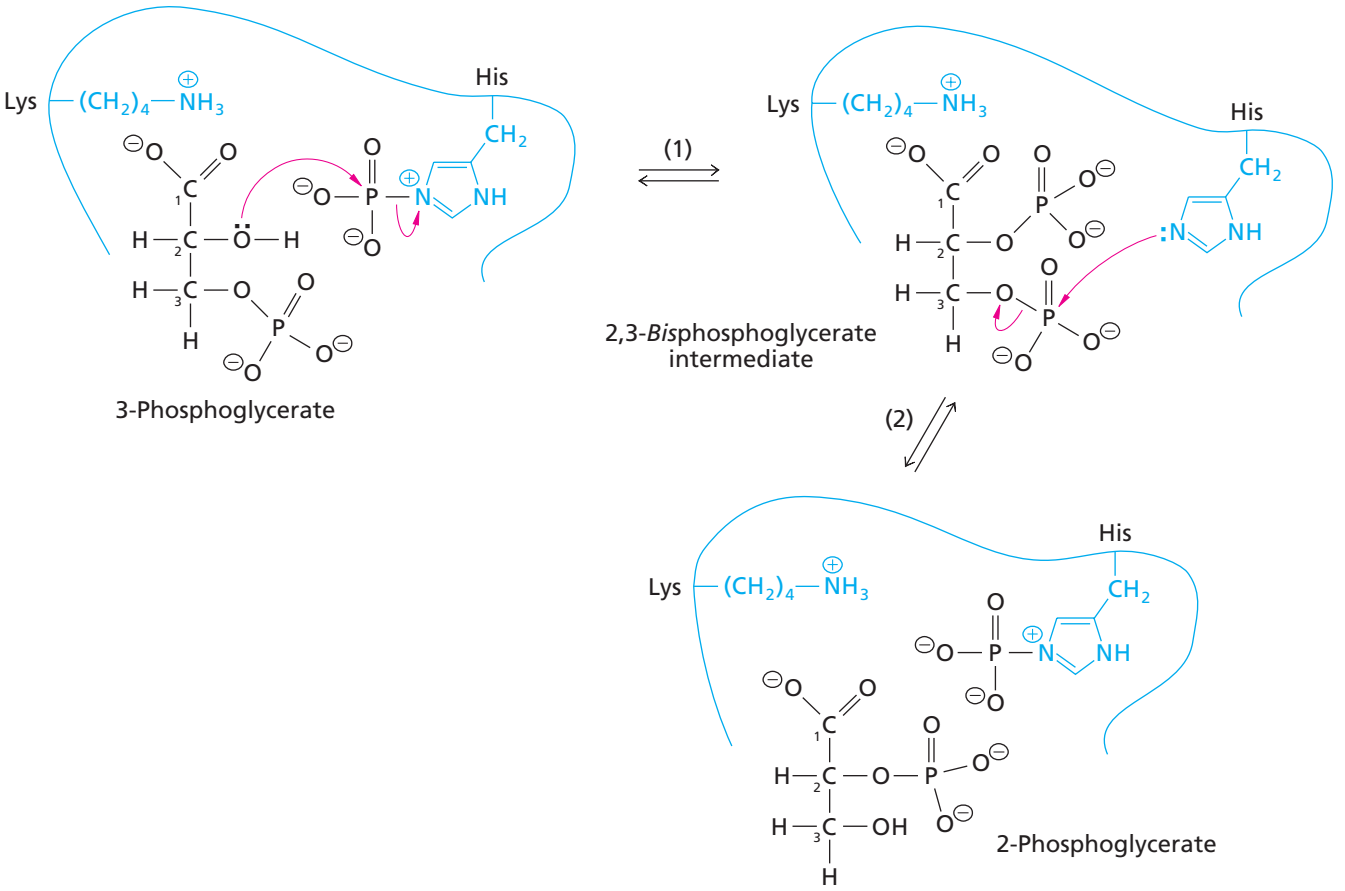
3-Phosphoglycerate — 2-Phosphoglycerate (11.9)

Mutases are isomerases that catalyze the transfer of a phosphoryl group from one part of a substrate molecule to another. There are two different types of phosphoglycerate mutase enzymes. In one type, the phosphoryl group is first transferred to an amino

acid side chain of the enxyme. The enzyme phosphoryl group is then transferred to the second site of the substrate molecule. The dephosphorylated intermediate remains bound in the active site during this process.

Another type of phosphoglycerate mutase makes use of a 2,3-*bis*phosphoglycerate (2,3BPG) intermediate as shown in Figure 11.8. This mechanism also involves a phosphorylated enzyme intemediate but it differs from the other type of enzyme because at no time is there a dephosphorylated metabolite during the reaction. Small amounts of 2,3-*bis*phosphoglycerate are required for full activity of this second type of enzyme. This is because 2,3BPG is required to phosphorylate the enzyme if it becomes dephosphorylated. The enzyme will lose its phosphate group whenever 2,3BPG is released from the active site before it can be converted to 2-phosphoglycerate or 3-phosphoglycerate. The second type of phosphoglycerate mutase is called cofactor-dependent PGM, or dPGM. The first type of enzyme is called cofactor-independent PGM, or iPGM.

dPGM and iPGM are not evolutionarily related. The cofactor-dependent enzyme (dPGM) belongs to a family of enzymes that include acid phosphatases and fructose 2,6-*bis*phosphatase. It is the major form of phosphoglycerate mutase in fungi, some bacteria, and most animals. The cofactor-independent enzyme (iPGM) belongs to the alkaline phosphatase family of enzymes. This version of phosphoglycerate mutase is found in plants and some bacteria. Some species of bacteria have both types of enzyme.

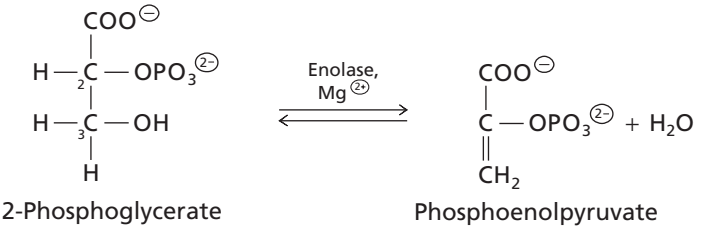


▲ **Figure 11.8**

Mechanism of the conversion of 3-phosphoglycerate to 2-phosphoglycerate in animals and fungi. (1) A lysine residue at the active site of phosphoglycerate mutase binds the carboxylate anion of 3-phosphoglycerate. A histidine residue, which is phosphorylated before the substrate binds, donates its phosphoryl group to form the 2,3-*bis*phosphoglycerate intermediate. (2) Rephosphorylation of the enzyme with a phosphoryl group from the C-3 position of the intermediate yields 2-phosphoglycerate.

9. Enolase

2-Phosphoglycerate is dehydrated to phosphoenolpyruvate in a near-equilibrium reaction catalyzed by enolase. The systematic name of enolase is 2-phosphoglycerate dehydratase.



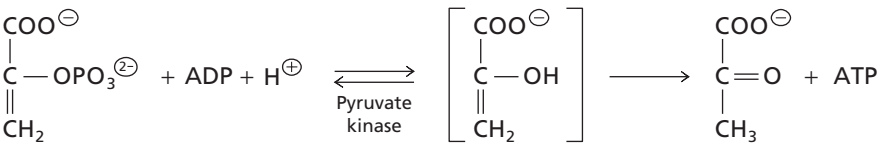
2-Phosphoglycerate Phosphoenolpyruvate (11.10)

In this reaction, the phosphomonoester 2-phosphoglycerate is converted to an enol–phosphate ester, phosphoenolpyruvate, by the reversible elimination of water from C-2 and C-3. Phosphoenolpyruvate has an extremely high phosphoryl group transfer potential because the phosphoryl group holds pyruvate in its unstable enol form (Section 10.7B).

Enolase requires Mg^{2+} for activity. Two magnesium ions participate in this reaction: a "conformational" ion binds to the carboxylate group of the substrate, and a "catalytic" ion participates in the dehydration reaction.

10. Pyruvate Kinase

The second substrate level phosphorylation of glycolysis is catalyzed by pyruvate kinase. Phosphoryl group transfer to ADP generates ATP in this metabolically irreversible reaction. The unstable enol tautomer of pyruvate is an enzyme-bound intermediate.



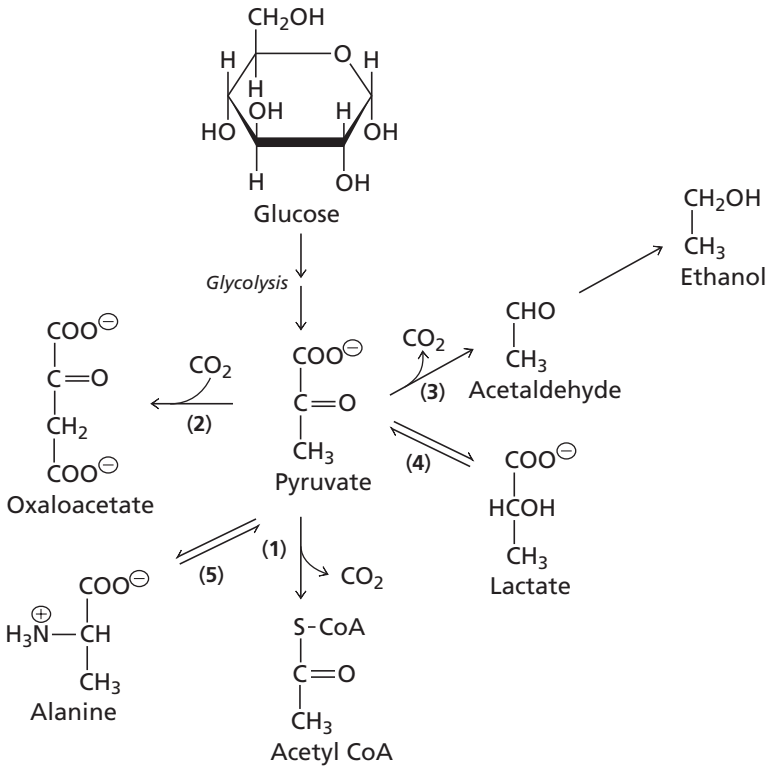
Phosphoenolpyruvate Enolpyruvate Pyruvate (11.11)

Transfer of the phosphoryl group from phosphoenolpyruvate to ADP is the third regulated reaction of glycolysis. Pyruvate kinase is regulated both by allosteric modulators and by covalent modification. In addition, expression of the pyruvate kinase gene in mammals is regulated by various hormones and nutrients. Recall from Chapter 10 that phosphoenolpyruvate hydrolysis has a higher standard Gibbs free energy change than ATP hydrolysis (Table 10.3). Because pyruvate kinase is regulated, the concentration of phosphoenolpyruvate is maintained at a high enough level to drive ATP formation during glycolysis.

11.3 The Fate of Pyruvate

The formation of pyruvate from phosphoenolpyruvate is the last step of glycolysis. Further metabolism of pyruvate typically takes one of five routes (Figure 11.9).

1. Pyruvate can be converted to acetyl CoA and acetyl CoA can be used in a number of metabolic pathways. In one important pathway it is completely oxidized to CO_2 in the citric acid cycle. This fate of pyruvate is described in Chapter 13. This is a route that operates efficiently in the presence of oxygen.



◀ Figure 11.9
Five major fates of pyruvate: (1) Under aerobic conditions, pyruvate is oxidized to the acetyl group of acetyl CoA, which can enter the citric acid cycle for further oxidation. **(2)** Pyruvate can be converted to oxaloacetate, which can be a precursor in gluconeogenesis. **(3)** Under anaerobic conditions, certain microorganisms ferment glucose to ethanol via pyruvate. **(4)** Glucose undergoes anaerobic glycolysis to lactate in vigorously exercising muscles, red blood cells, and certain other cells. **(5)** Pyruvate is converted to alanine.

2. Pyruvate can be carboxylated to produce oxaloacetate. Oxaloacetate is one of the citric acid cycle intermediates but it is also an intermediate in the synthesis of glucose. The fate of pyruvate as a precursor in gluconeogenesis is covered in Chapter 12.
3. In some species, pyruvate can be reduced to ethanol, which is then excreted from cells. This reaction normally takes place under anaerobic conditions where entry of acetyl CoA into the citric acid cycle is unfavorable.
4. In some species, pyruvate can be reduced to lactate. Lactate can be transported to cells that convert it back to pyruvate for entry into one of the other pathways. This is also an anaerobic pathway.
5. In all species, pyruvate can be converted to alanine.

The fate of pyruvate as a precursor in amino acid biosynthesis is discussed in Chapter 17.

In some species, pyruvate can be converted to phosphoenolpyruvate (Section 12.1B).

During glycolysis, $NAD^{\oplus}$ is reduced to NADH at the glyceraldehyde 3-phosphate dehydrogenase reaction (Step 6). In order for glycolysis to operate continuously, the cell must be able to regenerate $NAD^{\oplus}$. Otherwise, all the coenzyme would rapidly accumulate in the reduced form, and glycolysis would stop. Under aerobic conditions, NADH can be oxidized by the membrane-associated electron transport system (Chapter 14), which requires molecular oxygen. Under anaerobic conditions, the synthesis of ethanol or lactate consumes NADH and regenerates the $NAD^{\oplus}$ essential for continued glycolysis.

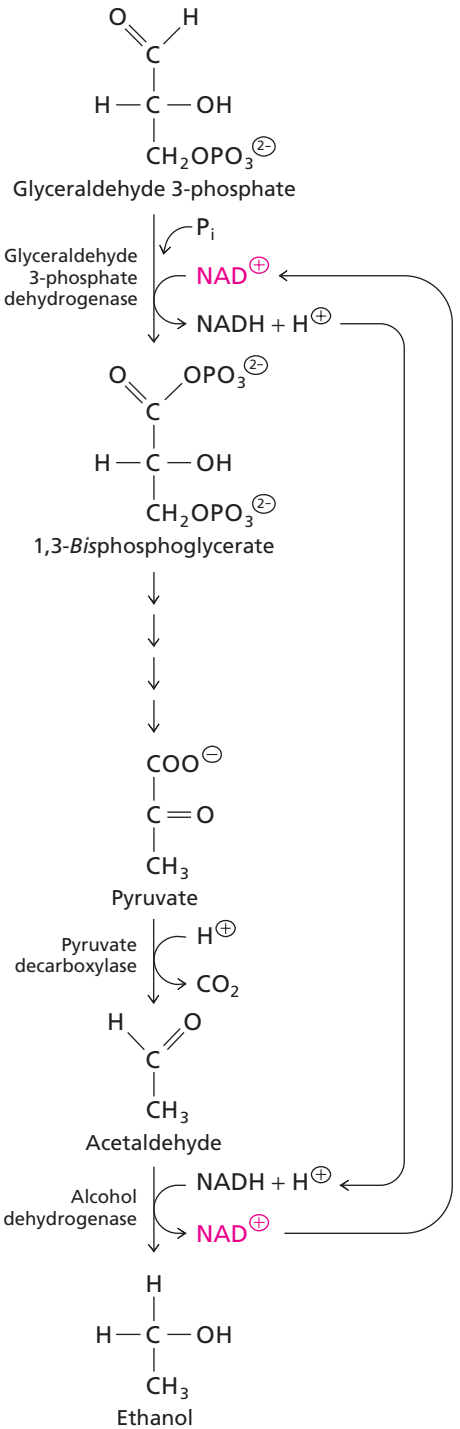
A. Metabolism of Pyruvate to Ethanol

Many bacteria, and some eukaryotes, are capable of surviving in the absence of oxygen. They convert pyruvate to a variety of compounds that are secreted. Ethanol is one of these compounds. It assumes significance in biochemistry because the synthesis of ethanol by highly selected strains of yeast is important in the production of beer and wine. Yeast cells convert pyruvate to ethanol and CO_2 and oxidize NADH to $NAD^{\oplus}$. Two reactions are required. First, pyruvate is decarboxylated to acetaldehyde in a reaction catalyzed by pyruvate decarboxylase. This enzyme requires the coenzyme thiamine diphosphate (TDP); its mechanism was described in the coenzymes chapter (Section 7.7).

Alcohol dehydrogenase catalyzes the reduction of acetaldehyde to ethanol. This oxidation–reduction reaction is coupled to the oxidation of NADH. These reactions and

KEY CONCEPT

In the absence of oxygen, eukaryotes have to give up the net gain of 2 NADH molecules in order to make lactate or ethanol.



▲ **Figure 11.10**
Anaerobic conversion of pyruvate to ethanol in yeast.

the cycle of $NAD^{\oplus}$/NADH reduction and oxidation in alcoholic fermentation are shown in Figure 11.10. *Fermentation* refers to a process where electrons from glycolysis—in the form of NADH—are passed to an organic molecule such as ethanol instead of being passed on to the membrane-associated electron transport chain and ultimately oxygen (*respiration*).

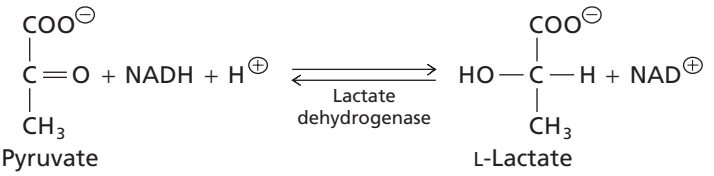
The sum of the glycolytic reactions and the conversion of pyruvate to ethanol is

$$\text{Glucose} + 2\ P_i^{2\ominus} + 2\ ADP^{3\ominus} + 2\ H^{\oplus} \longrightarrow 2\ \text{Ethanol} + 2\ CO_2 + 2\ ATP^{4\ominus} + 2\ H_2O \qquad (11.12)$$

These reactions have familiar commercial roles in the manufacture of beer and bread. In the brewery, the carbon dioxide produced during the conversion of pyruvate to ethanol can be captured and used to carbonate the final alcoholic brew; this gas produces the foamy head. In the bakery, carbon dioxide is the agent that causes bread dough to rise.

B. Reduction of Pyruvate to Lactate

Pyruvate is reduced to lactate in a reversible reaction catalyzed by lactate dehydrogenase. This reaction is common in anaerobic bacteria and also in mammals.



(11.13)

Lactate dehydrogenase is a classic dehydrogenase using $NAD^{\oplus}$ as a coenzyme; the mechanism was presented in Section 7.4. This is an oxidation–reduction reaction in which pyruvate is reduced to lactate by transfer of a hydride ion from NADH.

The lactate dehydrogenase reaction oxidizes the reducing equivalents generated in the glyceraldehyde 3-phosphate reaction and lowers the potential energy gain of glycolysis. It plays the same role that ethanol production accomplishes in other species (Figure 11.10). The net effect is to maintain flux in the glycolytic pathway and the production of ATP. In bacteria, lactate is secreted or converted to other end products, such as propionate. In mammals, lactate can only be reconverted to pyruvate.

The production of lactate in mammalian cells is essential in tissues where glucose is the main carbon source and reducing equivalents (NADH) are not needed in biosynthesis reactions or cannot be used to generate ATP by oxidative phosphorylation. A good example is the formation of lactate in skeletal muscle cells during vigorous exercise. Lactate formed in muscle cells is transported out of cells and carried via the bloodstream to the liver, where it is converted to pyruvate by the action of hepatic lactate dehydrogenase (see Cori cycle, Section 12.2A). Further metabolism of pyruvate requires oxygen. When the supply of oxygen to tissues is inadequate, all tissues produce lactate by anaerobic glycolysis.

The overall reaction for glucose degradation to lactate is

$$\text{Glucose} + 2\ P_i^{2\ominus} + 2\ ADP^{3\ominus} \longrightarrow 2\ \text{Lactate}^{\ominus} + 2\ ATP^{4\ominus} + 2\ H_2O \qquad (11.14)$$

Lactic acid is also produced by *Lactobacillus* and certain other bacteria when they ferment the sugars in milk. The acid denatures the proteins in milk, causing the curdling necessary for cheese and yogurt production.

Regardless of the final product—ethanol or lactate—glycolysis generates two molecules of ATP per molecule of glucose consumed. Oxygen is not required in either case. This feature is essential not only for anaerobic organisms but also for some specialized cells in multicellular organisms. Some tissues (such as kidney medulla and parts of the brain), termed obligatory glycolytic tissues, rely on glycolysis for all their energy. In the

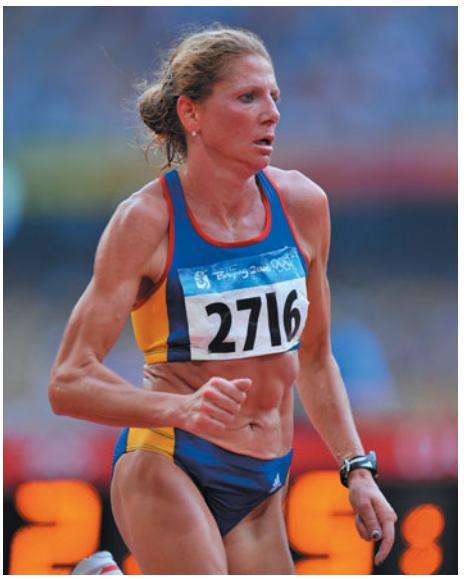
BOX 11.4 THE LACTATE OF THE LONG-DISTANCE RUNNER

Most of you have heard stories about lactate buildup during strenuous exercise. It all sounds so plausible. When muscle cells are working hard they use up glucose to generate ATP, which is required for muscle contraction. During very strenuous activity, the production of pyruvate may outstrip its ability to be oxidized by the citric acid cycle. If muscle cells aren't getting enough oxygen, then pyruvate is converted to lactic acid and the accumulation of lactic acid causes acidosis leading to muscle pain and reduced efficiency.

It's a nice story, but it's wrong.

Lactate concentration in muscle cells and in the bloodstream does increase but lactate is not an acid. It cannot donate a proton, so the increase in protons (acidosis) must come from another source. Lactate really is the product of the lactate dehydrogenase reaction, not lactic acid (which can donate a proton).

There is no net production of protons in the pathway leading from glucose to lactate. The acidosis seen after strenuous exercise is mostly due to the release of protons during ATP hydrolysis associated with muscle contraction. This is a temporary imbalance since ATP is soon regenerated in order to maintain a high steady state concentration. Lactate may indirectly contribute to some acidosis because, as a potent anion, it may affect buffering capacity but the effect is not large. Lactate has been getting a bum rap for decades, including previous editions of this textbook.



cornea of the eye, for example, oxygen availability is limited by poor blood circulation. Anaerobic glycolysis provides the necessary ATP for such tissues.

11.4 Free Energy Changes in Glycolysis

When the glycolytic pathway is operating, the flow of metabolites is from glucose to pyruvate. Under these conditions, the Gibbs free energy change for every single reaction must be either negative or zero. It is interesting to compare the *standard* Gibbs free energy changes ($\Delta G°'$) and the *actual* Gibbs free energy changes (ΔG) under conditions where flux through the glycolytic pathway is high. Such conditions occur in erythrocytes where blood glucose is the main source of energy and there is very little synthesis of carbohydrates (or any other molecules). The actual concentrations of the intermediates in glycolysis have been measured and the Gibbs free energy changes have been calculated. The standard Gibbs free energy changes for each of the ten reactions of glycolysis are shown in Table 11.2, The first column lists $\Delta G°'$ values under typical standard conditions (25°C and zero ionic strength) and the second column corrects those standard Gibbs free energy changes to mammalian physiological conditions (37°C in the presence of Mg^{2+}, Ca^{2+}, Na^{+} and K^{+}).

Figure 11.11 shows the cumulative standard Gibbs free energy changes and actual free energy changes for the glycolytic reactions in erythrocytes. The vertical axis indicates cumulative Gibbs free energy changes for each of the steps of glycolysis. The figure illustrates the difference between the Gibbs free energy changes under standard physiological conditions ($\Delta G°'$) and actual free energy changes under cellular conditions (ΔG).

The blue plot tracks the actual cumulative free energy changes. It shows that each reaction has a Gibbs free energy change that is either negative or zero. This is an essential requirement for conversion of glucose to pyruvate. It follows that the overall pathway, which is the sum of the individual reactions, must also have a negative free energy change. The overall Gibbs free energy change for glycolysis is about -72 kJ mol^{-1} under the conditions found in erythrocytes.

Table 11.2 Standard Gibbs free energies for reactions of glycolysis

Glycolysis reaction	$\Delta G^{\circ\prime}$ (kJ mol^{-1}) (standard conditions)	$\Delta G^{\circ\prime}$ (kJ mol^{-1}) (physiological conditions)
1	−17.2	−19.4
2	+2.0	+2.8
3	−18.0	−15.6
4	+28.0	+24.6
5	+7.9	+7.6
6	+6.7	+2.6
7	−18.8	−16.4
8	+4.4	+6.4
9	−2.7	−4.5
10	−25.5	−27.2

Data from Minakami and de Verdier (1976) and Li et al. (2010).

The actual Gibbs free energy changes are large only for Steps 1, 3, and 10, which are catalyzed by hexokinase, phosphofructokinase-1, and pyruvate kinase, respectively—the steps that are both metabolically irreversible and regulated. The ΔG values for the other steps are very close to zero. In other words, these other steps are near-equilibrium reactions in cells.

In contrast, the standard Gibbs free energy changes for the same ten reactions exhibit no consistent pattern. Although the three reactions with large negative Gibbs free energy changes in cells also have large standard Gibbs free energy changes, this may be coincidental since some of the near-equilibrium reactions in cells also have large values for $\Delta G^{\circ\prime}$. Furthermore, some of the $\Delta G^{\circ\prime}$ values for the reactions of glycolysis are positive, indicating that under standard conditions, flux through these reactions occurs toward substrate rather than product. This is especially obvious in Step 4 (aldolase) and Step 6 (glyceraldehyde 3-phosphate dehydrogenase). In other types of cells these near-equilibrium reactions might operate in the opposite direction during glucose synthesis.

KEY CONCEPT

The net production of product in a metabolic pathway (flux) will only occur if: (a) the overall Gibbs free energy change is negative, and (b) the Gibbs free energy change of each step in the pathway is either negative or zero.

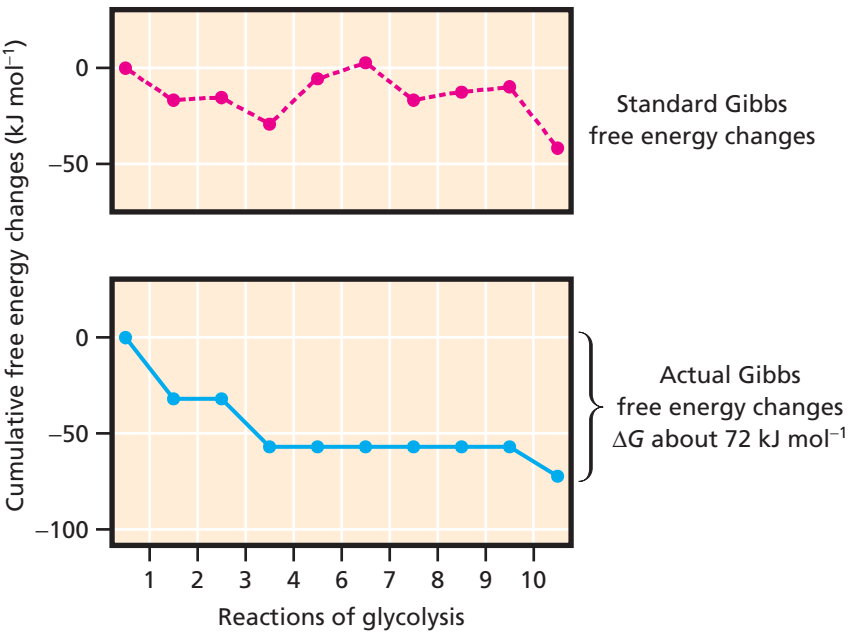


Figure 11.11 ▶
Cumulative standard and actual Gibbs free energy changes for the reactions of glycolysis. The vertical axis indicates free energy changes in kJ mol^{-1}. The reactions of glycolysis are plotted in sequence horizontally. The upper plot (red) tracks the standard free energy changes, and the bottom plot (blue) shows actual free energy changes in erythrocytes. The interconversion reaction catalyzed by triose phosphate isomerase (Reaction 5) is not shown. [Adapted from Hamori, E. (1975). Illustration of free energy changes in chemical reactions. *J. Chem. Ed.* 52:370–373.]

11.5 Regulation of Glycolysis

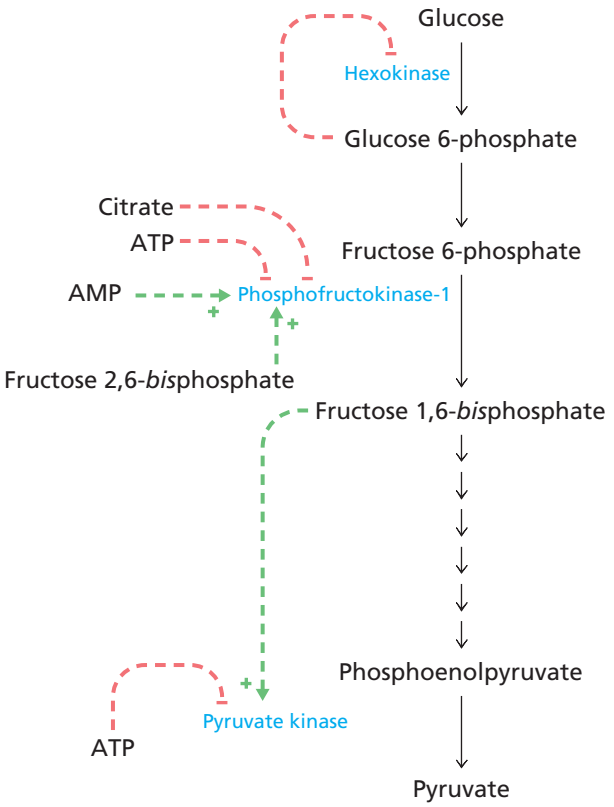
The regulation of glycolysis has been examined more thoroughly than that of any other pathway. Data on regulation come primarily from two types of biochemical research: enzymology and metabolic biochemistry. In enzymological approaches, metabolites are tested for their effects on isolated enzymes and the structure and regulatory mechanisms of individual enzymes are studied. Metabolic biochemistry analyzes the concentrations of pathway intermediates *in vivo* and stresses pathway dynamics under cellular conditions. We sometimes find that *in vitro* studies are deceptive as indicators of pathway dynamics *in vivo*. For instance, a compound may modulate enzyme activity *in vitro*, but only at concentrations not found in the cell. Accurate interpretation of biochemical data greatly benefits from a combination of enzymological and metabolic expertise.

In this section, we examine each regulatory site of glycolysis. Our primary focus is on the regulation of glycolysis in mammalian cells—in particular, those cells where glycolysis is an important pathway. Variations on the regulatory themes presented here can be found in other species.

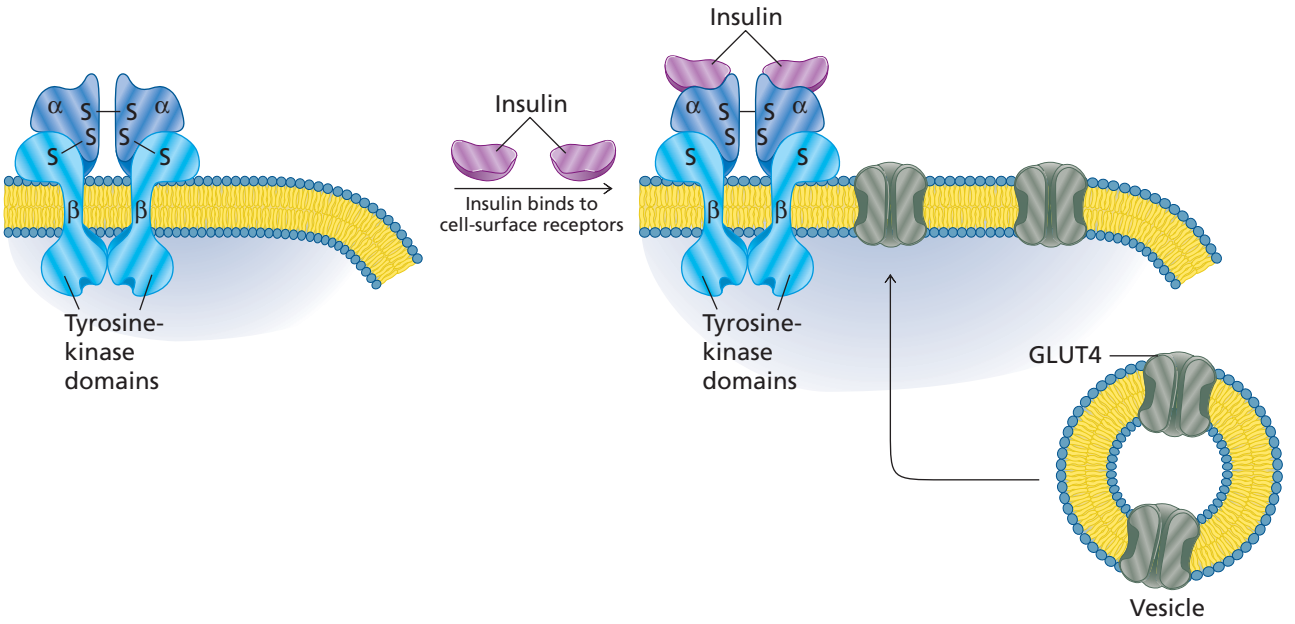
The regulatory effects of metabolites on glycolysis are summarized in Figure 11.12. The activation of glycolysis is desirable when ATP is required by processes such as muscle contraction. Hexokinase is inhibited by excess glucose 6-phosphate, and PFK-1 is inhibited by the accumulation of ATP and citrate (an intermediate in the energy-producing citric acid cycle). ATP and citrate both signal an adequate energy supply. Consumption of ATP leads to the accumulation of AMP, which relieves the inhibition of PFK-1 by ATP. Fructose 2,6-*bis*phosphate also relieves this inhibition. The rate of formation of fructose 1,6-*bis*phosphate then increases, which in certain tissues activates pyruvate kinase. Glycolytic activity decreases when its products are no longer required.

A. Regulation of Hexose Transporters

The first potential step for regulating glycolysis is the transport of glucose into the cell. In most mammalian cells, the intracellular glucose concentration is far lower than the blood glucose concentration, and glucose moves into the cells, down its concentration gradient, by passive transport. All mammalian cells possess membrane-spanning



◀ Figure 11.12
Summary of the metabolic regulation of the glycolytic pathway in mammals. Not shown are the effects of ADP on PFK-1, which vary among species.



▲ **Figure 11.13**
Regulation of glucose transport by insulin. The binding of insulin to cell-surface receptors stimulates intracellular vesicles containing membrane-embedded GLUT4 transporters to fuse with the plasma membrane. This delivers GLUT4 transporters to the cell surface and thereby increases the capacity of the cell to transport glucose.

Membrane transport systems are described in Section 9.11.

glucose transporters. Intestinal and kidney cells have a $Na^{\oplus}$-dependent cotransport system called SGLT1 for absorbing dietary glucose and urinary glucose, respectively. Other mammalian cells contain transporters from the GLUT family of passive hexose transporters. Each of the six members of the GLUT family has unique properties suitable for the metabolic activities of the tissues in which it is found.

The hormone insulin stimulates high rates of glucose uptake into skeletal and heart muscle cells and adipocytes via the transporter GLUT4. When insulin binds to receptors on the cell surface, intracellular vesicles that have GLUT4 embedded in their membranes fuse with the cell surface by exocytosis (Section 9.11D), thereby increasing the capacity of the cells to transport glucose (Figure 11.13). Because GLUT4 is found at high levels only in striated muscle and adipose tissue, insulin-regulated uptake of glucose occurs only in these tissues.

In most tissues, a basal level of glucose transport in the absence of insulin is maintained by GLUT1 and GLUT3. GLUT2 transports glucose into and out of the liver, and GLUT5 transports fructose in the small intestine. GLUT7 transports glucose 6-phosphate from the cytoplasm into the endoplasmic reticulum.

Once inside a cell, glucose is rapidly phosphorylated by the action of hexokinase. This reaction traps the glucose inside the cell since phosphorylated glucose cannot cross the plasma membrane. As we will see, phosphorylated glucose can also be used in glycogen synthesis or in the pentose phosphate pathway (Chapter 12).

B. Regulation of Hexokinase

The reaction catalyzed by mammalian hexokinase is metabolically irreversible (because it is regulated) but in bacteria and many other eukaryotes hexokinase is not regulated. In those species, the concentrations of reactants and products reach equilibrium. In mammals, the various forms of hexokinase are subject to complex regulation.

At physiological concentrations, the enzyme product, glucose 6-phosphate, allosterically inhibits hexokinase isozymes I, II, and III, but not glucokinase (isozyme IV). Glucokinase is more abundant than the other hexokinases in the liver and the insulin-secreting cells of the pancreas. The concentration of glucose 6-phosphate increases when glycolysis is inhibited at sites further along the pathway. The inhibition of hexokinases I, II, and III by glucose 6-phosphate therefore coordinates the activity of hexokinase with the activity of subsequent enzymes of glycolysis.

Glucokinase is suited to the physiological role of the liver in managing the supply of glucose for the entire body. In most cells, glucose concentrations are maintained far

BOX 11.5 GLUCOSE 6-PHOSPHATE HAS A PIVOTAL METABOLIC ROLE IN THE LIVER

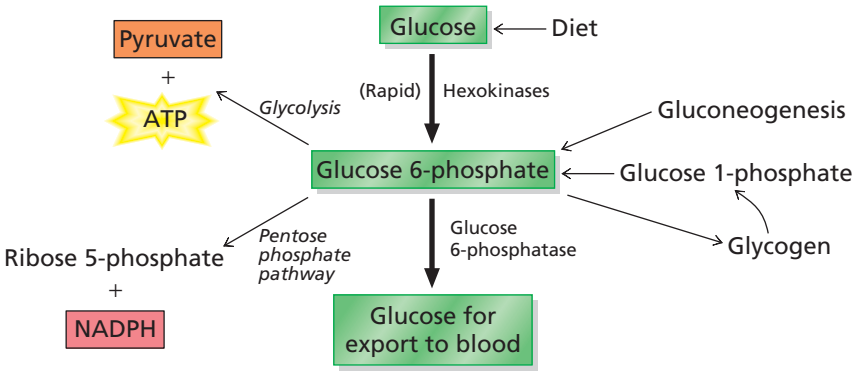
Glucose 6-phosphate is an initial substrate for several metabolic pathways (figure below). We have already seen that it is the initial intermediate in glycolysis. Glucose 6-phosphate is formed rapidly in liver cells from dietary glucose or newly synthesized glucose (from gluconeogenesis in liver cells; Section 12.1).

The principal use of liver glucose 6-phosphate is to maintain a constant concentration of blood glucose. Glucose 6-phosphatase is the enzyme responsible for catalyzing hydrolysis of glucose 6-phosphate to glucose. (This reaction is also the last step in gluconeogenesis.)

Glucose 6-phosphate that is not required for blood glucose is stored as liver glycogen (Section 12.6). Glycogen is subsequently degraded when a supply of glucose is needed. Hormones regulate both the synthesis and degradation of glycogen.

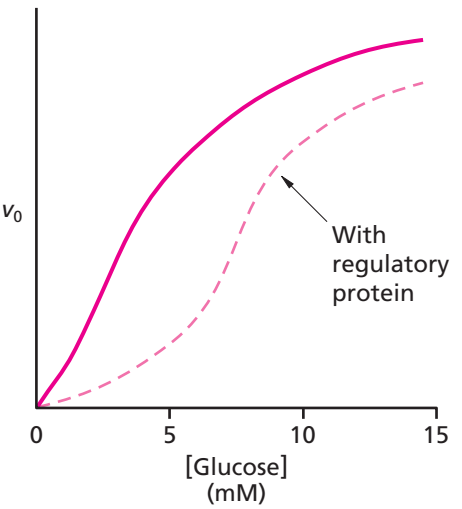
In addition to using it for balancing the blood glucose concentration, the liver metabolizes glucose 6-phosphate by the pentose phosphate pathway (Section 12.5) to produce ribose 5-phosphate (for nucleotides) and NADPH (for synthesis of fatty acids). We have seen in this chapter that glucose 6-phosphate can also enter the glycolytic pathway, where it is converted initially to pyruvate, which leads to another major metabolite—acetyl CoA.

▼ Glucose 6-phosphate is at a pivotal position in carbohydrate metabolism in the liver.



below the concentration in blood. However, glucose freely enters the liver via GLUT2, and the concentration of glucose in liver cells matches the concentration in blood. The blood glucose concentration is typically 5 mM, though after a meal it can rise as high as 10 mM. Most hexokinases have K_m values for glucose of about 0.1 mM or less. In contrast, glucokinase has a K_m of 2 to 5 mM for glucose; in addition, it is not significantly inhibited by glucose 6-phosphate. Therefore, liver cells can form glucose 6-phosphate (for glycogen synthesis) by the action of glucokinase when glucose is abundant and other tissues have sufficient glucose.

The activity of glucokinase is modulated by fructose phosphates. In liver cells, a regulatory protein inhibits glucokinase in the presence of fructose 6-phosphate, lowering its affinity for glucose to about 10 mM (Figure 11.14). Note that the v_0 vs. [S] curves for glucokinase are sigmoidal and not the hyperbolic curves expected for an enzyme obeying Michaelis–Menten kinetics. This is a common feature of allosterically regulated proteins. It means that there is no true K_m value for glucokinase. We can say that the effect of the regulatory protein is to raise the *apparent* K_m of the enzyme. Flux through glucokinase is usually low because liver cells always contain considerable fructose 6-phosphate. The flux can increase after a meal, when fructose 1-phosphate—derived only from dietary fructose—relieves the inhibition of glucokinase by the regulatory protein. Therefore, the liver can respond to increases in blood carbohydrate concentrations with proportionate increases in the rate of phosphorylation of glucose.



▲ Figure 11.14
Plot of initial velocity (v_0) versus glucose concentration for glucokinase. The addition of a regulatory protein lowers the enzyme's affinity for glucose. The blood glucose concentration is 5 to 10 mM.

C. Regulation of Phosphofructokinase-1

The second site of allosteric regulation is the reaction catalyzed by phosphofructokinase-1. PFK-1 is a large, oligomeric enzyme with a molecular weight ranging in different species from about 130,000 to 600,000. The quaternary structure of PFK-1 also varies among species. The bacterial and mammalian enzymes are both tetramers; the yeast

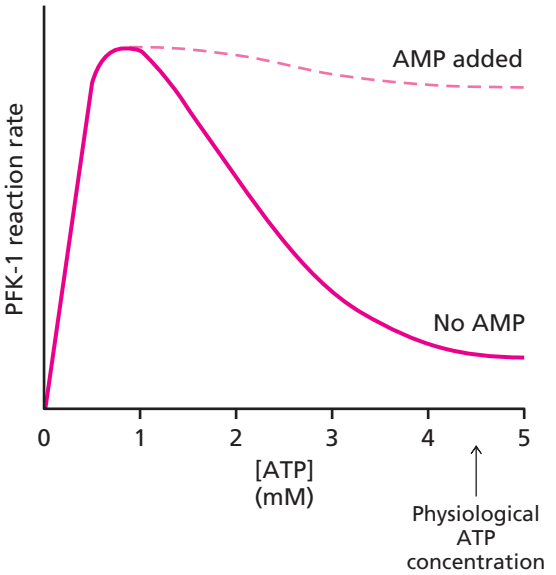


Figure 11.15 ▶
Regulation of PFK-1 by ATP and AMP. In the absence of AMP, PFK-1 is almost completely inhibited by physiological concentrations of ATP. In the range of AMP concentrations found in the cell, the inhibition of PFK-1 by ATP is almost completely relieved. [Adapted from Martin, B. R. (1987). *Metabolic Regulation: A Molecular Approach* (Oxford: Blackwell Scientific Publications), p. 222.]

enzyme is an octamer. This complex enzyme has several regulatory sites. The regulatory properties of the *Escherichia coli* phosphofructokinase-1 are described in Section 5.10A.

ATP is both a substrate and, in most species, an allosteric inhibitor of PFK-1. ATP increases the apparent K_m of PFK-1 for fructose 6-phosphate. The bacterial enzyme is activated by ADP but in mammals AMP is the allosteric activator of PFK-1. AMP acts by relieving the inhibition caused by ATP (Figure 11.15). ADP activates mammalian PFK-1 but inhibits the plant kinase; in bacteria, protists, and fungi, the regulatory effects of purine nucleotides vary among species.

The concentration of ATP does not change very much in most mammalian cells despite large changes in the rate of its formation and utilization. However, as discussed in Section 10.6, significant changes in the concentrations of ADP and AMP do occur because these molecules are present in cells in much lower concentrations than ATP and small changes in the level of ATP cause proportionally larger changes in the levels of ADP and AMP. The steady state concentrations of these compounds are therefore able to control flux through PFK-1.

Recall that activation by ADP (or AMP) makes sense in light of the net production of ATP in glycolysis. Elevated levels of ADP or AMP indicate a deficiency of ATP that can be offset by increasing the rate of degradation of glucose (Section 5.9A).

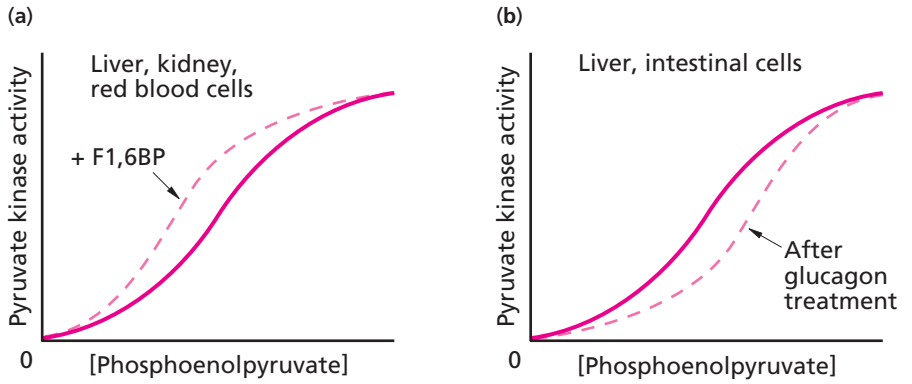
Citrate, an intermediate of the citric acid cycle, is another physiologically important inhibitor of mammalian PFK-1. An elevated concentration of citrate indicates that the citric acid cycle is blocked and further production of pyruvate would be pointless. The regulatory effect of citrate on PFK-1 is an example of feedback inhibition that regulates the supply of pyruvate to the citric acid cycle. (Phosphoenolpyruvate, not citrate, inhibits the bacterial enzyme.)

As shown in Figure 11.12, fructose 2,6-*bis*phosphate is a potent activator of PFK-1, effective in the micromolar range. This compound is present in mammals, fungi, and plants, but not prokaryotes. We will return to the role of fructose 2,6-*bis*phosphate in the next chapter after we have described gluconeogenesis and glycogen metabolism.

D. Regulation of Pyruvate Kinase

The third site of allosteric regulation of glycolysis is the reaction catalyzed by pyruvate kinase. Single-cell species, such as bacteria, and protists, have a single pyruvate kinase gene. The enzyme is allosterically regulated in a simple manner—its activity is affected by pyruvate and/or fructose 1,6-*bis*phosphate. Regulation is much more complex in mammals because different organs have different requirements for glucose and glycolysis.

Four different isozymes of pyruvate kinase are present in mammalian tissues. The isozymes found in liver, kidney, and red blood cells yield a sigmoidal curve when initial velocity is plotted against phosphoenolpyruvate concentration (Figure 11.16a). This indicates that PEP is an allosteric activator. These enzymes are also allosterically

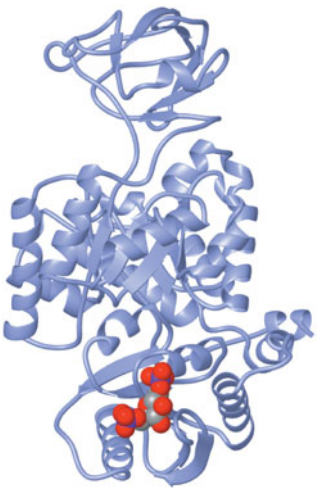


◀ Figure 11.16
Plots of initial velocity (v_0) versus phosphoenolpyruvate concentration for pyruvate kinase. (a) For isozymes in some cells, the presence of fructose 1,6-*bis*phosphate shifts the curve to the left, indicating that fructose 1,6-*bis*phosphate is an activator of the enzymes. **(b)** When liver or intestinal cells are incubated with glucagon, pyruvate kinase is phosphorylated by the action of protein kinase A. The curve shifts to the right, indicating less activity for pyruvate kinase.

activated by fructose 1,6-*bis*phosphate and inhibited by ATP. In the absence of fructose 1,6-*bis*phosphate, physiological concentrations of ATP almost completely inhibit the isolated enzyme. The presence of fructose 1,6-*bis*phosphate—probably the most important modulator *in vivo*—shifts the curve to the left. With sufficient fructose 1,6-*bis*phosphate, the curve becomes hyperbolic. Figure 11.16a shows that for a range of substrate concentrations, enzyme activity is greater in the presence of the allosteric activator. Recall that fructose 1,6-*bis*phosphate is the product of the reaction catalyzed by PFK-1. Its concentration increases when the activity of PFK-1 increases. Since fructose 1,6-*bis*phosphate activates pyruvate kinase, the activation of PFK-1 (which catalyzes Step 3 of the glycolytic pathway) causes subsequent activation of pyruvate kinase (the last enzyme in the pathway). This is an example of feed-forward activation.

The predominant isozyme of pyruvate kinase found in mammalian liver and intestinal cells is subject to an additional type of regulation, covalent modification by phosphorylation. Protein kinase A, which also catalyzes the phosphorylation of PFK-2 (Figure 11.17), catalyzes the phosphorylation of pyruvate kinase. Pyruvate kinase is less active in the phosphorylated state. The change in kinetic behavior is shown in Figure 11.16b, which depicts a plot of pyruvate kinase activity in liver and intestinal cells in the presence and absence of glucagon, a stimulator of protein kinase A. Dephosphorylation of pyruvate kinase is catalyzed by a protein phosphatase.

The pyruvate kinase activity of liver cells decreases on starvation and increases on ingestion of a diet high in carbohydrate. These long term changes are due to changes in the rate of synthesis of pyruvate kinase and not allosteric regulation or covalent modification.



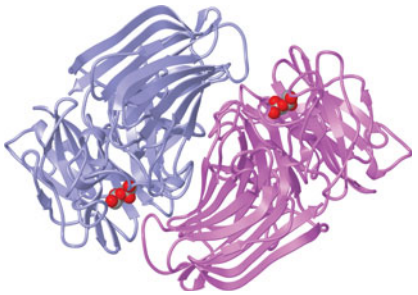
▲ Figure 11.17
Pyruvate kinase from the yeast *Saccharomyces cerevisiae*, with the activator fructose 1,6-*bis*phosphate (red). The active site is in the large central domain. [PDB 1A3W]

E. The Pasteur Effect

Louis Pasteur observed that when yeast cells grow anaerobically, they produce much more ethanol and consume much more glucose than when they grow aerobically. Similarly, skeletal muscle accumulates lactate under anaerobic conditions but not when it metabolizes glucose aerobically. In both yeast and muscle, the rate of conversion of glucose to pyruvate is much higher under anaerobic conditions. The slowing of glycolysis in the presence of oxygen is called the **Pasteur effect**. As we will see in Chapter 13, the complete aerobic metabolism of a glucose molecule produces much more ATP than the two molecules of ATP produced by glycolysis alone. Therefore, for any given ATP requirement, fewer glucose molecules must be consumed under aerobic conditions. Cells sense the state of ATP supply and demand, and they modulate glycolysis by several mechanisms. For example, the availability of oxygen leads to the inhibition of PFK-1 (and thus glycolysis), probably through an increase in the ATP/AMP ratio.

11.6 Other Sugars Can Enter Glycolysis

Glucose and glucose 6-phosphate are the most common substrates for glycolysis, especially in vertebrates where glucose is circulated in the bloodstream. However, a variety of other sugars can be degraded by the glycolytic pathway. In this section, we will see how sucrose, fructose, lactose, galactose, and mannose can be metabolized.



▲ **Invertase from the yeast *Schwanniomyces occidentalis*.** The active form of the enzyme is a dimer of identical subunits. Fructose (space-filling representation) is bound at the active site. [PDB 3KF3]

A. Sucrose Is Cleaved to Monosaccharides

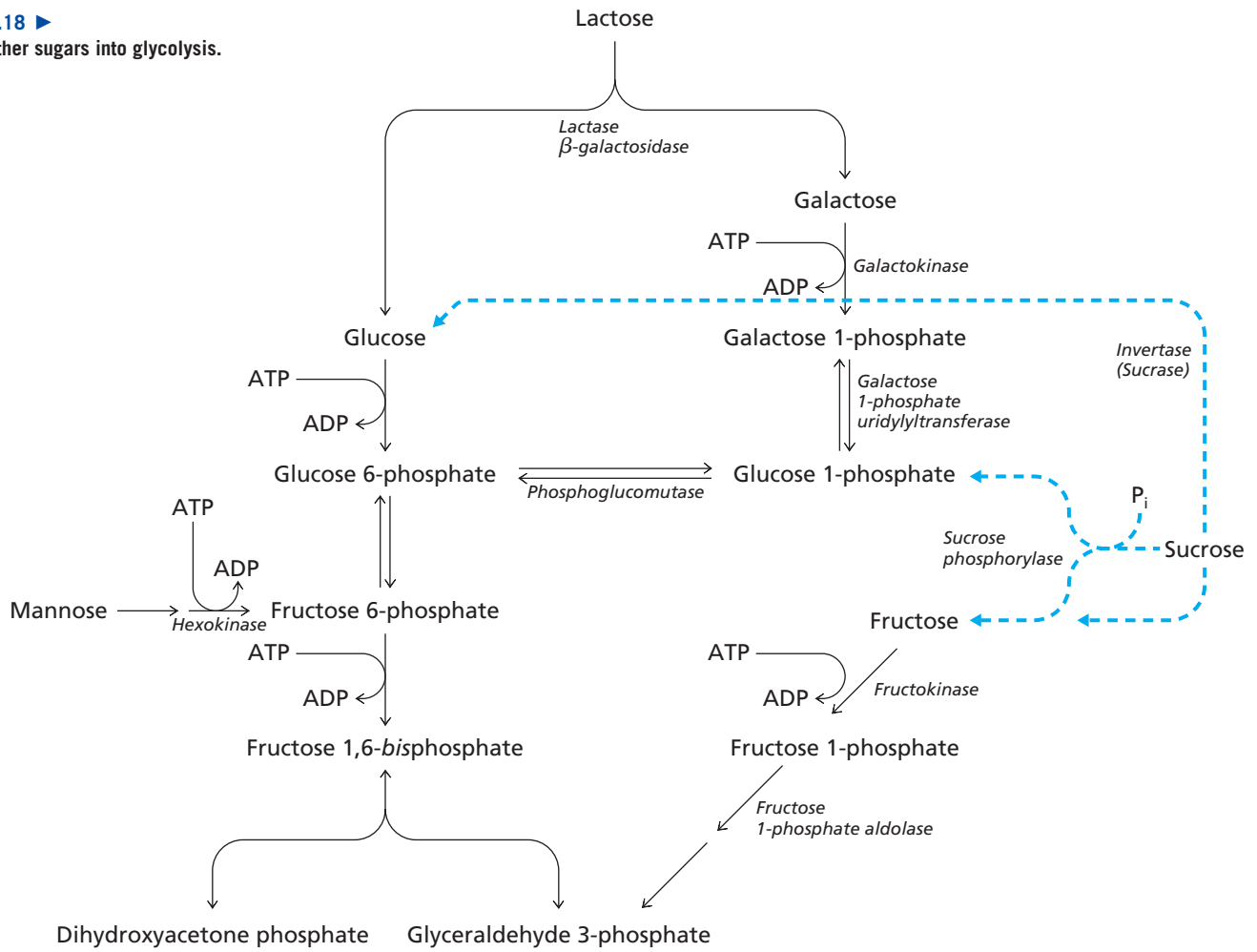
The disacharide sucrose can be degraded to its two component monosaccharides: fructose and glucose. This cleavage is catalyzed by a class of enzymes called sucrases. Invertase (β-fructofuranosidease) is one of the most common sucrases. It catalyzes a hydrolytic cleavage of the glycosidic linkage between the oxygen and the glucose residue, producing fructose and glucose (Figure 11.18). The glucose residues are then phosphorylated by hexokinase and the fructose residues enter the pathway as described below.

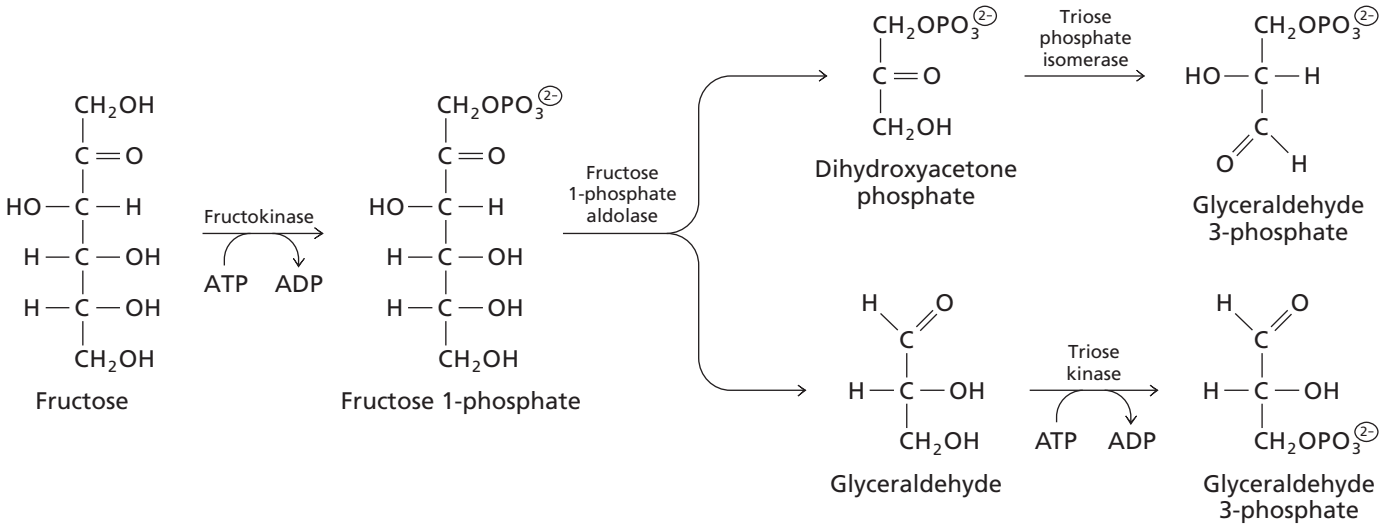
Some bacteria have a very interesting enzyme called sucrose phosphorylase. It cleaves sucrose in the presence of inorganic phosphate converting it to a molecule of fructose and a molecule of glucose 1-phosphate (Figure 11.18). All sugars entering glycolysis need to be phosphorylated at some stage and this step almost always involves the expenditure of one ATP equivalent. Sucrose phosphorylase is an important exception because it produces glucose 1-phosphate without spending any ATP currency.

B. Fructose Is Converted to Glyceraldehyde 3-Phosphate

Fructose is phosphorylated to fructose 1-phosphate by the action of a specific ATP-dependent fructokinase (Figure 11.19). In mammals, this step occurs in the liver after fructose has been absorbed in the intestine and transferred in the bloodstream. Fructose 1-phosphate aldolase catalyzes the cleavage of fructose 1-phosphate to dihydroxyacetone phosphate and glyceraldehyde. The glyceraldehyde is then phosphorylated to glyceraldehyde 3-phosphate in a reaction catalyzed by triose kinase, consuming a second molecule of ATP. Dihydroxyacetone phosphate is converted to a second molecule of glyceraldehyde 3-phosphate by the action of triose phosphate isomerase.

Figure 11.18 ►
Entry of other sugars into glycolysis.





▲ **Figure 11.19**
Conversion of fructose to two molecules of glyceraldehyde 3-phosphate.

The two molecules of glyceraldehyde 3-phosphate produced can then be metabolized to pyruvate by the remaining steps of glycolysis. The metabolism of one molecule of fructose to two molecules of pyruvate produces two molecules of ATP and two molecules of NADH. This is the same yield as the conversion of glucose to pyruvate. Fructose catabolism bypasses phosphofructokinase-1 and its associated regulation. Regulation of pyruvate kinase can still control flux in the pathway.

C. Galactose Is Converted to Glucose 1-Phosphate

The disaccharide lactose, present in milk, is a major source of energy for nursing mammals. In newborns, intestinal lactase catalyzes the hydrolysis of lactose to its components, glucose and galactose, both of which are absorbed from the intestine and transported in the bloodstream.

As shown in Figure 11.20, galactose—the C-4 epimer of glucose—can be converted to glucose 1-phosphate by a pathway in which the nucleotide sugar UDP-glucose (Section 7.2A) is recycled. In the liver, galactokinase catalyzes transfer of a phosphoryl group from ATP to galactose. The galactose 1-phosphate formed in this reaction exchanges with the glucose 1-phosphate moiety of UDP-glucose by cleavage of the pyrophosphate bond of UDP-glucose. This reaction is catalyzed by galactose 1-phosphate uridylyltransferase and produces glucose 1-phosphate and UDP-galactose. Glucose

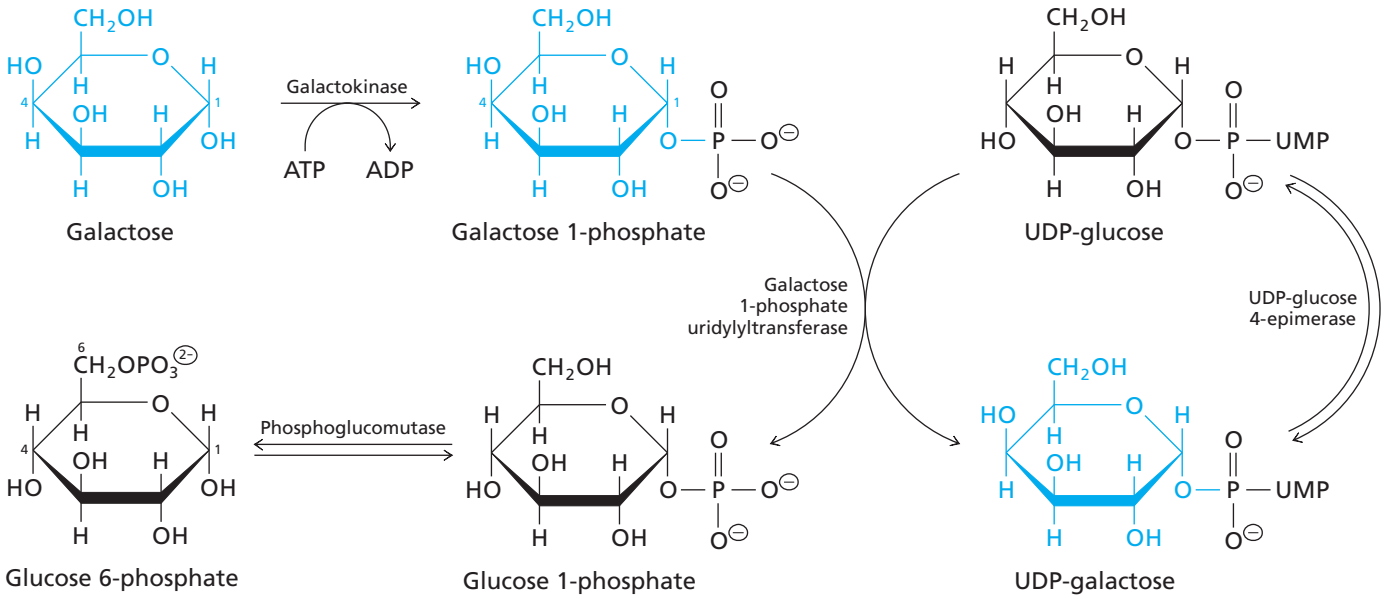
BOX 11.6 A SECRET INGREDIENT

Purified invertase is frequently used in the candy industry to convert sucrose to fructose and glucose. Fructose is sweeter than sucrose and therefore more appealing in some food. The liquid, creamy centers of some chocolates are produced by adding invertase—purified from yeast—to a sucrose mixture. In addition to tasting sweeter, fructose is much less likely to form crystals. The catalytic breakdown of sucrose inside the chocolate usually takes several days or weeks at room temperature.

Look for "invertase" on the labels of food to see more examples of this industrial application of biochemistry, but keep in mind that not all liquid centers in chocolates are due to added invertase.



▶ **Cherry Blossom by Lowney's (Hershey Canada).** The liquid center is due to the presence of added invertase.



▲ Figure 11.20
Conversion of galactose to glucose 6-phosphate. The metabolic intermediate UDP-glucose is recycled in the process. The overall stoichiometry for the pathway is galactose + ATP –> glucose 6-phosphate + ADP.

UDP-Galactose is required for biosynthesis of gangliosides (Section 16.11).

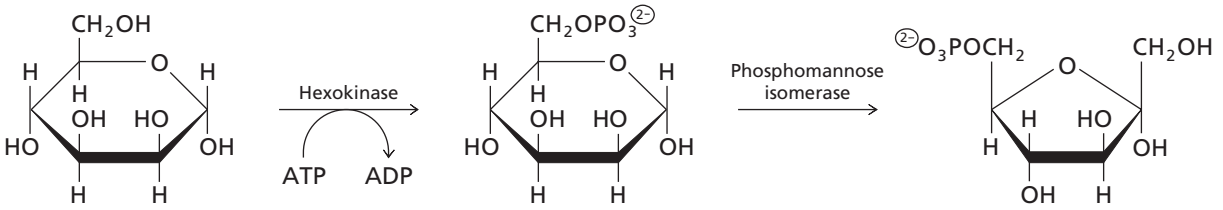
1-phosphate can enter glycolysis after conversion to glucose 6-phosphate in a reaction catalyzed by phosphoglucomutase. UDP-galactose is recycled to UDP-glucose by the action of UDP-glucose 4-epimerase.

The conversion of one molecule of galactose to two molecules of pyruvate produces two molecules of ATP and two molecules of NADH, the same yield as the conversions of glucose and fructose. The required UDP-glucose is formed from glucose and the ATP equivalent UTP, but only small (catalytic) amounts of it are needed since it is recycled.

Infants fed an exclusive diet of milk rely on galactose metabolism for about 20% of their caloric intake. In the most common form of the genetic disorder galactosemia (the inability to properly metabolize galactose), infants are deficient in galactose 1-phosphate uridylyltransferase. In such cases, galactose 1-phosphate accumulates in the cells and this can lead to a compromise in liver function, recognized by the appearance of jaundice (yellowing of the skin). The liver damage is potentially fatal. Other effects include damage to the central nervous system. Screening for galactose 1-phosphate uridylyltransferase in the red blood cells of the umbilical cord allows detection of galactosemia at birth. Many of the most severe effects of this genetic deficiency can be mitigated by a special diet that contains very little galactose and lactose.

The majority of humans undergo a reduction in the level of lactase at about 5 to 7 years of age. This is the normal situation found in most other primates. It parallels the switch from childhood, where mother's milk is a major source of nourishment, to adulthood, where milk is not consumed. In some human populations the production of lactase is not turned off during adolescence. These populations have acquired a mutant gene that continues to synthesize lactase in adults. As a result, individuals in these populations can consume milk products throughout their lives. Northern European populations and their descendants have high proportions of lactase-producing adults.

In normal adults, lactose is metabolized by bacteria in the large intestine, with the production of gases such as CO_2 and H_2 and short-chain acids. The acids can cause diarrhea by increasing the ionic strength of the intestinal fluid. Milk and milk products are usually avoided by people who do not synthesize lactase. Since they do not tolerate diets rich in milk products, they are said to be lactose intolerant although it's worth keeping in mind that this is the normal condition in most mammals, and most humans. Some lactose-intolerant individuals can eat yogurt, in which the lactose has been partially hydrolyzed by the action of an endogenous β-galactosidase of the microorganism in the yogurt culture. A commercially prepared enzyme supplement that contains β-galactosidase from a microorganism can be used to pretreat milk to reduce the lactose content or can be taken when milk products are ingested by lactase-deficient individuals.



▲ **Figure 11.21**
Conversion of mannose to fructose 6-phosphate.

D. Mannose Is Converted to Fructose 6-Phosphate

The aldohexose mannose is obtained in the diet from glycoproteins and certain polysaccharides. Mannose is converted to mannose 6-phosphate by the action of hexokinase. In order to enter the glycolytic pathway, mannose 6-phosphate undergoes isomerization to fructose 6-phosphate in a reaction catalyzed by phosphomannose isomerase. These two reactions are depicted in Figure 11.21.

11.7 The Entner–Doudoroff Pathway in Bacteria

KEY CONCEPT

The classic glycolysis pathway evolved millions of years after the Entner–Douderoff and the gluconeogenesis pathways.

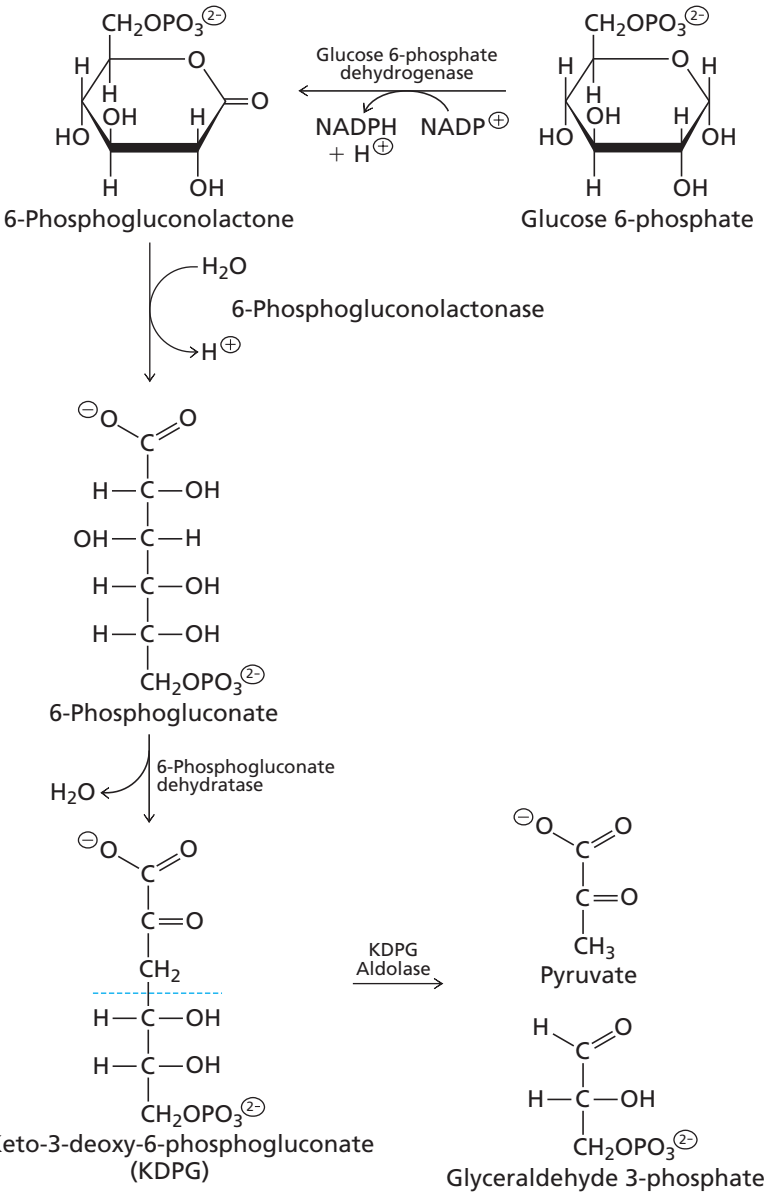
The classic glycolysis pathway is also called the Embden–Meyerhof–Parnas pathway. This pathway is found in all eukaryotes and many species of bacteria. However, a large number of bacterial species do not have phosphofructokinase-1 and cannot convert glucose 6-phosphate to fructose 1,6-*bis*phosphate in the hexose stage of glycolysis.

The hexose stage of classic glycolysis can be bypassed by the Entner–Doudoroff pathway. This pathway begins with the conversion of glucose 6-phosphate to 6-phosphogluconate, a reaction that is catalyzed by two enzymes: glucose 6-phosphate dehdrogenase and 6-phosphogluconolactonase (Figure 11.22). The oxidation of glucose 6-phosphate by glucose 6-phosphate dehydrogenase is coupled to the reduction of $NADP^{\oplus}$. The dehydrogenase and 6-phosphogluconolactonase enzymes are common in almost all species since they are required in the pentose phosphate pathway (Section 12.5). The Entner–Douderoff pathway is the earliest pathway for glucose degradation. The classic glycolysis pathway (EMP) evolved later.

6-Phosphogluconate is converted to 2-keto-3-deoxy-6-phosphogluconate (KDPG) in an unusual dehydration (dehydratase) reaction. KDPG is then split by the action of KDPG aldolase to one molecule of pyruvate and one molecule of glyceraldehyde 3-phosphate. Pyruvate is the end product of glycolysis and glyceraldehyde 3-phosphate can be converted to another molecule of pyruvate by the triose stage of glycolysis. The enzymes of the triose stage of the EMP pathway are found in all species since they are essential for glucose synthesis as well as glycolysis. Note that only one molecule of glyceraldehyde 3-phosphate passes down the bottom half of the glycolytic pathway for every glucose 6-phosphate molecule that enters the Entner–Doudoroff pathway. This means that only one molecule of ATP is produced for every glucose molecule degraded, whereas two ATP molecules are synthesized during glycolysis. Two reducing equivalents (NADH) are produced during glycolysis and two in the ED pathway (NADPH in the first reaction and one molecule of NADH when glyceraldehyde 3-phosphate is converted to 1,3-*bis*phosphoglycerate).

In addition to being the main pathway for glucose degradation in some species, the Entner–Doudoroff pathway is also important in species that possess a complete Embden–Meyeroff–Parnas pathway. The Entner–Doudoroff pathway is used in the metabolism of gluconate and other related organic acids. These metabolites cannot be shunted into the normal glycolytic pathway. Many bacterial species, including *E. coli*, can grow on gluconate as their sole carbon source. Under these conditions the main energy-producing degradation pathway is the Entner–Doudoroff pathway. The first reaction in the ED pathway produces NADPH instead of NADH and many species use the glucose 6-phosphate dehydrogenase reaction as an important source of NADPH reducing equivalents (Section 12.4).

Figure 11.22 ▶
The Entner–Doudoroff pathway.



In Box 12.2 we discuss metabolic diseases associated with glucose 6-phosphate dehydrogenase in humans.

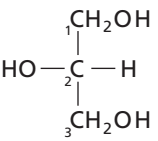
Aldolases cleave hexoses to two 3-carbon compounds. KDPG is the third aldolase we have described.

Summary

1. Glycolysis is a ten-step pathway in which glucose is catabolized to pyruvate. Glycolysis can be divided into a hexose stage and a triose stage. The products of the hexose stage are glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. The triose phosphates interconvert, and glyceraldehyde 3-phosphate is metabolized to pyruvate.
2. For each molecule of glucose converted to pyruvate, there is a net production of two molecules of ATP from ADP + P_i and two molecules of $NAD^{\oplus}$ are reduced to NADH.
3. Under anaerobic conditions in yeast, pyruvate is metabolized to ethanol and CO_2. In some other organisms, pyruvate can be converted to lactate under anaerobic conditions. Both processes use NADH and regenerate $NAD^{\oplus}$.
4. The overall Gibbs free energy change for glycolysis is negative. The steps catalyzed by hexokinase, phosphofructokinase-1, and pyruvate kinase are metabolically irreversible.
5. Glycolysis is regulated at four steps: the transport of glucose into some cells and the reactions catalyzed by hexokinase, phosphofructokinase-1, and pyruvate kinase.
6. Fructose, galactose, and mannose can enter the glycolytic pathway via conversion to glycolytic metabolites.
7. The Entner–Doudoroff pathway is an alternate pathway for glucose catabolism in some bacteria.

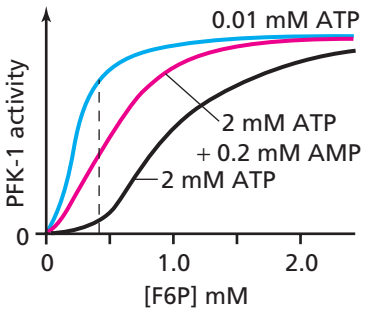
Problems

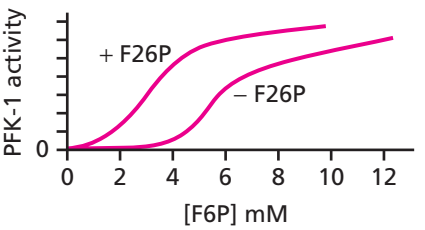
1. Calculate the number of ATP molecules obtained from the anaerobic conversion of each of the following sugars to lactate: (a) glucose, (b) fructose, (c) mannose, and (d) sucrose.
2. (a) Show the positions of the six glucose carbons in the two lactate molecules formed by anaerobic glycolysis. (b) Under aerobic conditions, pyruvate can be decarboxylated to yield acetyl CoA and CO_2. Which carbons of glucose must be labeled with ^{14}C to yield $^{14}CO_2$?
3. If ^{32}P (i.e., isotopically labeled phosphorus) is added to a cell-free liver preparation undergoing glycolysis, will this label be directly incorporated in any glycolytic intermediate or pathway product?
4. Huntington's disease is a member of the "glutamine-repeat" family of diseases. In middle-aged adults the disease causes neurodegenerative conditions, including involuntary movements and dementia. The mutated protein (Huntington protein) contains a polyglutamine region with 40 to 120 glutamines that is thought to mediate a tight binding of this protein to glyceraldehyde 3-phosphate dehydrogenase (GAPDH). If the brain relies almost solely on glucose as an energy source, suggest a role for the Huntington protein in this disease.
5. Fats (triacylglycerols) are a significant source of stored energy in animals and are metabolized initially to fatty acids and glycerol. Glycerol can be phosphorylated by the action of a kinase to produce glycerol 3-phosphate, which is oxidized to produce dihydroxyacetone phosphate.
   (a) Write the reactions for the conversion of glycerol to dihydroxyacetone phosphate.
   (b) The kinase that acts on the prochiral molecule glycerol is stereospecific, leading to production of *L*-glycerol 3-phosphate. Which carbons of glycerol 3-phosphate must be labeled with ^{14}C so that aerobic glycolysis yields acetyl CoA with both carbons labeled?



Glycerol

6. Tumor cells often lack an extensive capillary network and must function under conditions of limited oxygen supply. Explain why these cancer cells take up far more glucose and may overproduce some glycolytic enzymes.
7. Rapid glycolysis during strenuous exercise provides the ATP needed for muscle contraction. Since the lactate dehydrogenase reaction does not produce any ATP, would glycolysis be more efficient if pyruvate rather than lactate were the end product?
8. Why are both hexokinase and phosphofructokinase-1 inhibited by an ATP analog in which the oxygen atom joining the β- and γ-phosphorus atoms is replaced by a methylene group (—CH_2—)?
9. The $\Delta G^{\circ\prime}$ for the aldolase reaction in muscle is +22.8 kJ mol^{-1}. In view of this, why does the aldolase reaction proceed in the direction of glyceraldehyde 3-phosphate and dihydroxyacetone phosphate during glycolysis?
10. For the aldolase reaction, calculate the concentration of fructose 1,6-*bis*phosphate if the concentrations of DHAP and G3P were each: (a) 5 μM, (b) 50 μM, (c) 500 μM.
11. The following plot shows the rate of mammalian phosphofructokinase-1 (PFK-1) activity versus fructose 6-phosphate (F6P) concentration in (a) the presence of ATP, AMP, or both and (b) in the absence or presence of fructose 2,6-*bis*phosphate (F26P). Explain these effects on the reaction rates of PFK-1.





12. Draw a diagram showing how increased intracellular [cAMP] affects the activity of pyruvate kinase in mammalian liver cells.
13. In response to low levels of glucose in the blood, the pancreas produces glucagon, which triggers the adenylyl cyclase signaling pathway in liver cells. As a result, flux through the glycolytic pathway decreases.
   (a) Why is it advantageous for glycolysis to decrease in the liver in response to low blood glucose levels?
   (b) How are the effects of glucagon on glycolysis reversed when the level of glucagon decreases in response to adequate blood glucose levels?
14. Chemoautotrophs growing in the ocean will sometimes have all the enzymes needed for glycolysis even though they will never encounter external glucose. Why?

Selected Readings

Metabolism of Glucose

Alberty, R. A. (1996) Recommendations for nomenclature and tables in biochemical thermodynamics. *Eur. J. Biochem.* 240:1–14.

Cullis, P. M. (1987). Acyl group transfer–phosphoryl group transfer. In *Enzyme Mechanisms*, M. I. Page and A. Williams, eds. (London: Royal Society of Chemistry), pp. 178–220.

Hamori, E. (1975). Illustration of free energy changes in chemical reactions. *J. Chem. Ed.* 52:370–373.

Hoffmann-Ostenhof, O., ed. (1987). *Intermediary Metabolism* (New York: Van Nostrand Reinhold).

Li X, Dash RK, Pradhan RK, Qi F, Thompson M, Vinnakota KC, Wu F, Yang F, Beard DA. (2010) A database of thermodynamic quantities for the reactions of glycolysis and the tricarboxylic acid cycle. J Phys Chem B. 114:16068–16082.

Minakami S. and de Verdier, C-H. (1976) Colorimetric study on human erythrocyte glycolysis. Eur. J. Biochem. 65: 451–460.

Ronimus, R. S., and Morgan, H. W. (2003), Distribution and phylogenies of enzymes of the Embden–Meyerof–Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* 1:199–221.

Seeholzer, S. H., Jaworowski, A., and Rose, I. A. (1991). Enolpyruvate: chemical determination as a pyruvate kinase intermediate. *Biochem.* 30:727–732.

St-Jean, M., Blonski, C., and Sygush, J. (2009). Charge stabilization and entropy reduction of central lysine residues in fructose-*bis*phosphate aldolase. *Biochem.* 48:4528–4537.

Regulation of Glycolysis

Depré, C., Rider, M. H., and Hue, L. (1998). Mechanisms of control of heart glycolysis. *Eur. J. Biochem.* 258:277–290.

Engström, L., Ekman, P., Humble, E., and Zetterqvist, Ö. (1987). Pyruvate kinase. In *The Enzymes*, Vol. 18, P. D. Boyer and E. Krebs, eds. (San Diego: Academic Press), pp. 47–75.

Gould, G. W., and Holman, G. D. (1993). The glucose transporter family: structure, function and tissue-specific expression. *Biochem. J.* 295:329–341.

Pessin, J. E., Thurmond, D. C., Elmendorf, J. S., Coker, K. J., and Okada, S. (1999). Molecular basis of insulin-stimulated GLUT4 vesicle trafficking. Location! Location! Location! *J. Biol. Chem.* 274:2593–2596.

Pilkis, S. J., Claus, T. H., Kurland, I. J., and Lange, A. J. (1995). 6-Phosphofructo-2-kinase/fructose-2,6-*bis*phosphatase: a metabolic signaling enzyme. *Annu. Rev. Biochem.* 64:799–835.

Pilkis, S. J., El-Maghrabi, M. R., and Claus, T. H. (1988). Hormonal regulation of hepatic gluconeogenesis and glycolysis. *Annu. Rev. Biochem.* 57:755–783.

Pilkis, S. J., and Granner, D. K. (1992). Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annu. Rev. Physiol.* 54:885–909.

Van Schaftingen, E. (1993). Glycolysis revisited. *Diabetologia* 36:581–588.

Yamada, K., and Noguchi, T. (1999). Nutrient and hormonal regulation of pyruvate kinase gene expression. *Biochem. J.* 337:1–11.

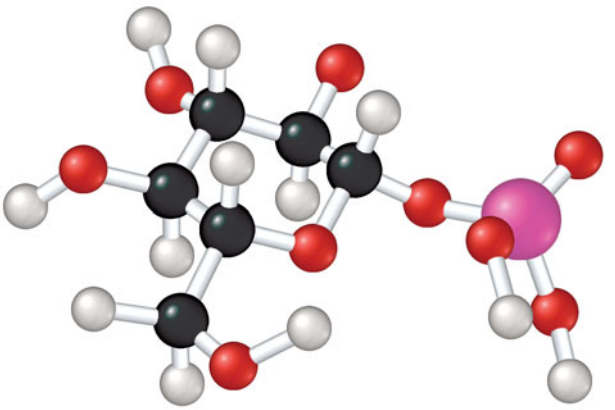
Metabolism of Other Sugars

Álvaro-Benito, M., Polo, A., González, B., Fernández-Lobato, M., and Sanz-Aparicio, J. (2010). Structural and kinetic analysis of *Schwanniomyces occidentalis* invertase reveals a new oligomerization pattern and the role of its supplementary domain in substrate binding. *J. Biol. Chem.* 285:13930–13941; doi:10.1074/jbc.M109.095430

Frey, P. A. (1996). The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J.* 10:461–470.

Itan, Y., Jones, B. L., Ingram, C. J. E., Swallow, D. M., and Thomas, M. G. (2010). A worldwide correlation of lactase persistence phenotypes and genotypes. *BMC Evol. Biol.* 10:36; www.biomedcentral.com/1471-2148/10/36





Gluconeogenesis, The Pentose Phosphate Pathway, and Glycogen Metabolism

We have seen that the catabolism of glucose is central to energy metabolism in some cells. In contrast, all species can synthesize glucose from simple two-carbon and three-carbon precursors by gluconeogenesis (literally, the formation of new glucose). Some species, notably photosynthetic organisms, can make these precursors by fixing carbon dioxide leading to the net synthesis of glucose from inorganic compounds. In our discussion of gluconeogenesis in this chapter we must keep in mind that every glucose molecule used in glycolysis had to be synthesized in some species.

The pathway for gluconeogenesis shares some steps with glycolysis, the pathway for glucose degradation, but four reactions specific to the gluconeogenic pathway are not found in the degradation pathway. These reactions replace the metabolically irreversible reactions of glycolysis. These opposing sets of reactions are an example of separate, regulated pathways for synthesis and degradation (Section 10.2).

In addition to fueling the production of ATP (via glycolysis and the citric acid cycle), glucose is also a precursor of the ribose and deoxyribose moieties of nucleotides and deoxynucleotides. The pentose phosphate pathway is responsible for the synthesis of ribose as well as the production of reducing equivalents in the form of NADPH.

Glucose availability is controlled by regulating the uptake and synthesis of glucose and related molecules and by regulating the synthesis and degradation of storage polysaccharides composed of glucose residues. Glucose is stored as glycogen in bacteria and animals and as starch in plants. Glycogen and starch can be degraded to release glucose monomers that can fuel energy production via glycolysis or serve as precursors in biosynthesis reactions. The metabolism of glycogen will illustrate another example of opposing, regulated pathways.

Although the reaction we had found would be viewed today as utterly trivial, it came nevertheless as a great surprise, because, at that time, nobody could imagine that the phosphorylation of an enzyme could be involved in its regulation.

—Eddy Fischer, *Memories of Ed Krebs* (2010)

Top: The Cori ester, α-D-glucopyranose 1-phosphate.

In mammals, gluconeogenesis, the pentose phosphate pathway, and glycogen metabolism are closely and coordinately regulated in accordance with the moment-to-moment requirements of the organism. In this chapter, we review these pathways and examine some of the mechanisms for regulating glucose metabolism in mammalian cells. The regulation of glucose and glycogen metabolism in mammals is important from a historical perspective because it was the first example of a signal transduction mechanism.

12.1 Gluconeogenesis

As stated in the introduction, all organisms have a pathway for glucose biosynthesis, or gluconeogenesis. This is true even for animals that use exogenous glucose as an important energy source because glucose may not always be available from external sources or intracellular stores. For example, large mammals that have not eaten for 16 to 24 hours have depleted their liver glycogen reserves and need to synthesize glucose to stay alive because glucose is required for the metabolism of certain tissues, for example, brain. Some mammalian tissues, primarily liver and kidney, can synthesize glucose from simple precursors such as lactate and alanine. Under fasting conditions, gluconeogenesis supplies almost all of the body's glucose. When exercising under anaerobic conditions, muscle converts glucose to pyruvate and lactate, which travel to the liver and are converted to glucose. Brain and muscle consume much of the newly formed glucose. Bacteria can convert many nutrients to phosphate esters of glucose and to glycogen.

It is convenient to consider pyruvate as the starting point for the synthesis of glucose. The pathway for gluconeogenesis from pyruvate is compared to the glycolytic pathway in Figure 12.1. Note that many of the intermediates and enzymes are identical. All seven of the near-equilibrium reactions of glycolysis proceed in the reverse direction during gluconeogenesis. Enzymatic reactions unique to gluconeogenesis are required for the three metabolically irreversible reactions of glycolysis. These irreversible glycolytic reactions are catalyzed by pyruvate kinase, phosphofructokinase-1, and hexokinase. In the biosynthesis direction these reactions are catalyzed by different enzymes.

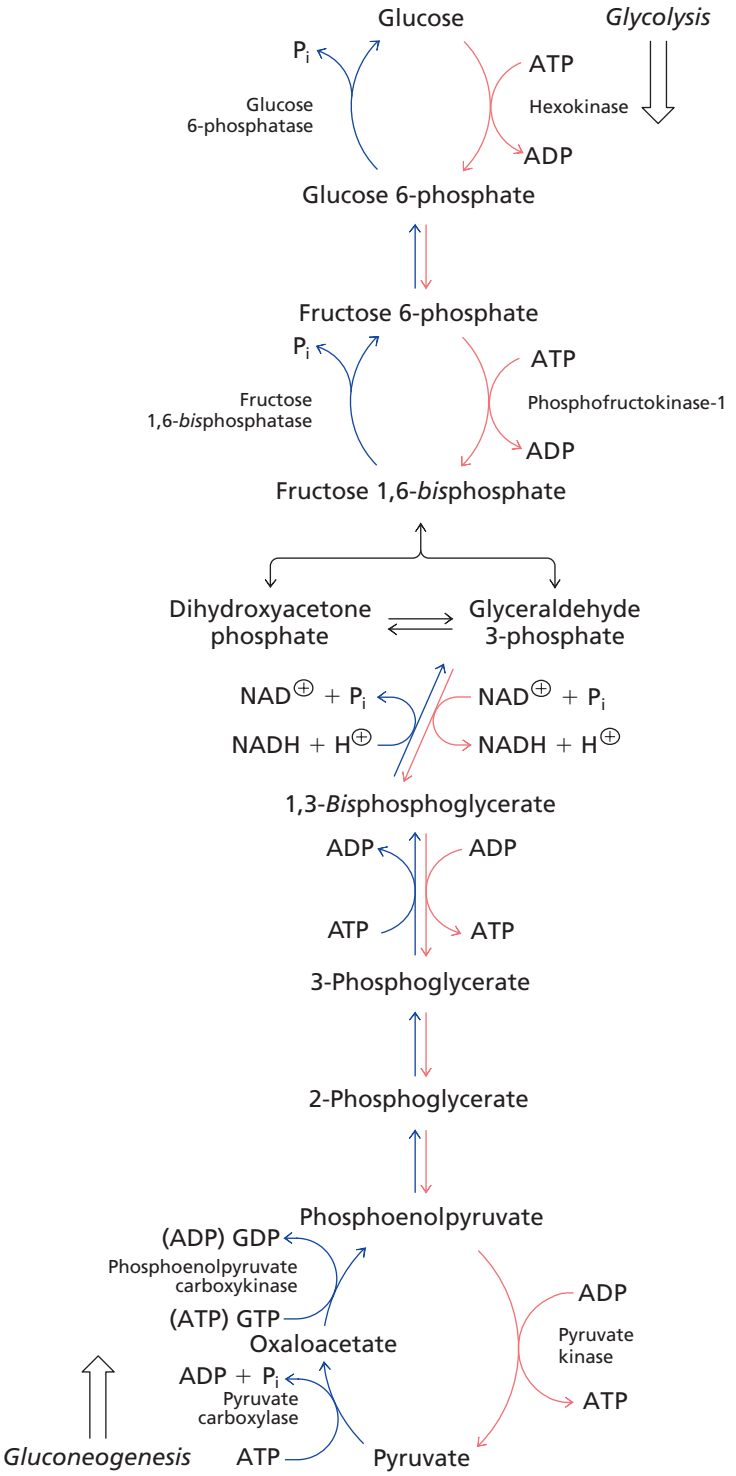
Although all species have a gluconeogenesis pathway, they don't all have the glycolysis pathway (Section 11.7). This is especially true of bacterial species that diverged very early in the evolution of prokaryotes. Thus, it seems like gluconeogenesis is the more ancient pathway, which makes sense since there has to be a source of glucose before pathways for its degradation can evolve. Since the biosynthesis pathway evolved first, it is appropriate to think of the glycolytic enzymes as bypass enzymes. These enzymes, especially phosphofructokinase-1, evolved in order to bypass the metabolically irreversible reactions of gluconeogenesis.

In the next section, we discuss how other precursors enter the pathway.

The synthesis of one molecule of glucose from two molecules of pyruvate requires four ATP and two GTP molecules as well as two molecules of NADH. The net equation for gluconeogenesis is

$$\begin{aligned} 2\ \text{Pyruvate} + 2\ \text{NADH} + 4\ \text{ATP} + 2\ \text{GTP} + 6\ \text{H}_2\text{O} + 2\ \text{H}^{\oplus} \longrightarrow \\ \text{Glucose} + 2\ \text{NAD}^{\oplus} + 4\ \text{ADP} + 2\ \text{GDP} + 6\ \text{P}_i \end{aligned} \qquad (12.1)$$

Four ATP equivalents are needed to overcome the thermodynamic barrier to the formation of two molecules of the energy-rich compound phosphoenolpyruvate from two molecules of pyruvate. Recall that in glycolysis the conversion of phosphoenolpyruvate to pyruvate is a metabolically irreversible reaction catalyzed by pyruvate kinase. In the catabolic direction this reaction is coupled to the synthesis of ATP. Two ATP molecules are required to carry out the reverse of the glycolytic reaction catalyzed by phosphoglycerate kinase. In the hexose stage of gluconeogenesis, no energy is recovered in the steps that convert fructose 1,6-*bis*phosphate to glucose because fructose 1,6-*bis*phosphate is not a "high energy" intermediate. Recall that glycolysis consumes two ATP molecules and generates four, for a net yield of two ATP equivalents and two molecules of NADH. Contrast this with the synthesis of one molecule of glucose by gluconeogenesis consuming a total of six ATP equivalents and two molecules of NADH. As expected, the biosynthesis of glucose requires energy and its degradation releases energy.



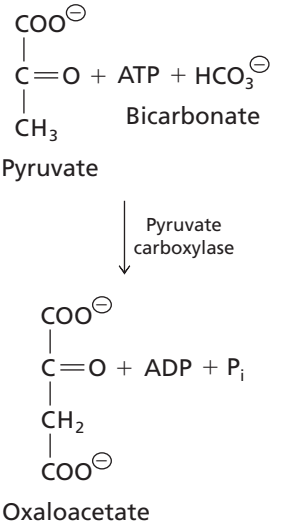
◀ **Figure 12.1**
Comparison of gluconeogenesis and glycolysis. There are four metabolically irreversible reactions of gluconeogenesis (blue). These are the reactions catalyzed by three different enzymes in glycolysis (red). Both pathways include a triose stage and a hexose stage. Two molecules of pyruvate are therefore required to produce one molecule of glucose.

A. Pyruvate Carboxylase

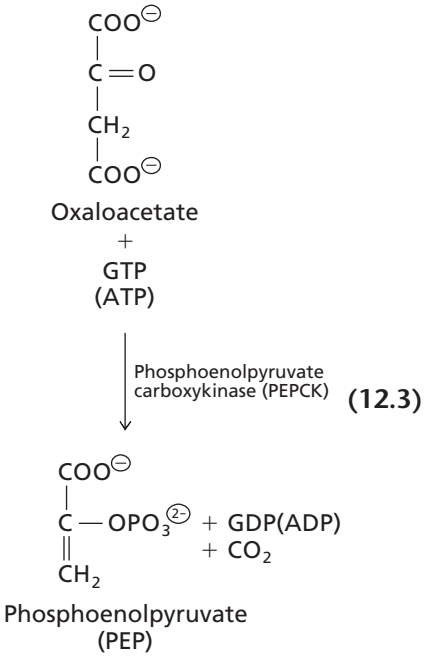
We begin our examination of the individual steps in the conversion of pyruvate to glucose with the two enzymes required for synthesis of phosphoenolpyruvate. The two steps involve a carboxylation followed by decarboxylation. In the first step, pyruvate carboxylase catalyzes the conversion of pyruvate to oxaloacetate. The reaction is coupled to the hydrolysis of one molecule of ATP (Figure 12.2).

Pyruvate carboxylase is a large, complex, enzyme composed of four identical subunits. Each subunit has a biotin prosthetic group covalently linked to a lysine residue. The biotin is required for the addition of bicarbonate to pyruvate. Pyruvate carboxylase catalyzes a metabolically irreversible reaction—it can be allosterically activated by acetyl CoA. This is the only regulatory mechanism known for the enzyme. Accumulation of

Pyruvate carboxylase is a biotin-containing enzyme. The reaction mechanism was described in Section 7.10.



▲ **Figure 12.2**
Pyruvate carboxylase reaction.



▲ **Figure 12.3**
Phosphoenolpyruvate carboxykinase reaction.

acetyl CoA indicates that it is not being efficiently metabolized by the citric acid cycle. Under these conditions, pyruvate carboxylase is stimulated in order to direct pyruvate to oxaloacetate instead of acetyl CoA. Oxaloacetate can enter the citric acid cycle or serve as a precursor for glucose biosynthesis.

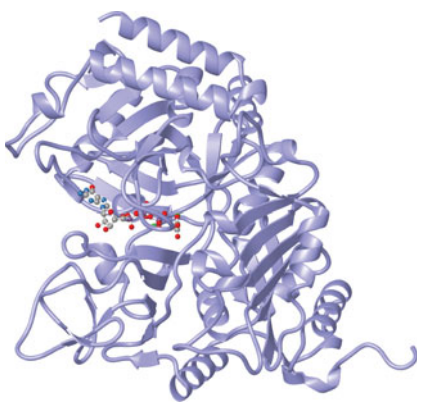
Bicarbonate is one of the substrates in the reaction shown in Figure 12.2. Bicarbonate is formed when carbon dioxide dissolves in water so the reaction is sometimes written with CO_2 as a substrate. The pyruvate carboxylase reaction plays an important role in fixing carbon dioxide in bacteria and some eukaryotes. This role may not be so obvious when we examine gluconeogenesis since the carbon dioxide is released in the very next reaction; however, much of the oxaloacetate that is made is not used for gluconeogenesis. Instead, it replenishes the pool of citric acid cycle intermediates that serve as precursors to the biosynthesis of amino acids and lipids (Section 13.7).

B. Phosphoenolpyruvate Carboxykinase

Phosphoenolpyruvate carboxykinase (PEPCK) catalyzes the conversion of oxaloacetate to phosphoenolpyruvate (Figure 12.3). This is a well-studied enzyme with an induced-fit binding mechanism similar to that described for yeast hexokinase (Section 6.5C) and citrate synthase (Section 13.3A).

There are two different versions of PEPCK. The enzyme found in bacteria, protists, fungi, and plants uses ATP as the phosphoryl group donor in the decarboxylation reaction. The animal version uses GTP. In most species, the enzyme displays no allosteric kinetic properties and has no known physiological modulators. Its activity is most often affected by controls at the level of transcription of its gene. The level of PEPCK activity in cells influences the rate of gluconeogenesis. This is especially true in mammals where gluconeogenesis is mostly confined to cells in the liver, kidneys, and small intestine. During fasting in mammals, prolonged release of glucagon from the pancreas leads to continued elevation of intracellular cAMP, that triggers increased transcription of the PEPCK gene in the liver and increased synthesis of PEPCK. After several hours, the amount of PEPCK rises and the rate of gluconeogenesis increases. Insulin, abundant in the fed state, acts in opposition to glucagon at the level of the gene reducing the rate of synthesis of PEPCK.

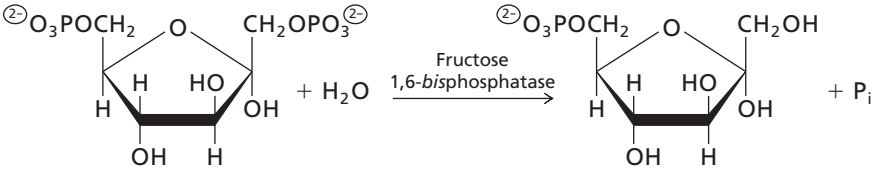
The two-step synthesis of phosphoenolpyruvate from pyruvate is common in most eukaryotes, including humans. This is the main reason why it's usually shown when the gluconeogenesis pathway is described (Figure 12.1). However, many species of bacteria can convert pyruvate directly to phosphoenolpyruvate in an ATP-dependent reaction catalyzed by phosphoenolpyruvate synthetase (Figure 12.4). The products of this reaction include AMP and P_i. The second phosphoryl from ATP is transferred to pyruvate. Thus, two ATP equivalents are used in the conversion of pyruvate to phosphoenolpyruvate. This is a much more efficient route than the eukaryotic two-step pathway catalyzed by pyruvate carboxylase and PEPCK. The presence of phosphoenolpyruvate synthetase in bacterial cells is due to the fact that efficient gluconeogenesis is much more important in bacteria than in eukaryotes.



▲ **Phosphoenolpyruvate carboxykinase from rat (*Rattus norvegicus*).** The closed active site contains a bound GTP molecule, a molecule of oxaloacetate, and two Mn^{2+} ions (pink). [PDB 3DT4]

C. Fructose 1,6-*bis*phosphatase

The reactions of gluconeogenesis between phosphoenolpyruvate and fructose 1,6-*bis*phosphate are simply the reverse of the near-equilibrium reactions of glycolysis. The next reaction in the glycolysis pathway—catalyzed by phosphofructokinase-1—is metabolically irreversible. In the biosynthesis direction, this reaction is catalyzed by the third enzyme specific to gluconeogenesis, fructose 1,6-*bis*phosphatase. This enzyme catalyzes the conversion of fructose 1,6-*bis*phosphate to fructose 6-phosphate.



Fructose 1,6-*bis*phosphate — Fructose 6-phosphate

(12.2)

BOX 12.1 SUPERMOUSE

Richard Hanson's group at Case Western Reserve University in Cleveland, Ohio, USA, created a form of supermouse by adding extra copies of the cytoplasmic phosphoenopyruvate carboxykinase gene. The homozygous transgenic mice expressed 10× more PEPCK in their skeletal muscle. They were hyperactive, aggressive, and capable of running for extended periods of time on a mouse treadmill (up to 5 km without stopping!). They ate more than control mice but were significantly smaller.

The rodent athletes converted prodigious amounts of oxaloacetate into phosphoenolpyruvate and subsequently to intermediates in the gluconeogenesis pathway, including glucose. Their muscle cells had many more mitochondria than the cells of normal mice.

The biochemical explanation of this hyperactivity is not completely understood. It's probably due to effects on the citric acid cycle (Chapter 13). This allows increased flux in that pathway leading ultimately to higher levels of ATP. When asked whether this genetic modification would be a good way of creating superior human athletes, Hanson and Hakimi (2008) replied, "The PEKCK-C^{mus} mice are very aggressive; the world needs *less*, not more aggression," besides, the creation of such transgenic humans is ". . . neither ethical nor possible."

Watch the video at: youtube.com/watch?v=4PXC_mctsgY

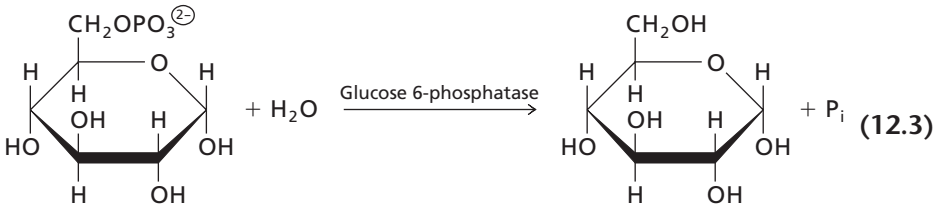


▲ Mighty Mouse © CBS Operations.

As you might expect, hydrolysis of the phosphate ester in this reaction is associated with a large negative standard Gibbs free energy change ($\Delta G°'$). The actual Gibbs free energy change *in vivo* is also negative because this reaction is metabolically irreversible. The mammalian enzyme displays sigmoidal kinetics and is allosterically inhibited by AMP and by the regulatory molecule fructose 2,6-*bis*phosphate. Thus, the reaction cannot reach equilibrium. Recall that fructose 2,6-*bis*phosphate is a potent activator of phosphofructokinase-1, the enzyme that catalyzes the formation of fructose 1,6-*bis*phosphate in glycolysis (Section 11.5C). The two enzymes that catalyze the interconversion of fructose 6-phosphate and fructose 1,6-*bis*phosphate are reciprocally controlled by the concentration of fructose 2,6-*bis*phosphate (see Section 12.6C).

D. Glucose 6-phosphatase

The final step of gluconeogenesis is the hydrolysis of glucose 6-phosphate to form glucose. The enzyme is glucose 6-phosphatase.



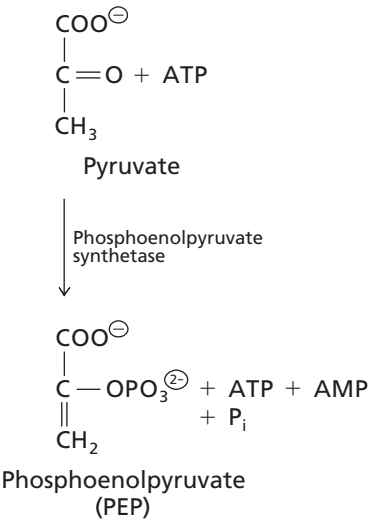
(12.3)

Glucose 6-phosphate Glucose

Although we present glucose as the final product of gluconeogenesis, this is not true in all species. In most cases, the biosynthetic pathway ends with glucose 6-phosphate. This product is an activated form of glucose. It becomes the substrate for additional carbohydrate pathways leading to synthesis of glycogen (Section 12.6), starch and sucrose (Section 15.11), pentose sugars (Section 12.5), and other hexoses.

In mammals, glucose is an important end product of gluconeogenesis since it serves as an energy source for glycolysis in many tissues. Glucose is made in the cells of the liver, kidneys, and small intestine and exported to the bloodstream. In these cells, glucose 6-phosphatase is bound to the endoplasmic reticulum with its active site in the lumen. The enzyme is part of a complex that includes a glucose 6-phosphate transporter (G6PT) and a phosphate transporter. G6PT moves glucose 6-phosphate from the

▼ Figure 12.4
Phosphoenolpyruvate synthetase reaction.



Additional effects of glucagon and insulin are described in Section 12.6C.

Defects in the activities of glucose 6-phosphatase or glucose 6-phosphate transporter cause von Gierke disease (Section 12.8).

cytosol to the interior of the ER where it is hydrolyzed to glucose and inorganic phosphate. Phosphate is returned to the cytosol and glucose is transported to the cell surface (and the bloodstream) via the secretory pathway.

The other enzymes required for gluconeogenesis are found, at least in small amounts, in many mammalian tissues. Glucose 6-phosphatase is found only in cells from the liver, kidneys, and small intestine, so only these tissues can synthesize free glucose. Cells of tissues that lack glucose 6-phosphatase retain glucose 6-phosphate for internal carbohydrate metabolism.

12.2 Precursors for Gluconeogenesis

KEY CONCEPT

Mammalian fuel metabolism is an important subset of biochemistry because it helps us to understand our own bodies.

The main substrates for glucose 6-phosphate synthesis are pyruvate, citric acid cycle intermediates, three-carbon intermediates in the pathway (e.g. glyceraldehyde 3-phosphate), and two-carbon compounds such as acetyl CoA. Acetyl CoA is converted to oxaloacetate in the glyoxylate cycle, that operates in bacteria, protists, fungi, plants, and some animals (Section 13.8). Some organisms can fix inorganic carbon by incorporating it into two-carbon and three-carbon organic compounds (e.g., Calvin cycle, Section 15.4). These compounds enter the gluconeogenesis pathway resulting in net synthesis of glucose from CO_2.

Mammalian biochemistry is focused on fuel metabolism and biosynthesis of glucose from simple precursors and is it usually discussed in that context. The major gluconeogenic precursors in mammals are lactate and most amino acids, especially alanine. Glycerol, which is produced from the hydrolysis of triacylglycerols, is also a substrate for gluconeogenesis. Glycerol enters the pathway after conversion to dihydroxyacetone phosphate. Precursors arising in nongluconeogenic tissues must first be transported to the liver to be substrates for gluconeogenesis.

A. Lactate

Glycolysis generates large amounts of lactate in active muscle and red blood cells. Lactate from these and other sources enters the bloodstream and travels to the liver where it is converted to pyruvate by the action of lactate dehydrogenase. Pyruvate can then be a substrate for gluconeogenesis. Glucose produced by the liver enters the bloodstream for delivery to peripheral tissues, including muscle and red blood cells. This sequence is known as the **Cori cycle** (Figure 12.5). The conversion of lactate to glucose requires energy, most of which is derived from the oxidation of fatty acids in the liver. Thus, the Cori cycle transfers chemical potential energy in the form of glucose from the liver to the peripheral tissues.

B. Amino Acids

The carbon skeletons of most amino acids are catabolized to pyruvate or intermediates of the citric acid cycle. The end products of these catabolic pathways can serve directly as precursors for synthesis of glucose 6-phosphate in cells that are capable of gluconeogenesis. In peripheral mammalian tissues, pyruvate formed from glycolysis or amino acid catabolism

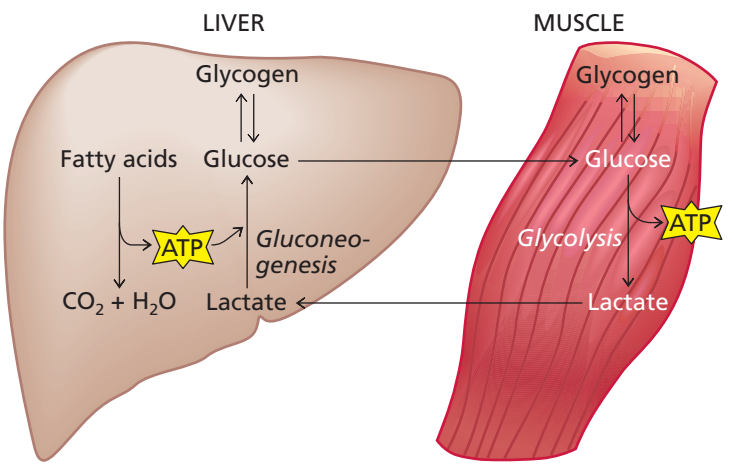


Figure 12.5 ►
Cori cycle. Glucose is converted to L-lactate in muscle cells. Some of this lactate is secreted and passes via the bloodstream to the liver. Lactate is converted to glucose in the liver and the glucose is secreted into the bloodstream where it is taken up by muscle cells. Both tissues are capable of synthesizing glycogen and mobilizing it.

must be transported to the liver before it can be used in glucose synthesis. The Cori cycle is one way of accomplishing this transfer by converting pyruvate to lactate in muscle and re-converting it to pyruvate in liver cells. The glucose–alanine cycle is a similar transport system (Section 17.9B). Pyruvate can also accept an amino group from an α-amino acid, such as glutamate, forming alanine by the process of transamination (Section 7.2B) (Figure 12.6).

Alanine travels to the liver, where it undergoes transamination with α-ketoglutarate to re-form pyruvate for gluconeogenesis. Amino acids become a major source of carbon for gluconeogenesis during fasting when glycogen supplies are depleted.

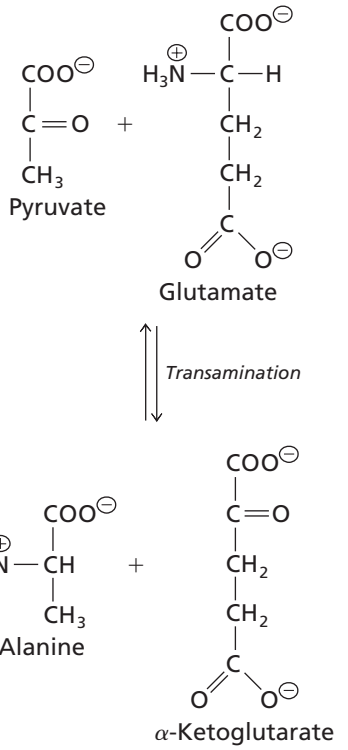
The carbon skeleton of aspartate is also a precursor of glucose. Aspartate is the amino group donor in the urea cycle, a pathway that eliminates excess nitrogen from the cell (Section 17.9B). Aspartate is converted to fumarate in the urea cycle and then fumarate is hydrated to malate that is oxidized to oxaloacetate. In addition, the transamination of aspartate with α-ketoglutarate directly generates oxaloacetate.

C. Glycerol

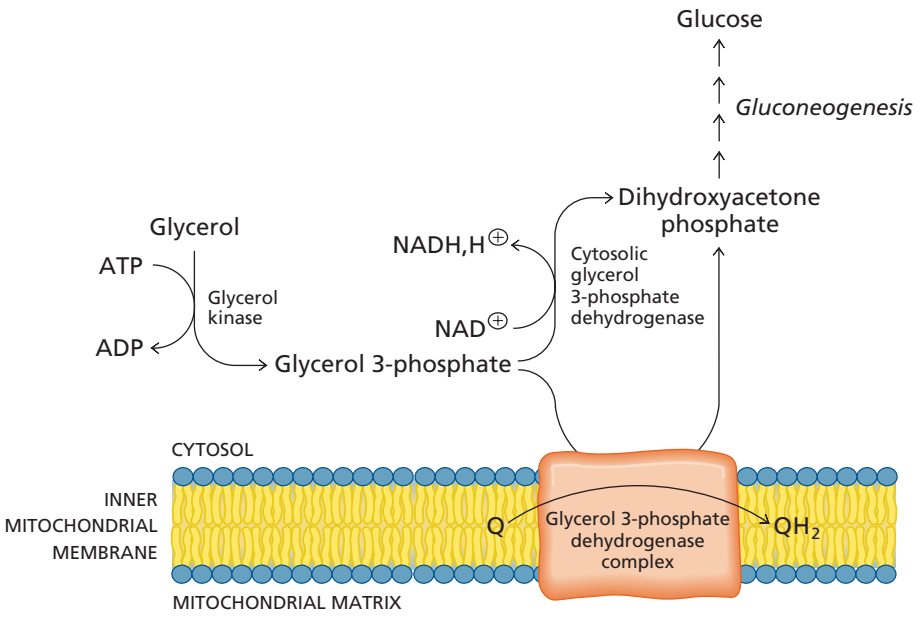
The catabolism of triacylglycerols produces glycerol and acetyl CoA. As mentioned earlier, acetyl CoA contributes to the net formation of glucose through reactions of the glyoxylate cycle (Section 13.8). The glyoxylate cycle does not contribute to net synthesis of glucose from lipids in mammalian cells. Glycerol, however, can be converted to glucose by a route that begins with phosphorylation to glycerol 3-phosphate, catalyzed by glycerol kinase (Figure 12.7). Glycerol 3-phosphate enters gluconeogenesis after conversion to dihydroxyacetone phosphate. This oxidation can be catalyzed by a flavin containing glycerol 3-phosphate dehydrogenase complex embedded in the inner mitochondrial membrane. The outer face of this enzyme binds glycerol 3-phosphate and electrons are passed to ubiquinone (Q) and subsequently to the rest of the membrane-associated electron transport chain. The oxidation of glycerol 3-phosphate can also be catalyzed by the $NAD^{\oplus}$ requiring cytosolic glycerol 3-phosphate dehydrogenase, although this enzyme is usually associated with the reverse reaction for making glycerol. Both enzymes are found in the liver, the site of most gluconeogenesis in mammals.

D. Propionate and Lactate

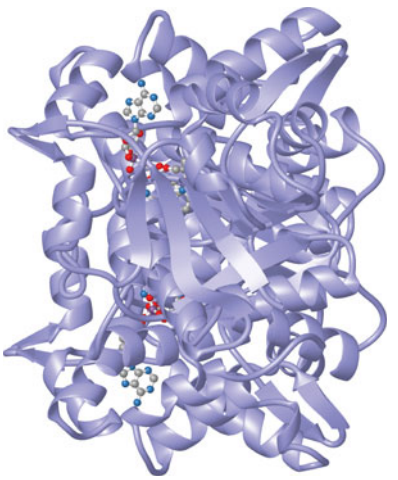
In ruminants—cattle, sheep, giraffes, deer, and camels—the propionate and lactate produced by the microorganisms in the rumen (chambered stomach) are absorbed and



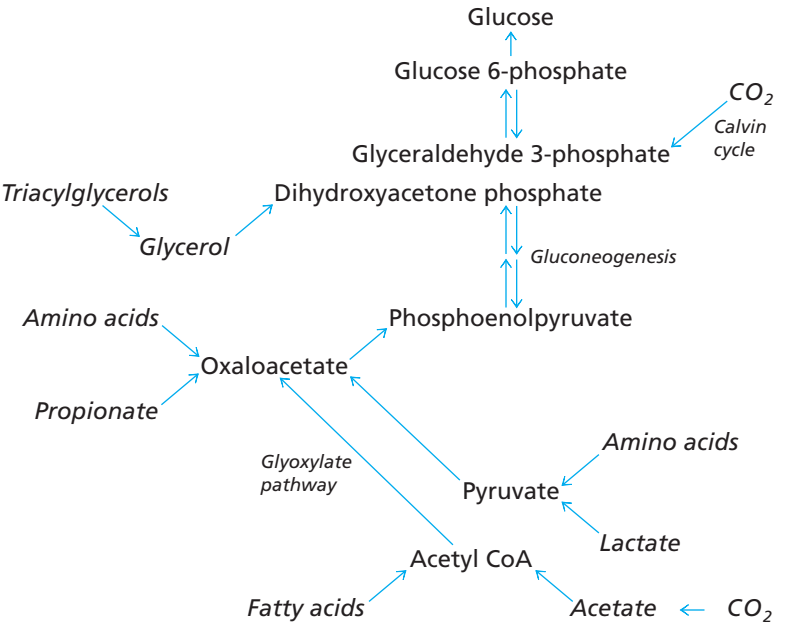
▲ **Figure 12.6**
Conversion of pyruvate to alanine. Pyruvate can be converted to alanine in peripheral tissues. Alanine is secreted into the bloodstream where it is taken up by liver cells and converted back to pyruvate by the same transamination reaction. Pyruvate then serves as a precursor for gluconeogenesis.



▲ **Figure 12.7**
Gluconeogenesis from glycerol. Glycerol 3-phosphate can be oxidized by a glycerol 3-phosphate dehydrogenase complex in the mitochondrial membrane. A cytoplasmic version of this enzyme interconverts dihydroxyacetone phosphate and glycerol 3-phosphate.



▲ **Glycerol 3-phosphate dehydrogenase.** This is the human (*Homo sapiens*) version of the cytosolic enzyme containing DHAP and $NAD^{\oplus}$ at the active site. The structure of the membrane-bound version is not known. [PDB 1WPQ]



► **Precursors for gluconeogenesis.** The glyoxylate pathway, the Calvin cycle, and fixation of CO_2 into acetate, do not occur in mammals. Propionate is produced by microorganisms in the rumen of ruminants.

enter the gluconeogenesis pathway. Propionate is converted to propionyl CoA and then to succinyl CoA. These reactions will be covered in the chapter on lipid metabolism (Section 16.3). Succinyl CoA is an intermediate of the citric acid cycle that can be metabolized to oxaloacetate. Lactate from the rumen is oxidized to pyruvate.

E. Acetate

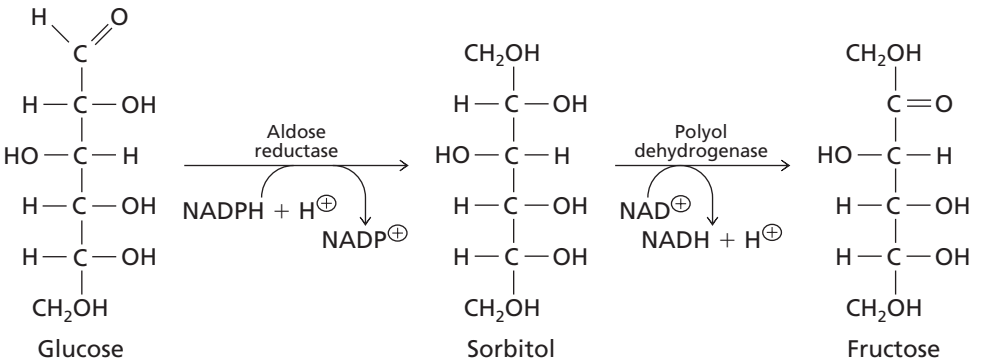
Many species can utilize acetate as their main source of carbon. They can convert acetate to acetyl CoA that serves as the precursor to oxaloacetate. Bacteria and

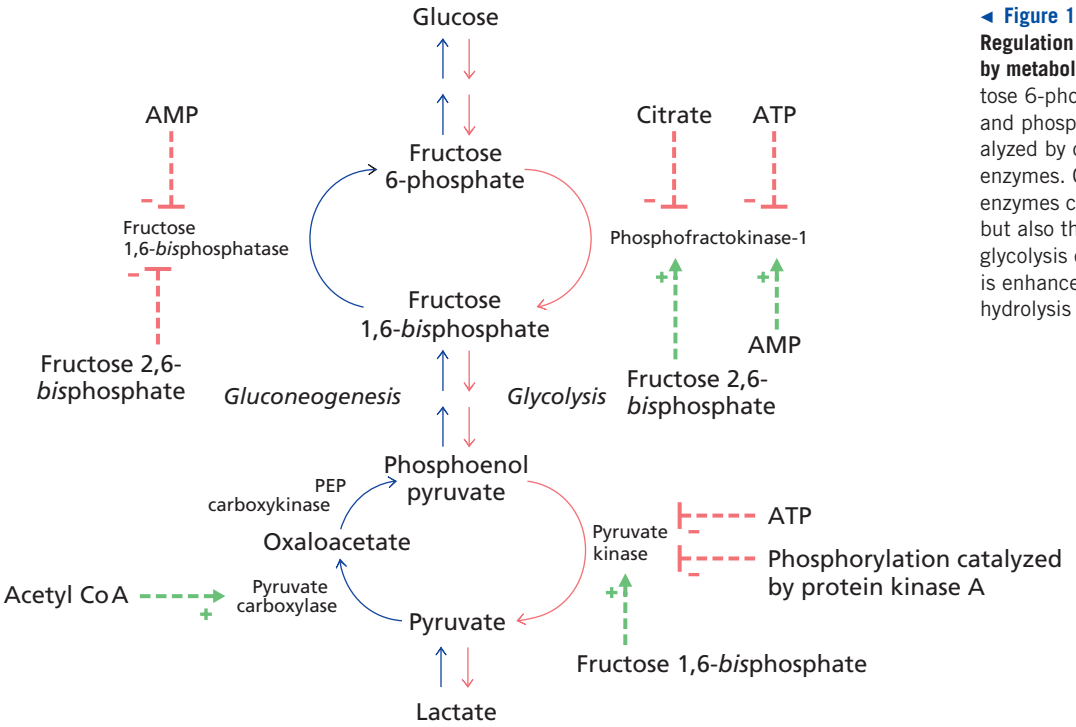
BOX 12.2 GLUCOSE IS SOMETIMES CONVERTED TO SORBITOL

In most animals, glucose—whether from gluconeogenesis, food, or glycogenolysis—is usually oxidized or reincorporated into glycogen. However, in some mammalian tissues (including, testes, pancreas, brain and the lens of the eye), glucose can be converted to fructose as shown in the pathway below. Aldose reductase catalyzes the reduction of glucose to produce sorbitol and polyol dehydrogenase catalyzes the oxidation of sorbitol to fructose. This short pathway supplies essential fructose for some cells. For example, fructose is the main fuel for sperm cells.

Aldose reductase has a high K_m value for glucose so flux through this pathway is normally low and glucose is usually metabolized by glycolysis. When the concentration of glucose is higher than usual (e.g., in individuals with diabetes), increased amounts of sorbitol are produced in tissues such as the lens. There is less polyol dehydrogenase activity than aldose reductase activity so sorbitol can accumulate. Since membranes are relatively impermeable to sorbitol, the resulting change in the osmolarity of the cells causes aggregation and precipitation of lens proteins leading to cataracts—opaque regions in the lens.

▼ **Production of sorbitol from glucose.**





◀ **Figure 12.8**
Regulation of glycolysis and gluconeogenesis by metabolites. The interconversions of fructose 6-phosphate/fructose 1,6-*bis*phosphate and phosphoenolpyruvate/pyruvate are catalyzed by different metabolically irreversible enzymes. Changing the activity of any of the enzymes can affect not only the rate of flux but also the direction of flux toward either glycolysis or gluconeogenesis. The net effect is enhanced regulation at the expense of the hydrolysis of ATP.

single-celled eukaryotes such as yeast utilize acetate as a precursor for gluconeogenesis. Some species of bacteria can synthesize acetate directly from CO_2. In those species the gluconeogenesis pathway provides a route for the synthesis of glucose from inorganic substrates.

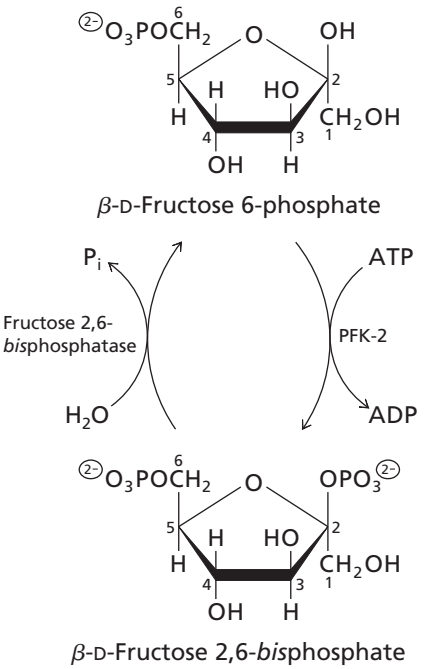
12.3 Regulation of Gluconeogenesis

Gluconeogenesis is carefully regulated *in vivo*. Glycolysis and gluconeogenesis are opposing catabolic and anabolic pathways that share some enzymatic steps but certain reactions are unique to each pathway. For example, phosphofructokinase-1 catalyzes a reaction in glycolysis and fructose 1,6-*bis*phosphatase catalyzes the opposing reaction in gluconeogenesis; both reactions are metabolically irreversible. Usually, only one of the enzymes is active at any given time.

Short-term regulation of gluconeogenesis (regulation that occurs within minutes and does not involve the synthesis of new protein) is exerted at two sites—the reactions involving pyruvate and phosphoenolpyruvate and those that interconvert fructose 1,6-*bis*phosphate and fructose 6-phosphate (Figure 12.8). When there are two enzymes catalyzing the same reaction (in different directions), modulating the activity of either enzyme can alter the flux through the two opposing pathways. For example, inhibiting phosphofructokinase-1 stimulates gluconeogenesis since more fructose 6-phosphate enters the pathway leading to glucose rather than being converted to fructose 1,6-*bis*phosphate. Simultaneous control of fructose 1,6-*bis*phosphatase also regulates the flux of fructose 1,6-*bis*phosphate toward either glycolysis or gluconeogenesis.

We've encountered phosphofructokinase-1 (PFK-1) several times, most notably in the previous chapter (Section 11.5C) and in our discussion of allostery (Section 5.9). Now it's time to examine the effect of the allosteric effector, fructose 2,6-*bis*phosphate, on the activity of PFK-1.

Fructose 2,6-*bis*phosphate is formed from fructose 6-phosphate by the action of the enzyme phosphofructokinase-2 (PFK-2) (Figure 12.9). In mammalian liver, a different



▲ **Figure 12.9**
Interconversion of β-D-fructose 6-phosphate and β-D-fructose 2,6-*bis*phosphate.

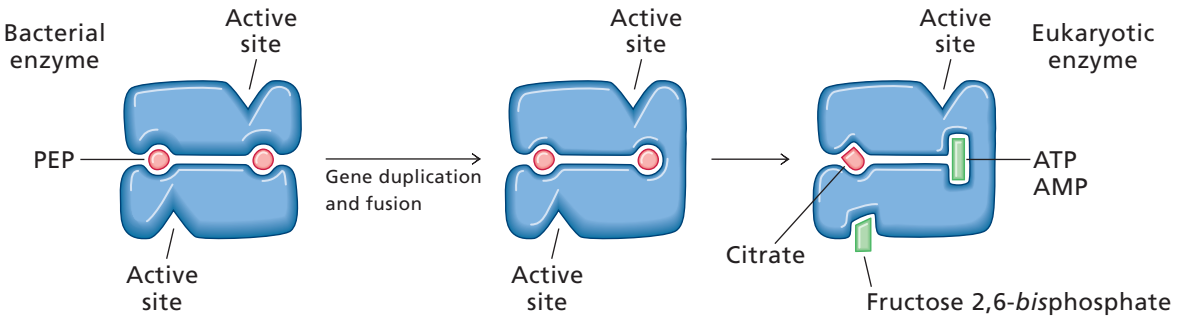
BOX 12.3 THE EVOLUTION OF A COMPLEX ENZYME

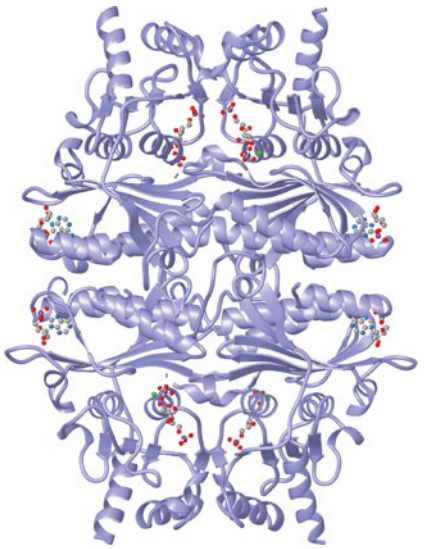
Bacterial versions of phosphofructokinase-1 are homotetramers (Figure 5.19). The functional unit is a head-to-tail dimer with two active sites and two regulatory sites in the interface between the monomers. Phosphoenolpyruvate (PEP) inhibits the enzyme.

In eukaryotes, a tandem gene duplication occurred in the fungi/animal lineage. This was followed by a fusion of the two genes leading to a monomer that was twice the size of the bacterial version. This larger monomer resembled the bacterial dimer with two active sites and two regulatory sites. Over a period of millions of years these sites became modified. One of the active sites continued to bind fructose 6-phosphate and ATP catalyzing the formation of fructose 1,6-*bis*phosphate. In the reverse reaction it binds fructose 1,6-*bis*phosphate. The other active site evolved to bind fructose 2,6-*bis*phosphate, which became an allosteric activator.

The two original regulatory sites also evolved to accommodate new ligands. Citrate became the new inhibitor at one of the sites and the other site became the allosteric site for regulation by ATP (inhibitor) or AMP (activator).

▼ **Evolution of the fungal and animal versions of phosphofructokinase-1.**





▲ **T conformation (inactive) of fructose 1,6-*bis*phosphatase.** This is the tetrameric enzyme from human (Homo sapiens) bound to the allosteric inhibitor AMP (space-filling) at the regulatory sites between the two dimers. The competitive inhibitor fructose 2,6-*bis*phosphate (ball-and-stick) is bound at the active sites of each monomer. [PDB 1EYJ]

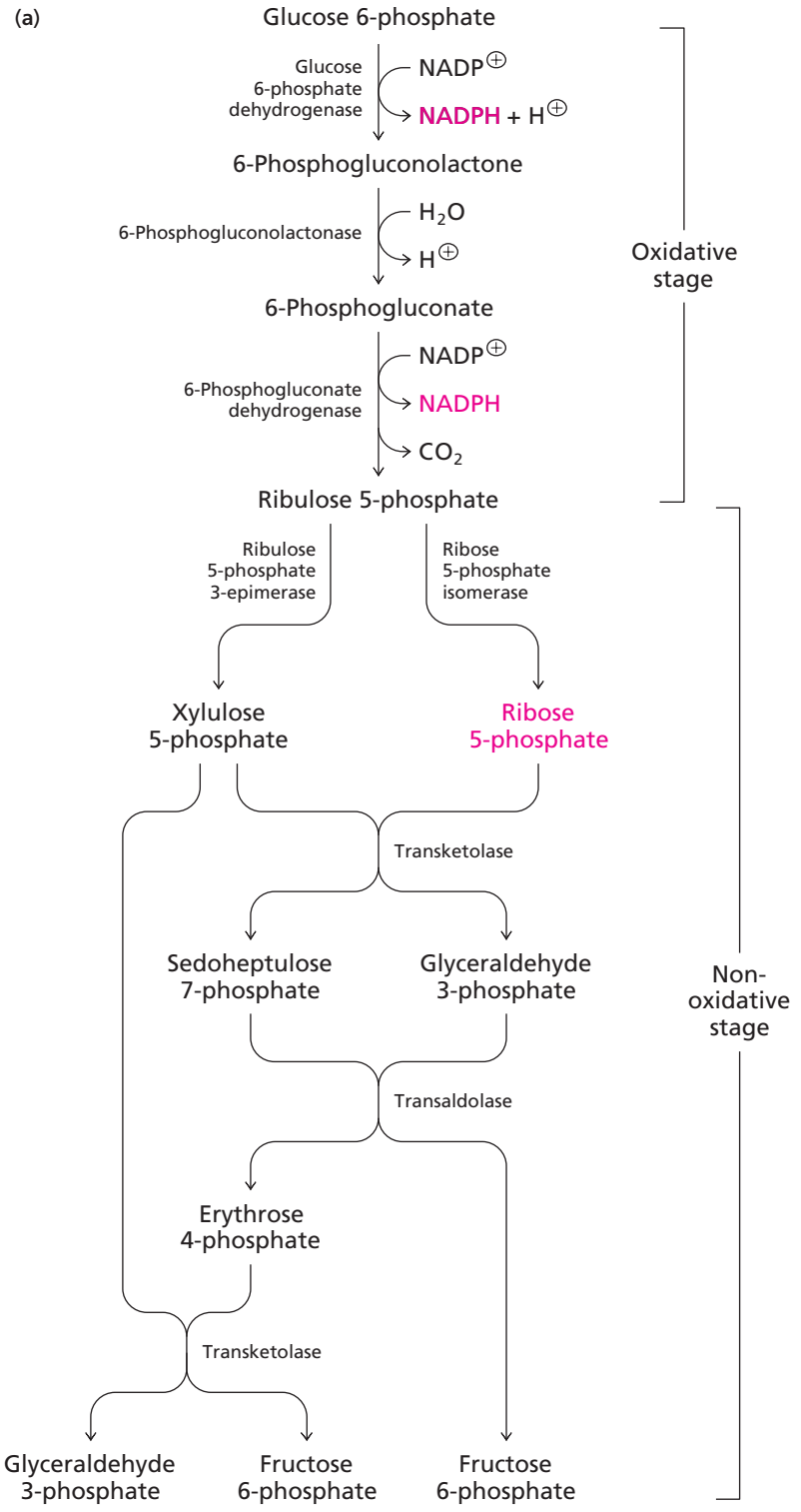
active site on the same protein catalyzes the hydrolytic dephosphorylation of fructose 2,6-*bis*phosphate, re-forming fructose 6-phosphate. This activity of the enzyme is called fructose 2,6-*bis*phosphatase. The dual activities of this bifunctional enzyme control the steady state concentration of fructose 2,6-*bis*phosphate and, ultimately, the switch between glycolysis and gluconeogenesis.

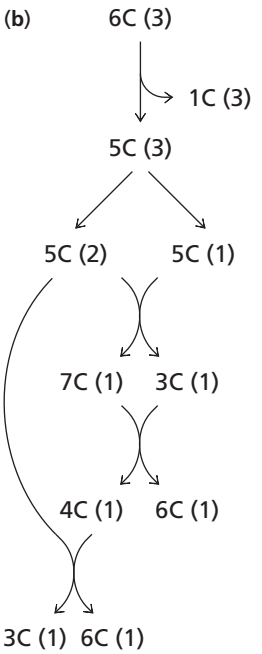
As shown in Figure 12.8, the allosteric effector fructose 2,6-*bis*phosphate activates PFK-1 and inhibits fructose 1,6-*bis*phosphatase. Note that an increase in fructose 2,6-*bis*phosphate has reciprocal effects: it stimulates glycolysis and inhibits gluconeogenesis. Similarly, AMP affects the two enzymes in a reciprocal manner; inhibiting fructose 1,6-*bis*phosphatase and activating phosphofructokinase-1. The regulation of the bifunctional enzyme PFK-2/fructose 2,6-*bis*phosphatase will be described after we cover glycogen metabolism.

12.4 The Pentose Phosphate Pathway

The pentose phosphate pathway is a pathway for the synthesis of three pentose phosphates: ribulose 5-phosphate, ribose 5-phosphate, and xylulose 5-phosphate. Ribose 5-phosphate is required for the synthesis of RNA and DNA. The complete pathway has two stages: an oxidative stage and a nonoxidative stage (Figure 12.10). In the oxidative stage, NADPH is produced when glucose 6-phosphate is converted to the five-carbon compound ribulose 5-phosphate.

$$\text{Glucose 6-phosphate} + 2\ \text{NADP}^{\oplus} + \text{H}_2\text{O} \longrightarrow$$
$$\text{Ribulose 5-phosphate} + 2\ \text{NADPH} + \text{CO}_2 + 2\ \text{H}^{\oplus} \qquad (12.4)$$

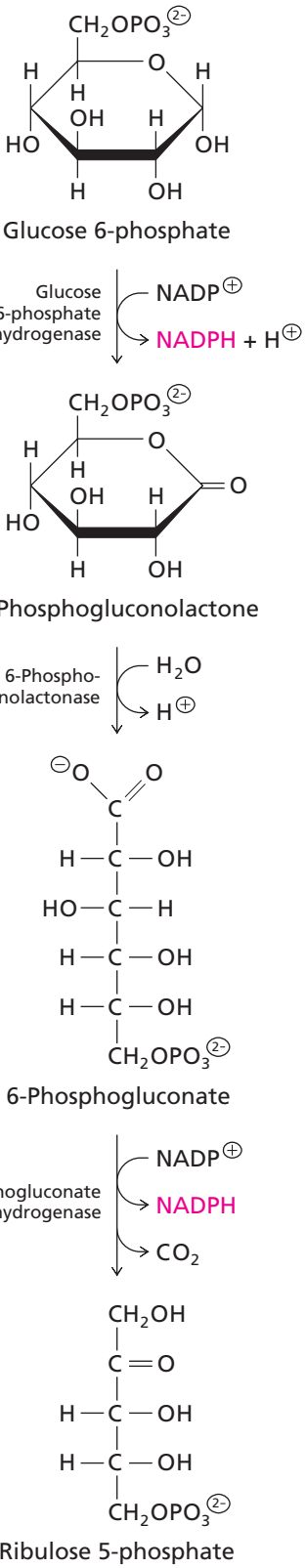




▲ **Figure 12.10**
Pentose phosphate pathway. (a) The oxidative stage of the pathway produces a five-carbon sugar phosphate, ribulose 5-phosphate, with concomitant production of NADPH. The nonoxidative stage produces the glycolytic intermediates glyceraldehyde 3-phosphate and fructose 6-phosphate. **(b)** The path of carbon in the pentose phosphate pathway. In the oxidative stage, three molecules of a six-carbon compound are converted to three molecules of a five-carbon sugar (ribulose 5-phosphate) with release of three molecules of CO_2. In the nonoxidative stage, three molecules of five-carbon sugars are interconverted to produce two molecules of a six-carbon sugar (fructose 6-phosphate) and one molecule of a three-carbon compound (glyceraldehyde 3-phosphate).

If a cell requires both NADPH and nucleotides then all the ribulose 5-phosphate is isomerized to ribose 5-phosphate and the pathway is completed at this stage. In some cases, more NADPH than ribose 5-phosphate is needed and most of the pentose phosphates are converted to intermediates in the gluconeogenesis pathway.

The nonoxidative stage of the pentose phosphate pathway disposes of the pentose phosphate formed in the oxidative stage by providing a route to gluconeogenesis



▲ **Figure 12.11**
Oxidative stage of the pentose phosphate pathway. Two molecules of $NADP^{\oplus}$ are reduced to two molecules of NADPH for each molecule of glucose 6-phosphate that enters the pathway.

or glycolysis. In this stage, ribulose 5-phosphate is converted to the intermediates fructose 6-phosphate and glyceraldehyde 3-phosphate. If all the pentose phosphate were converted to these intermediates, the sum of the nonoxidative reactions would be the conversion of three pentose molecules to two hexose molecules plus one triose molecule.

$$3 \text{ Ribulose 5-phosphate} \longrightarrow 2 \text{ Fructose 6-phosphate} + \text{Glyceraldehyde 3-phosphate} \quad \textbf{(12.5)}$$

Both fructose 6-phosphate and glyceraldehyde 3-phosphate can be metabolized by glycolysis or gluconeogenesis.

Let's take a closer look at the individual reactions of the pentose phosphate pathway.

A. Oxidative Stage

The three reactions of the oxidative stage of the pentose phosphate pathway are shown in Figure 12.11. The first two steps are the same as those in the bacterial Entner–Doudoroff pathway (Section 11.7). The first reaction, catalyzed by glucose 6-phosphate dehydrogenase (G6PDH), is the oxidation of glucose 6-phosphate to 6-phosphogluconolactone. This step is the major regulatory site for the entire pentose phosphate pathway. Glucose 6-phosphate dehydrogenase is allosterically inhibited by NADPH (feedback inhibition). This simple regulatory feature ensures that the production of NADPH by the pentose phosphate pathway is self-limiting.

The next enzyme of the oxidative phase is 6-phosphogluconolactonase that catalyzes the hydrolysis of 6-phosphogluconolactone to the sugar acid 6-phosphogluconate. Finally, 6-phosphogluconate dehydrogenase catalyzes the oxidative decarboxylation of 6-phosphogluconate. This reaction produces a second molecule of NADPH, ribulose 5-phosphate, and CO_2. In the oxidative stage, therefore, a six-carbon sugar is oxidized to a five-carbon sugar plus CO_2 and two molecules of $NADP^{\oplus}$ are reduced to two molecules of NADPH.

B. Nonoxidative Stage

The nonoxidative stage of the pentose phosphate pathway consists entirely of near equilibrium reactions. This stage of the pathway provides five-carbon sugars for biosynthesis and introduces sugar phosphates into glycolysis or gluconeogenesis. Ribulose 5-phosphate has two fates: an epimerase can catalyze the formation of xylulose 5-phosphate, or an isomerase can catalyze the formation of ribose 5-phosphate (Figure 12.12). (Note the difference between an epimerase and an isomerase.) Ribose 5-phosphate is the precursor of the ribose (or deoxyribose) portion of nucleotides. The remaining steps of the pathway convert the five-carbon sugars into glycolytic intermediates. Rapidly dividing cells that require both ribose 5-phosphate (as a precursor of ribonucleotide and deoxyribonucleotide residues) and NADPH (for the reduction of ribonucleotides to deoxyribonucleotides) generally have high pentose phosphate pathway activity.

The overall pentose phosphate pathway (Figure 12.10) shows that in the nonoxidative stage two molecules of xylulose 5-phosphate and one molecule of ribose 5-phosphate are interconverted to generate one three-carbon molecule (glyceraldehyde 3-phosphate) and two six-carbon molecules (fructose 6-phosphate). Thus, the carbon-containing products from the passage of three molecules of glucose through the pentose phosphate pathway are glyceraldehyde 3-phosphate, fructose 6-phosphate, and CO_2. The balanced equation for this process is

$$3 \text{ Glucose 6-phosphate} + 6\ NADP^{\oplus} + 3\ H_2O \longrightarrow 2 \text{ Fructose 6-phosphate} + \text{Glyceraldehyde 3-phosphate} + 6 \text{ NADPH} + 3\ CO_2 + 6\ H^{\oplus} \quad \textbf{(12.6)}$$

BOX 12.4 GLUCOSE 6-PHOSPHATE DEHYDROGENASE DEFICIENCY IN HUMANS

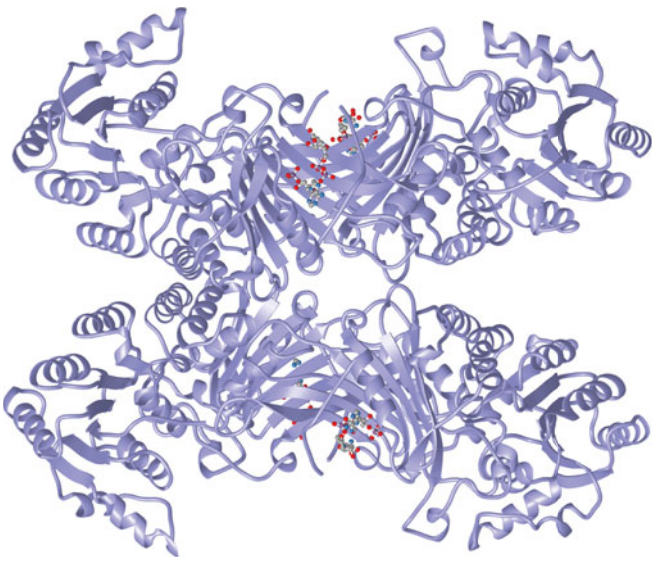
The genetics of human glucose 6-phosphate dehydrogenase has been the subject of much research. There are two different enzymes that can catalyze the reaction shown in Figure 12.11. One of the genes (G6PDH) is found on the X chromosome (Xq28) and it is expressed almost exclusively in red blood cells. The other gene (H6PDH) encodes an enzyme that is less specific; it can use other hexose substrates. Hexose 6-phosphate dehydrogenase is synthesized in many cells where it serves as the first enzyme in the oxidative stage of the pentose phosphate pathway.

The glucose 6-phosphate dehydrogenase reaction is the only reaction capable of reducing $NADP^{\oplus}$ in red blood cells; consequently, deficiencies of this enzyme have drastic effects on the metabolism of these cells. Other cells are not affected since they contain H6PDH. G6PDH deficiency in humans causes hemolytic anemia.

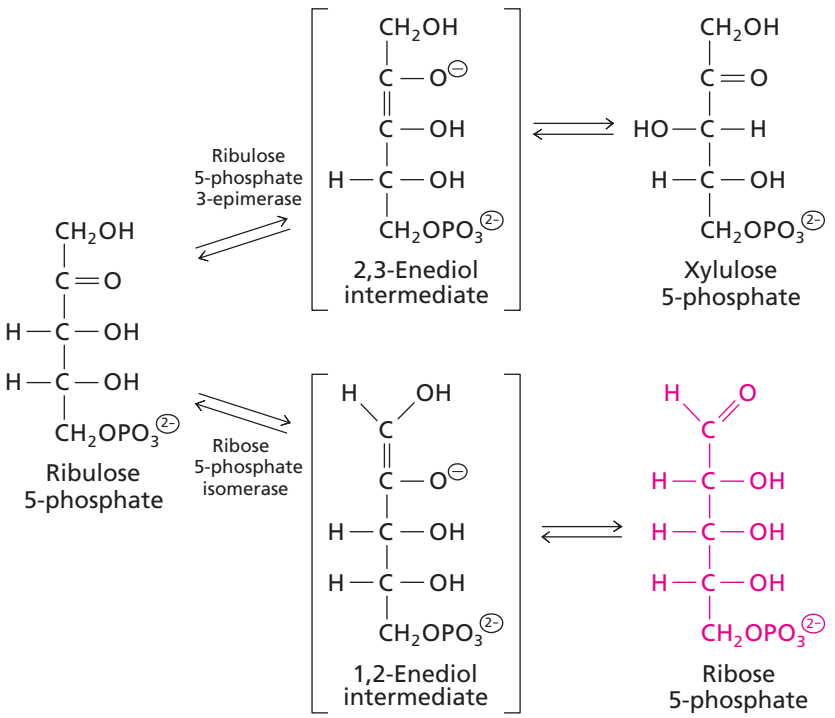
There are hundreds of different alleles of the X chromosome G6PDH gene. The variants produce lower amounts of the enzyme or they alter its catalytic efficiency. There are no known null mutants in the human population because the complete absence of G6PDH activity is lethal. Note that males are more likely to be affected since they have only a single copy of the gene on their one X chromosome.

It is estimated that 400 million people have some form of G6PDH deficiency and suffer from mild forms of hemolytic anemia. The symptoms can be life threatening if the patient is treated with certain drugs that are normally prescribed for other diseases. Many of these individuals have an increased resistance to malaria because the malarial parasite does not survive well in red blood cells that produce lowered amounts of NADPH. This explains why there are so many deficiency alleles segregating in the human population in spite of the fact that the pentose phosphate pathway is inefficient. It's an example of balanced selection like the familiar sickle cell anemia example.

Human genome database entries for these genes can be viewed on the Entrez Gene website [ncbi.nlm.nih.gov/gene]. Type in the entries for the G6PDH gene (2531) or the H6PDH gene (9563). The Online Mendelian Inheritance in Man (OMIM) webpage is at ncbi.nlm.nih.gov/omim. The entry for G6PDH is MIM=305900 and the entry for H6PDH is MIM=138090.

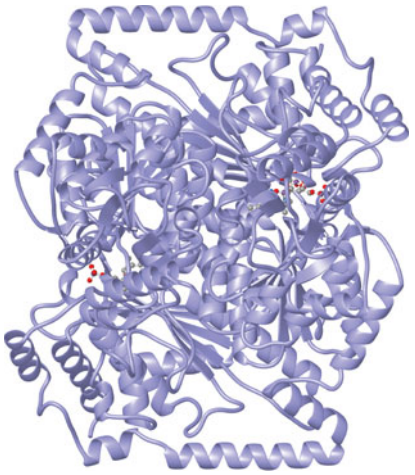


▲ **Human glucose 6-phosphate dehydrogenase, variant Canton R459L.** The enzyme is a dimer of dimers (tetramer). Two molecules of $NADP^{\oplus}$ are bound at the active sites in each dimer. [PDB 1QK1]



The reactions of the nonoxidative stage of the pentose phosphate pathway are similar to those of the regeneration stage of the reductive pentose phosphate cycle of photosynthesis (Section 15.8).

◀ **Figure 12.12**
Conversion of ribulose 5-phosphate to xylulose 5-phosphate or ribose 5-phosphate. In either case, the removal of a proton forms an enediol intermediate. Reprotonation forms either the ketose xylulose 5-phosphate or the aldose ribose 5-phosphate.



▲ **Transketolase from *Escherichia coli*.** The active site of each monomer contains one molecule of xylulose 5-phosphate (space-filling) and the TDP cofactor. [PDN 2R8O]

In most cells, the glyceraldehyde 3-phosphate and fructose 6-phosphate produced by the pentose phosphate pathway are used to resynthesize glucose 6-phosphate. This glucose 6-phosphate molecule can reenter the pentose phosphate pathway. In that case, the equivalent of one molecule of glucose is completely oxidized to CO_2 by six passages through the pathway. After six molecules of glucose 6-phosphate are oxidized, the six ribulose 5-phosphates produced can be rearranged by the reactions of the pentose phosphate pathway and part of the gluconeogenic pathway to form five glucose 6-phosphate molecules. (Recall that two glyceraldehyde 3-phosphate molecules are equivalent to one fructose 1,6-*bis*phosphate molecule.) If we disregard H_2O and $H^{\oplus}$, the overall stoichiometry for this process is

$$6 \text{ Glucose 6-phosphate} + 12 \text{ NADP}^{\oplus} \longrightarrow$$
$$5 \text{ Glucose 6-phosphate} + 12 \text{ NADPH} + 6 \text{ CO}_2 + \text{P}_i \qquad \textbf{(12.7)}$$

This net reaction emphasizes that most of the glucose 6-phosphate entering the pentose phosphate pathway could be recycled; one-sixth is converted to CO_2 and P_i. Indeed, an alternate name for the pathway is the *pentose phosphate cycle*.

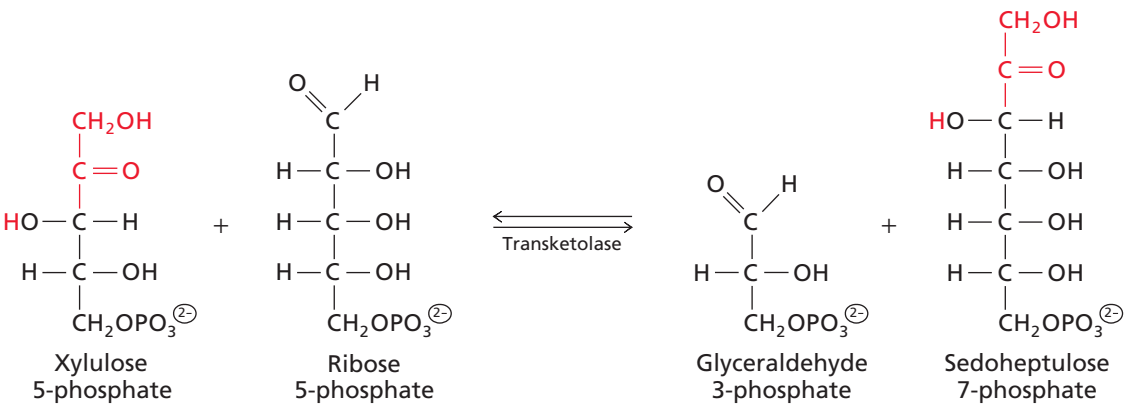
C. Interconversions Catalyzed by Transketolase and Transaldolase

The interconversions of the nonoxidative stage of the pentose phosphate pathway are catalyzed by two enymes called transketolase and transaldolase. These enzymes have broad substrate specificities.

Transketolase is also called glycoaldehydetransferase. It is a thiamine diphosphate (TDP)-dependent enzyme that catalyzes the transfer of a two-carbon glycoaldehyde group from a ketose phosphate to an aldose phosphate. The ketose phosphate is shortened by two carbons and the aldose phosphate is lengthened by two carbons (Figure 12.13).

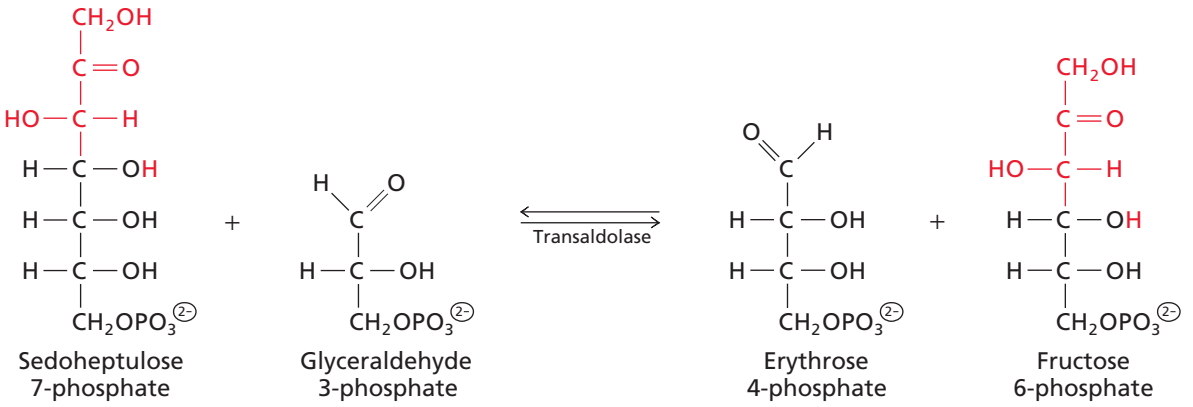
Transaldolase is also called dihydroxyacetonetransferase. It catalyzes the transfer of a three-carbon dihydroxyacetone group from a ketose phosphate to an aldose phosphate. The transaldolase reaction of the pentose phosphate pathway converts sedoheptulose 7-phosphate and glyceraldehyde 3-phosphate to erythrose 4-phosphate and fructose 6-phosphate (Figure 12.14).

Figure 12.13 ▼
Reaction catalyzed by transketolase. The reversible transfer of a glycoaldehyde group (shown in red) from xylulose 5-phosphate to ribose 5-phosphate generates glyceraldehyde 3-phosphate and sedoheptulose 7-phosphate. Note that the ketose–phosphate substrate (in either direction) is shortened by two carbon atoms, whereas the aldose–phosphate substrate is lengthened by two carbon atoms. In this example, 5C + 5C ⟹ 3C + 7C.



12.5 Glycogen Metabolism

Glucose is stored as the intracellular polysaccharides starch and glycogen. In Chapter 15 we discuss starch metabolism, which occurs mostly in plants. Glycogen is an important storage polysaccharide in bacteria, protists, fungi and animals. Large glycogen particles can be easily seen in the cytoplasm of these organisms. Most of the glycogen in vertebrates is found in muscle and liver cells. Muscle glycogen appears in electron micrographs as cytosolic granules with a diameter of 10 to 40 nm, about the size of ribosomes. Glycogen particles in liver cells are about three times larger. The glycogen particles in bacteria are smaller.



▲ **Figure 12.14**
Reaction catalyzed by transaldolase. The reversible transfer of a three-carbon dihydroxyacetone group (shown in red) from sedoheptulose 7-phosphate, to C-1 of glyceraldehyde 3-phosphate generates a new ketose phosphate, fructose 6-phosphate, and releases a new aldose phosphate, erythrose 4-phosphate. Note that the carbon atoms balance: 7C + 3C ⟺ 6C + 4C.

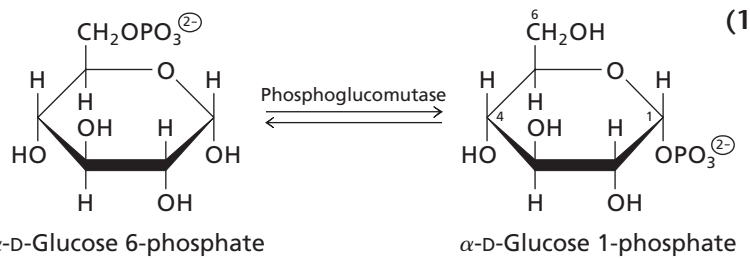
A. Glycogen Synthesis

De novo glycogen synthesis requires a preexisting primer of four to eight α-(1 → 4)-linked glucose residues. This primer is attached to a specific tyrosine residue of a protein called glycogenin (Figure 12.15) via the 1-hydroxyl group of the reducing end of the short polysaccharide. The primer is formed in two steps. The first glucose residue is attached to glycogenin by the action of a glucosyltransferase activity that requires UDP-glucose. Glycogenin itself catalyzes this reaction as well as the extension of the primer by up to seven more glucose residues. Thus, glycogenin is both a protein scaffold for glycogen and an enzyme. Each glycogen molecule (which can contain thousands of glucose residues) contains a single glycogenin protein at its center.

Further glycogen addition reactions begin with glucose 6-phosphate that can be converted to glucose 1-phosphate. We saw in Section 11.5 that glucose 6-phosphate can enter a number of pathways, including glycolysis and the pentose phosphate pathway. Glycogen synthesis and degradation is mostly a way of storing glucose 6-phosphate until it is needed by the cell. The synthesis and degradation of glycogen require separate enzymatic steps. We have already noted that it is a general rule of metabolism that biosynthesis pathways and degradation pathways follow different routes.

Three separate enzyme-catalyzed reactions are required to incorporate a molecule of glucose 6-phosphate into glycogen (Figure 12.16). First, phosphoglucomutase catalyzes the conversion of glucose 6-phosphate to glucose 1-phosphate. Glucose 1-phosphate is then activated by reaction with UTP, forming UDP-glucose and pyrophosphate (PP_i). In the third step, glycogen synthase catalyzes the addition of glucose residues from UDP-glucose to the nonreducing end of glycogen.

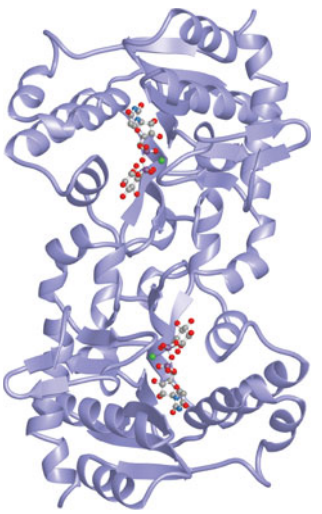
Phosphoglucomutase is a ubiquitous enzyme. It catalyzes a near-equilibrium reaction that converts α-D-glucose 6-phosphate to α-D-glucose 1-phosphate Glucose 1-phosphate is the famous "Cori ester" discovered by Gerty Cori and Carl Cori in the 1930s when the reactions of glycogen metabolism were first being elucidated.



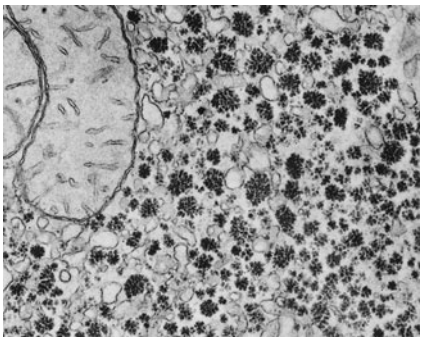
(12.12)



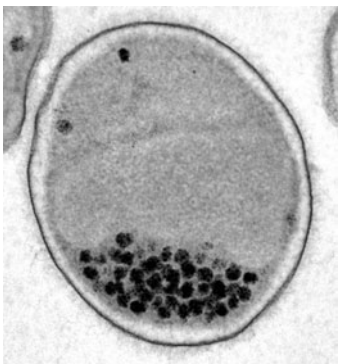
▲ **Gerty Cori, (1896–1957) biochemist.** Carl Cori and Gerty Cori won the Nobel Prize in 1947 "for their discovery of the course of the catalytic conversion of glycogen." This stamp depicts the "Cori ester" but it's slightly different than the structure we usually see in textbooks. Can you spot the difference?



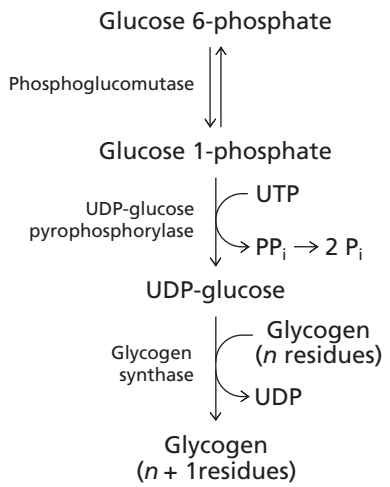
▲ **Figure 12.15**
Glycogenin from rabbit (*Oryctolagus cuniculus*). The molecule is a homodimer and each of the active sites contains a bound molecule of UDP-glucose. [PDB 1LL2]



▲ **Large glycogen particles in a section of a liver cell.** (Electron micrograph.)



▲ **Stained glycogen granules in bacteria (*Candidatus* spp.)**



▲ **Figure 12.16**
Synthesis of glycogen in eukaryotes.

The mechanism of this reaction is similar to that of cofactor-dependent phosphoglycerate mutase (Section 11.2 8). Glucose 6-phosphate binds to the phosphoenzyme, and glucose 1,6-*bis*phosphate is formed as an enzyme-bound intermediate. Transfer of the C-6 phosphate to the enzyme leaves glucose 1-phosphate.

Glucose 1-phosphate is activated by formation of UDP-glucose in the second step of glycogen synthesis. In this reaction a UMP group from UTP is transferred to the phosphate at C-1 with release of pyrophosphate (see Figure 7.6). The enzyme that catalyzes this reaction is called UDP glucose pyrophosphorylase and it is present in most eukaryotic species. Note that the activation of glucose requires UTP. The energy is stored in UDP-glucose where it can be used in many biosynthesis reactions. We saw in Section 11.6 that UDP-glucose can be a substrate for synthesis of UDP-galactose. (UDP-galactose is used in the synthesis of gangliosides.) The standard Gibbs free energy change in the UDP glucose pyrophospholylase reaction is close to zero. Under the steady state, near-equilibrium conditions found *in vivo*, $\Delta G = 0$ and the concentrations of glucose 1-phosphate and UDP-glucose are nearly equal. Flux in the direction of UDP-glucose synthesis is driven by subsequent hydrolysis of pyrophosphate (Section 10.6). Two ATP equivalents (UTP and PP_i) are used in the activation of glucose.

Glycogen synthesis is a polymerization reaction where glucose units are added one at a time to a growing polysaccharide chain. This reaction is catalyzed by glucogen synthase (Figure 12.17). Many polymerization reactions are **processive**—the enzyme remains bound to the end of the growing chain and addition reactions are very rapid (see Section 20.2B). The glycogen synthase reaction is **distributive**—the enzyme releases the growing glycogen chain after each reaction.

Glycogen synthases that use UDP-glucose as their substrate are present in protists, animals, and fungi. Some bacteria synthesize glycogen using ADP-glucose. Starch synthesis in plants also requires ADP-glucose. The glycogen synthase reaction is the major regulatory step of glycogen synthesis. In animals, there are hormones that control the rate of glycogen synthesis by altering the activity of glycogen synthase. We will describe regulation in the next section.

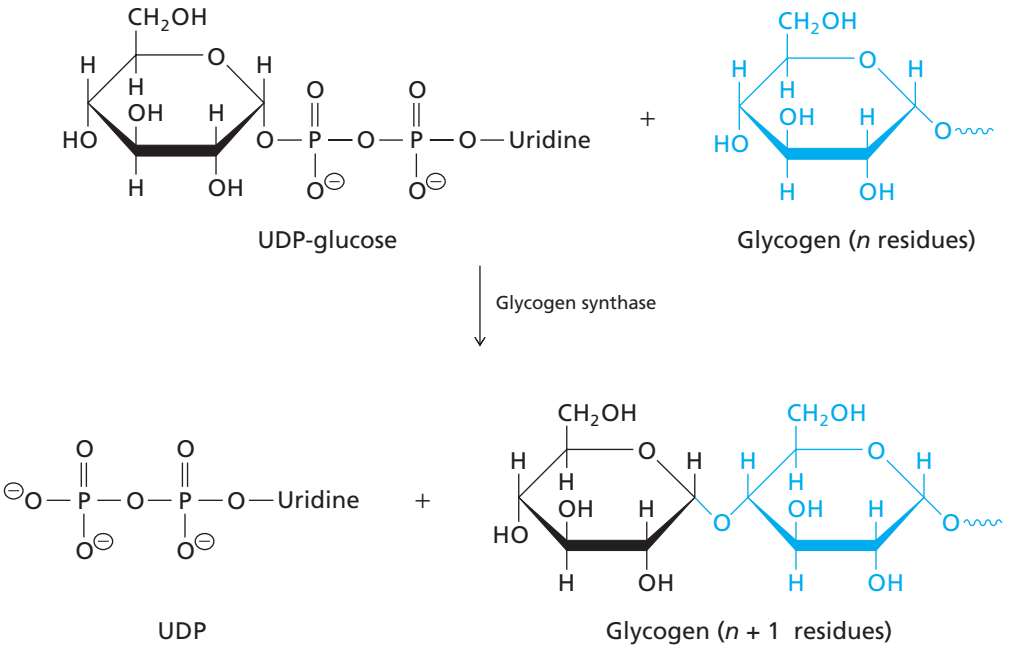
Another enzyme, amylo-(1,4 → 1,6)-transglycosylase, catalyzes branch formation in glycogen. This enzyme, also known as the branching enzyme, removes an oligosaccharide of at least six residues from the nonreducing end of an elongated chain and attaches it by an α-(1 → 6) linkage to a position at least four glucose residues from the nearest α-(1 → 6) branch point. These branches provide many sites for adding or removing glucose residues, thereby contributing to the speed with which glycogen can be synthesized or degraded.

The complete glycogen molecule has many layers of polysaccharide chains extending out from the glycogenin core (Figure 12.18). The large granules in liver cells, for example, have glycogen molecules with up to 120,000 glucose residues. There are usually two branches per chain and each chain is 8–14 residues in length. The molecule has about 12 layers of chains. If there were on average two branches per chain then each polysacharide unit would have thousands of free ends.

B. Glycogen Degradation

The glucose residues of starch and glycogen are released from storage polymers through the action of enzymes called polysaccharide phosphorylases: starch phosphorylase (in plants) and glycogen phosphorylase (in other organisms). These enzymes catalyze the removal of glucose residues from the nonreducing ends of starch or glycogen, provided the monomers are attached by α-(1 → 4) linkages. As the name implies, the enzymes catalyze phosphorolysis—cleavage of a bond by group transfer to an oxygen atom of phosphate. In contrast to hydrolysis (group transfer to water), phosphorolysis produces phosphate esters. Thus, the first product of polysaccharide breakdown is α-D-glucose 1-phosphate (the Cori ester), not free glucose.

$$\text{Polysaccharide}_{(n\ \text{residues})} + P_i \xrightarrow{\text{Polysaccharide phosphorylase}} \text{Polysaccharide}_{(n\text{-}1\ \text{residues})} + \text{Glucose 1-phosphate} \quad \textbf{(12.9)}$$



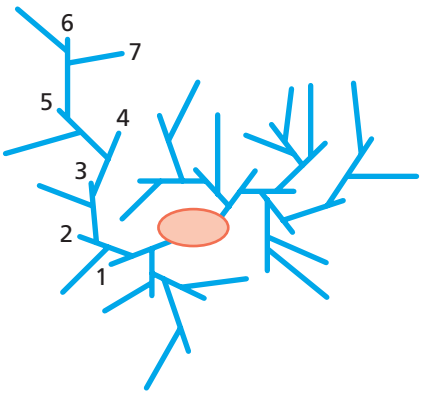
▲ **Figure 12.17**
The glycogen synthase reaction.

The phosphorolysis reaction catalyzed by glycogen phosphorylase is shown in Figure 12.19. Pyridoxal phosphate (PLP) is a prosthetic group in the active site of the enzyme. The phosphate group of PLP appears to relay a proton to the substrate phosphate to help cleave the scissile C—O bond of glycogen. Note that glycogen phosphorylase catalyzes a remarkable reaction since it only uses glycogen and inorganic phosphates as substrates in a reaction that produces a relatively "high energy" compound, glucose 1-phosphate (Table 10.1).

Glycogen phosphorylase is a dimer of identical subunits. The catalytic sites lie in the middle of each subunit. It binds phosphate and the end of a glycogen chain (Figure 12.20). The large glycogen particle binds to a nearby site and the chain being degraded passes along a groove on the surface of the enzyme. Four or five glucose residues can be cleaved sequentially before the enzyme has to release a glycogen particle and re-bind. Thus, in contrast to glycogen synthase, glycogen phosphorylase is partially processive.

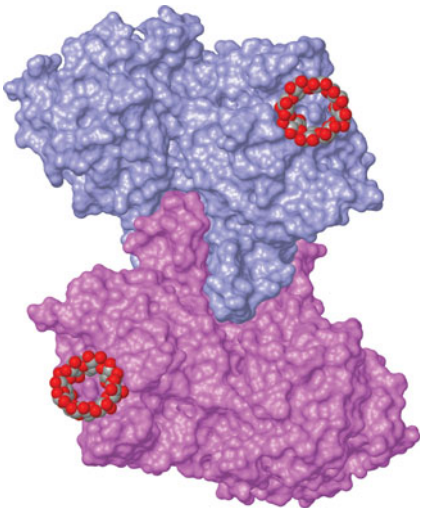
The enzyme stops four glucose residues from a branch point (an α-(1 → 6) glucosidic bond) leaving a limit dextrin. The limit dextrin can be further degraded by the action of the bifunctional glycogen debranching enzyme (Figure 12.21). A glucanotransferase activity of the debranching enzyme catalyzes the relocation of a chain of three glucose residues from a branch to a free 4-hydroxyl end of the glycogen molecule. Both the original linkage and the new one are α-(1 → 4). The other activity of glycogen debranching enzyme, amylo-1,6-glucosidase, catalyzes hydrolytic (not phosphorolytic) removal of the remaining α-(1 → 6)-linked glucose residue. The products are one free glucose molecule and an elongated chain that is again a substrate for glycogen phosphorylase. When a glucose molecule released from glycogen by the action of the debranching enzyme enters glycolysis, two ATP molecules are produced (Section 11.1). In contrast, each glucose molecule mobilized by the action of glycogen phosphorylase (representing about 90% of the residues in glycogen) yields three ATP molecules. The energy yield from glycogen is higher than from free glucose because glycogen phosphorylase catalyzes phosphorolysis rather than hydrolysis—no ATP is consumed as in the hexokinase-catalyzed phosphorylation of free glucose.

The product of glycogen degradation, glucose 1-phosphate, is rapidly converted to glucose 6-phosphate by phosphoglucomutase.

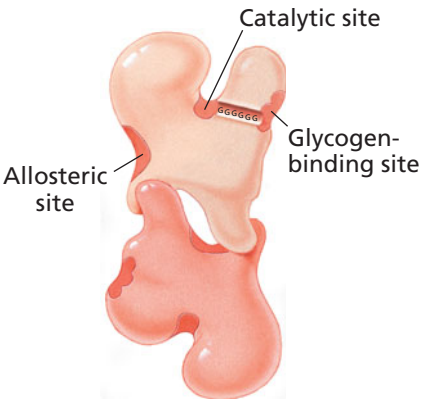


▲ **Figure 12.18**
A glycogen molecule. Two polysaccharides (blue) are attached to each core glycogenin molecule. Each chain core has 8–14 residues and two branches. Not all branches are shown. Seven layers are numbered but typical glycogen molecules have 8–12 layers, depending on the species.

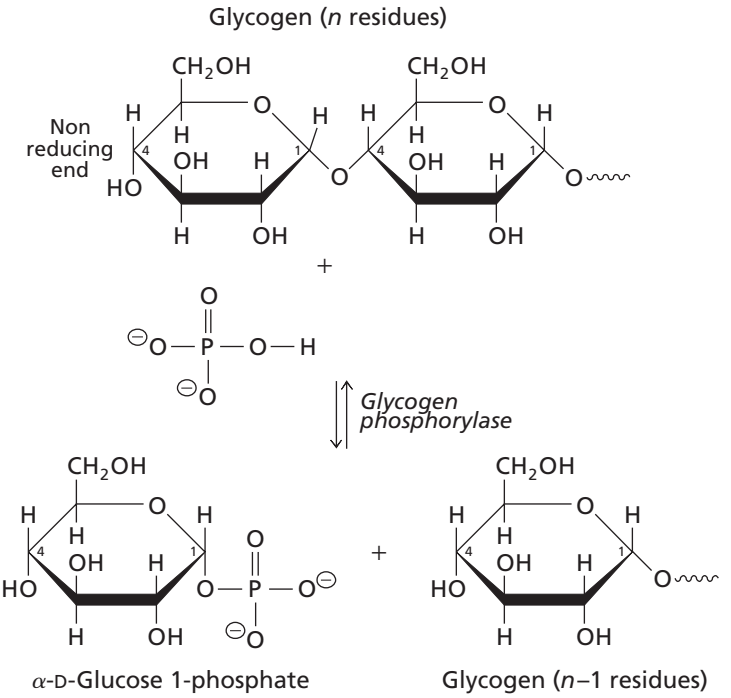
There's no magical net gain of energy by storing glucose as glycogen since the cost of incorporating glucose 6-phosphate into glycogen is two ATP equivalents (Figure 12.16).



▲ **Inhibiting glycogen phosphorylase.** The action of glycogen phosphorylase produces glucose in the liver. Insulin controls this activity by inactivating glycogen phosphorylase but in the absence of insulin (e.g., Type II diabetes), excess production of glucose can be dangerous. Many inhibitors of glycogen phosphorylase have been developed as possible treatments for diabetes. One of them is a cyclic maltose molecule shown here bound to the active sites of the rabbit (*Oryctolagus cuniculis*) enzyme. [PDB1P2G]



▲ **Figure 12.20**
Binding and catalytic sites on glycogen phosphorylase.



▲ **Figure 12.19**
Cleavage of a glucose residue from the nonreducing end of a glycogen chain, catalyzed by glycogen phosphorylase.

12.6 Regulation of Glycogen Metabolism in Mammals

Mammalian glycogen stores glucose in times of plenty (after feeding, a time of high glucose levels) and supplies glucose in times of need (during fasting or in "fight or flight" situations). In muscle, glycogen provides fuel for muscle contraction. In contrast, liver glycogen is largely converted to glucose that exits liver cells and enters the bloodstream for transport to other tissues that require it. Both the mobilization and synthesis of glycogen are regulated by hormones.

A. Regulation of Glycogen Phosphorylase

Glycogen phosphoryase is responsible for the breakdown of glycogen to produce glucose 1-phosphate. In muscle cells, glucose 1-phosphate is converted to glucose 6-phosphate that is used in glycolysis to produce ATP. In liver cells, glucose 6-phosphate is hydrolyzed to free glucose that is secreted into the bloodstream where it can be taken up by other tissues.

The activity of glycogen phosphorylase is regulated by several allosteric effectors and by covalent modification (phosphorylation). Let's take a few minutes to study the regulation of glycogen phosphorylase because not only is it important in glycogen metabolism, it's also historically important.

The enzyme exists in four different forms as shown in Figure 12.22. The unphosphorylated form is called glycogen phosphorylase b (GPb) and the phosphorylated form is called glycogen phosphorylase a (GPa). The enzyme is phosphorylated by a kinase enzyme and dephosphorylated by a phosphatase.

Like other allosterically regulated enzymes, glycogen phosphorylase adopts two conformations; the R conformation is the active conformation and the T conformation is much less active. This is depicted in Figure 12.22 as a change in the shape of the catalytic site: In the R conformation, inorganic phosphate (a substrate of the reaction) can bind and in the T conformation binding of inorganic phosphate is inhibited.

Unphosphorylated GPb can exist in both inactive T conformations and active R conformations. The allosteric site of the enzyme binds several effectors that cause a

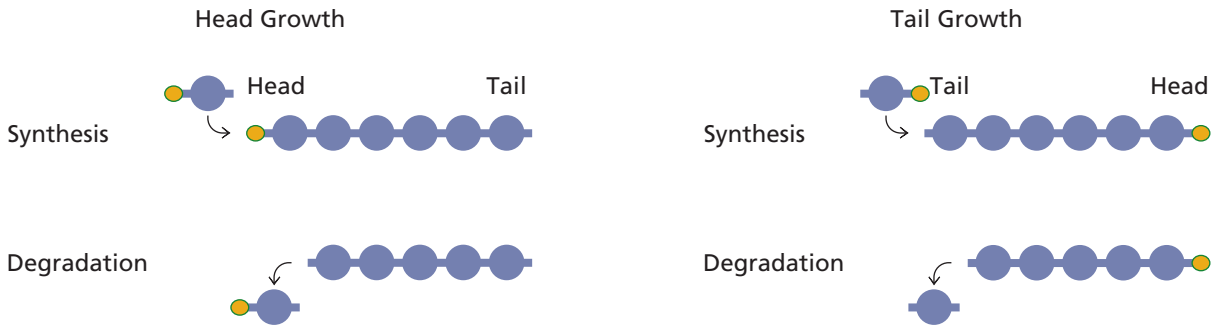
BOX 12.5 HEAD GROWTH AND TAIL GROWTH

Polymerization reactions can be described as either head growth or tail growth. In a head growth mechanism, the growing end of the chain is "activated" and cleavage of the "high energy" linkage at the head of the molecule provides the energy for the next addition of a monomer. In a tail growth mechanism, the growing end does not contain the high energy linkage; instead, the energy for the addition reaction comes from the activated monomer.

Glycogen synthesis is an example of a tail growth mechanism. The incoming monomer (UDP-glucose) is activated and, when the reaction is complete, the end of the glycogen chain is a simple hydroxyl group at the 4-carbon atom of a glucose residue. DNA and RNA synthesis are also examples of a tail growth mechanism. Protein synthesis and fatty acid synthesis are examples of head growth mechanism.

The differences between the two mechanisms become clear when you think of the reverse reaction: degradation. Glycogen and nucleic acids can be degraded by chopping off a single residue. In the case of glycogen, synthesis and degradation are part of an ongoing process since the glycogen particle serves as a storage molecule for glucose. In the case of nucleic acids, especially DNA, the degradation reaction is an essential part of DNA repair and proofreading that ensures DNA replication is extremely accurate (Section 20.2C). Removal of single residues does not prevent the polymer from serving immediately as a substrate for further addition reactions.

Protein synthesis and fatty acid synthesis utilize head growth mechanisms for synthesis. In this case, removal of an end residue also removes the activated head so further addition reactions are not possible without an additional step to "reactivate" the head. This is one reason why protein synthesis errors can't be repaired and one reason why fatty acid chains aren't used as energy storage molecules in the same way that glycogen is used.

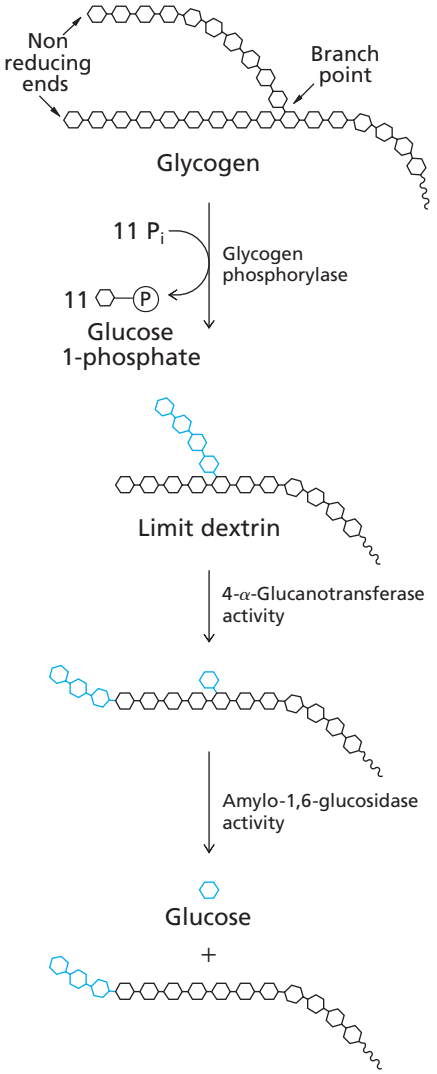


▲ **Head and tail growth.** In a head growth mechanism (left), incoming activated monomers are added to the "head" of the growing polymer. (The end that contains the activated residue.) After the addition reaction, the polymer still contains an activated residue at the growing end. In tail growth (right), the incoming activated monomer is added to the "tail" end of the growing polymer. The monomer substrate carries the energy for its own addition reaction. When the polymer is degraded, a single residue is removed. Polymers that use a head growth mechanism will no longer be a substrate for addition reactions following degradation because the activated head has been removed. Polymers that employ a tail growth mechanism are still able to act as substrates for addition reactions.

shift in conformation. The allosteric site is close to the dimer interface between the two monomers and both subunits change conformation simultaneously—a result that conforms to the *concerted model* of Monod, Wyman, and Changeux (Section 5.9C).

When ATP is bound, the activity of the enzyme is inhibited (T state). This is the normal state of activity since physiological concentrations of ATP are high and relatively constant. When the AMP concentration rises, it displaces ATP from the allosteric site causing a shift to the active R conformation and activation of glycogen breakdown. In muscle cells, increasing AMP concentration results from strenuous muscle activity and signals the need for more glucose 1-phosphate to stimulate ATP production by glycolysis. The enzyme is inhibited by glucose 6-phosphate (feedback inhibition). There's no need to continue glycogen breakdown if glucose 6-phosphate concentration is sufficient to fuel glycolysis.

The main difference between the R conformation and the T conformation is the position of a loop containing Asp-283 and nearby residues (the 280s loop). In the T conformation, the negatively charged side chain of Asp-283 lies close to the pyridoxal 5-phosphate (PLP) cofactor at the catalytic site. This proximity prevents binding of inorganic phosphate, inhibiting the reaction. In the R conformation, the position of this loop shifts allowing inorganic phosphate to enter the active site.



▲ **Figure 12.21**
Degradation of glycogen. Glycogen phosphorylase catalyzes the phosphorolysis of glycogen chains, stopping four residues from an α-(1 → 6) branch point and producing one molecule of glucose 1-phosphate for each glucose residue mobilized. Further degradation is accomplished by the two activities of the glycogen debranching enzyme. The 4-α-glucanotransferase activity catalyzes the transfer of a trimer from a branch of the limit dextrin to a free end of the glycogen molecule. The amylo-1,6-glucosidase activity catalyzes hydrolytic release of the remaining α-(1 → 6)-linked glucose residue.

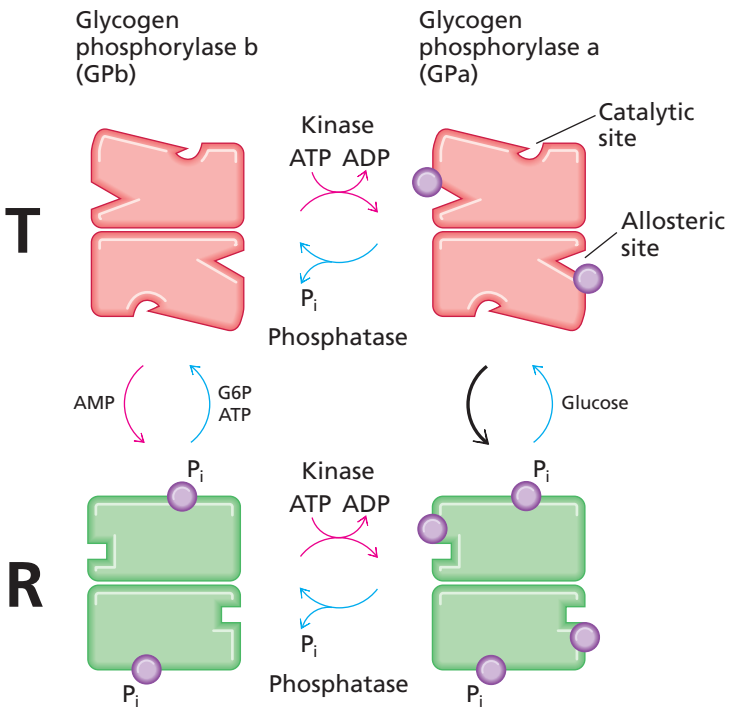
Phosphofructokinase-1 (PFK-1) is regulated in a similar manner by ATP and AMP.

The structures of GPa and GPb are shown in Figure 12.23 in order to illustrate the structural changes that take place when the enzyme is phosphorylated and dephosphorylated. The phosphoryl group is covalently attached to serine residue 14 (Ser-14) near the N-terminal end of the protein.

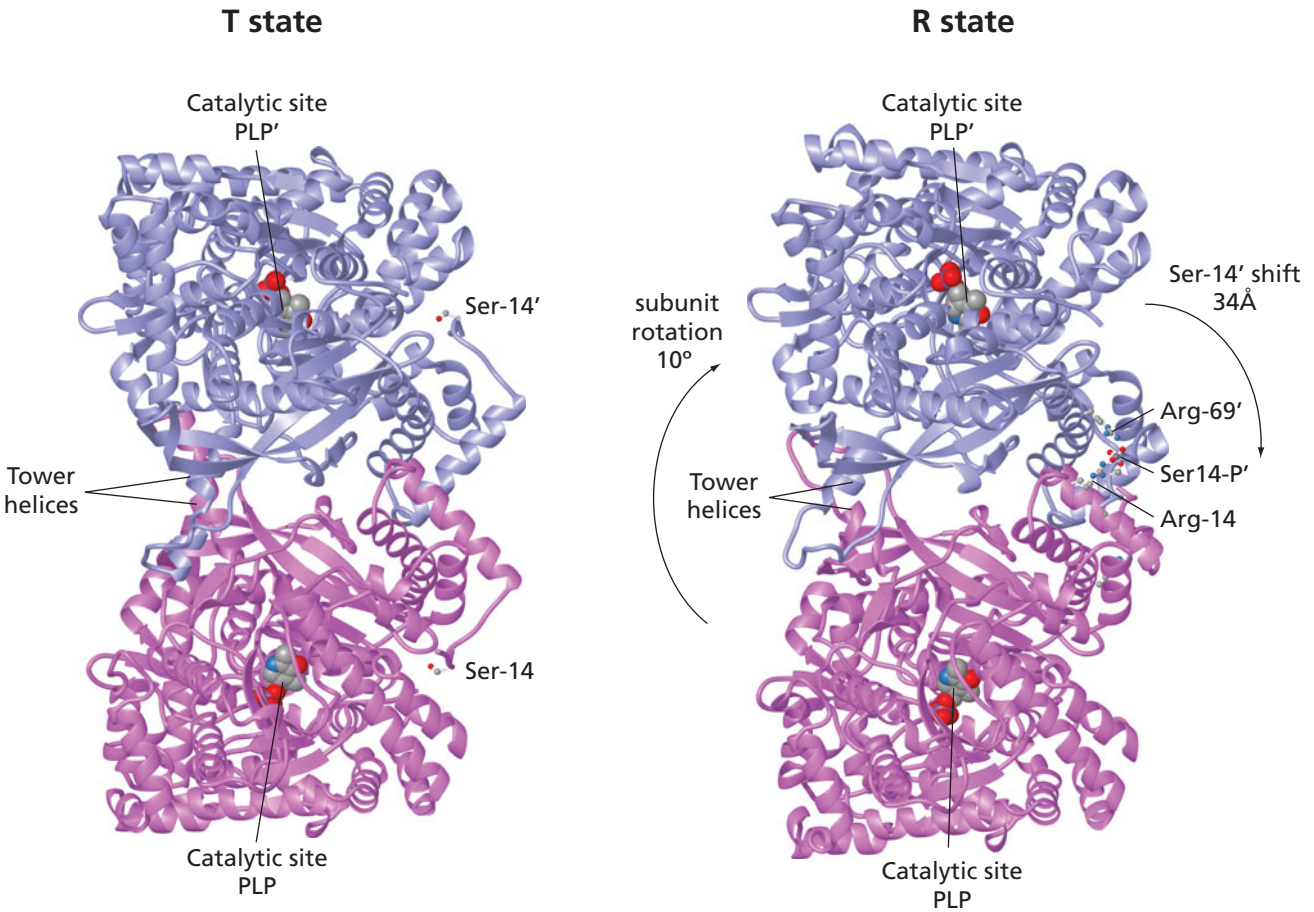
In the unphosphorylated state (GPb), the N-terminal residues, including Ser-14, associate with the surface near the catalytic site. In the phosphorylated state (GPa), phosphoserine-14 interacts with two positively charged arginine residues near the allosteric site. The remarkable shift in the location of the N-terminal end of the chain cause other conformation changes in the enzyme; notably, a reorientation of two α helices, the tower helices, on the other side of the dimer interface. This, in turn, affects the position of the 280s loop controlling the transition between the active R conformation and the inactive T conformation.

The equilibrium between T and R is greatly shifted in favor of the R conformation (active) when glycogen phosphorylase is phosphorylated (GPa). GPa is relatively insensitive to ATP, AMP, and glucose 6-phosphate. In muscle cells, GPa will be formed in response to hormones that signal the need for glucose and strenuous muscle activity. This promotes rapid mobilization of glycogen. In liver cells, the liver version of glycogen phosphorylase responds to the same hormones but in this case glycogen breakdown leads to excretion of glucose that can be taken up by muscle cells. Liver glycogen phosphorylase a is inhibited by glucose by shifting GPa to the T conformation. This makes sense since the presence of a high concentration of free glucose means that it's not necessary to continue producing glucose from glycogen.

The muscle version of glycogen phosphorylase is not inhibited by glucose since muscle cells rarely see significant concentrations of free glucose. Muscle cells don't convert glucose 6-phosphate to glucose and any glucose taken up from the bloodstream is quickly phosphoryated by hexokinase to glucose 6-phosphate.



▲ **Figure 12.22**
Regulation of glycogen phosphorylase. Glycogen phosphorylase b is the unphosphorylated form of the enzyme. Glycogen phosphorylase a is phosphorylated at a position near the allosteric site. Phosphorylation is indicated by a purple ball at that site. The T conformation (red) is mostly inactive and the R conformation (green) is active in glycogen breakdown as shown by binding of inorganic phosphate (purple ball) to the catalytic site. The R conformation is greatly favored when the enzyme is phosphorylated (glycogen phosphorylase a).



▲ **Figure 12.23**
Phosphorylated and unphosphoylated forms of glycogen phosphorylase. PLP at the catalytic site is shown as a space-filling molecule. The large shift in position of Ser-14 upon phospharylation to Ser-14-P causes a conformational change that allows access to the catalytic site [PDB 3CEH, 1Z8D].

Gerty Cori and Carl Cori discovered in 1938 that glycogen phosphorylase activity was regulated by AMP. Since then, glycogen phosphorylase has been one of the prime examples of allosterically regulated enzymes, exciting three generations of biochemistry students. Glycogen phosphorylase was the very first enzyme whose regulation by covalent modification was demonstrated. Eddy Fischer and Edwin Krebs published their result in 1956 and for a long time regulation by phosphorylation was thought to be an unusual form of regulation confined to glycogen metabolism. Today, we know that phosphorylation is a very common form of regulation in eukaryotes and it is the most important part of many signal transduction pathways. There are hundreds of labs studying signal transduction.



▲ Edmond ("Eddy") H. Fischer (1920–) (left) and Edwin G. Krebs (1918–2009) (right) received the Nobel Prize in Physiology or Medicine in 1992 "for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism."

B. Hormones Regulate Glycogen Metabolism

Insulin, glucagon, and epinephrine are the principal hormones that control glycogen metabolism in mammals. Insulin, a 51-residue protein synthesized by the β cells of the pancreas, is secreted when the concentration of glucose in the blood increases. High levels of insulin are associated with the fed state of an animal. Insulin increases the rate of glucose transport into muscle and adipose tissue via the GLUT4 glucose transporter (Section 11.5A). Insulin also stimulates glycogen synthesis in the liver.

Glucagon, a peptide hormone containing 29 amino acid residues, is secreted by the α cells of the pancreas in response to a low blood glucose concentration. Glucagon restores the blood glucose concentration to a steady state level by stimulating glycogen

degradation. Glucagon is extremely selective in its target because only liver cells are rich in glucagon receptors. The effect of glucagon is opposite that of insulin and an elevated glucagon concentration is associated with the fasted state.

The adrenal glands release the catecholamine epinephrine (also known as adrenaline), in response to neural signals that trigger the fight or flight response (Figure 3.5c). The epinephrine precursor, norepinephrine, also has hormone activity. Epinephrine stimulates the breakdown of glycogen. It triggers a response to a sudden energy requirement whereas glucagon and insulin act over longer periods to maintain a relatively constant concentration of glucose in the blood. Epinephrine binds to β-adrenergic receptors of liver and muscle cells and to α_1-adrenergic receptors of liver cells. The binding of epinephrine to β-adrenergic receptors or of glucagon to its receptors activates the adenylyl cyclase signaling pathway. The second messenger, cyclic AMP (cAMP), then activates protein kinase A (PKA).

PKA phosphorylates a number of other proteins causing significant changes in metabolism. Let's look first at the regulation of glycogen metabolism by glucagon (Figure 12.24). When glucagon binds to its receptor it stimulates adenylate cyclase causing an increase in cAMP that leads to activation of PKA. PKA phosphorylates glycogen synthase converting the "a" form to the inactive "b" form. This blocks glycogen synthesis. PKA also phosphorylates another kinase called phosphorylase kinase. As the name implies, this is the kinase that phosphorylates glycogen phosphorylase. PKA activates phosphorylase kinase leading to conversion of glycogen phosphorylase b to the the active form, glycogen phosphorylase a. The result is an increase in the rate of degradation of glycogen.

The net effect of glucagon (or epinephrine) is to block synthesis of glycogen and stimulate its breakdown. The reciprocal regulation of these two enzymes is an important feature of regulation in this pathway.

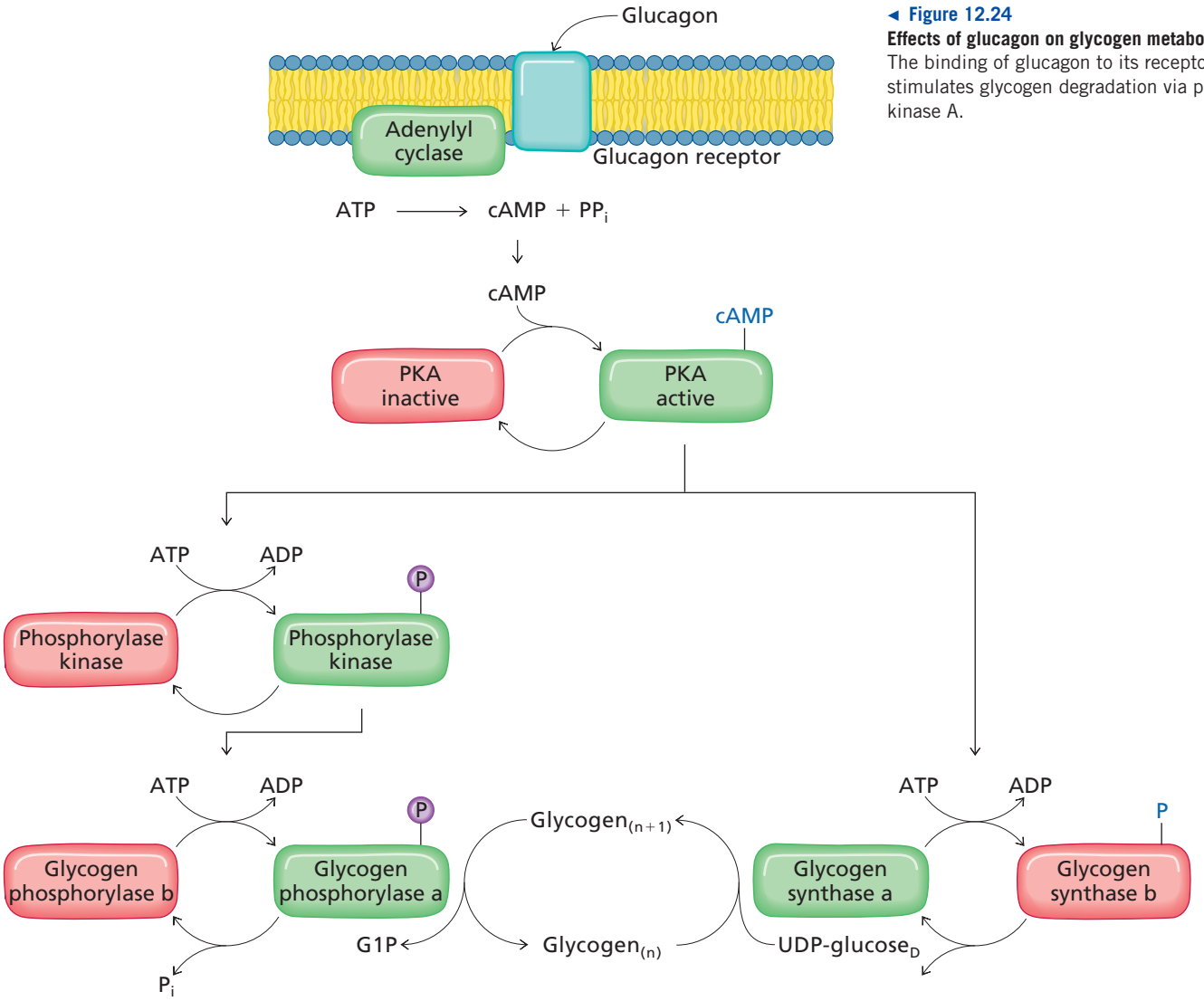
Glycogen synthase and glycogen phosphorylase are dephosphorylated by phosphoprotein phosphatase-1, an enzyme that acts on many other substrates. As shown in Figure 12.25, dephosphorylation leads to reciprocal inactivation of glycogen phosphorylase and activation of glycogen synthase. This results in synthesis of glycogen from UDP-glucose and inhibition of glycogen breakdown. Insulin stimulates the activity of phosphoprotein phosphatase-1, thus causing the uptake of glucose into glycogen and its depletion in the bloodstream. Prosphoprotein phosphatase-1 also acts on phosphorylase kinase blocking further activation of glycogen phosphorylase.

C. Hormones Regulate Gluconeogenesis and Glycolysis

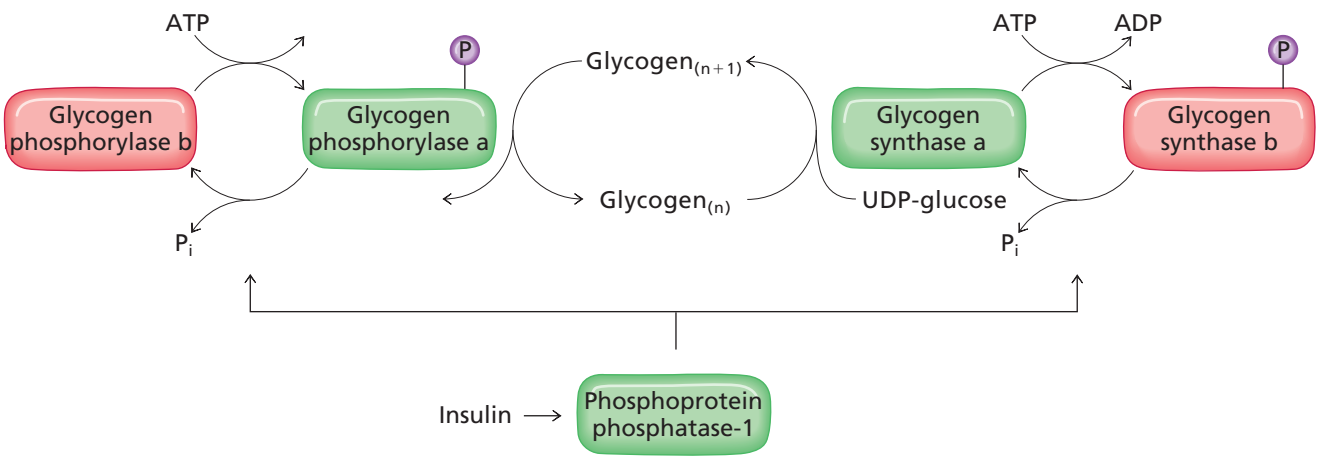
Now it's time to return to our discussion of the regulation of gluconeogenesis and glycolysis. Fructose 1,6-*bis*phosphatase (FBPase) and phosphofructokinase-1 (PFK-1) are the key enzymes involved in the decision to either degrade glucose or synthesize it (Section 12.3). Recall that these two enzymes are reciprocally regulated by the effector fructose 2,6-*bis*phosphate (Figure 12.8). This effector molecule is synthesized from fructose 6-phosphate by phosphofructokinase-2 (PFK-2) and it is dephosphorylated back to fructose 6-phosphate by fructose 2,6-*bis*phosphatase (F2,6BPase) (Figure 12.9). These two enzymatic activities are located on the same bifunctional protein. The relationship among the four enzymes and their products is summarized in Figure 12.26.

The F2,6BPase and PFK-2 activities in the bifunctional enzyme are regulated by phosphorylation in a reciprocal manner. When the protein is phosphorylated, the enyme acts as a fructose 2,6-*bis*phosphatase and the phosphofructokinase activity is inhibited. Conversely, when the enzyme is unphosphorylated it acts as a phosphofructokinase and the fructose 2,6-*bis*phosphatase activity is inhibited.

This is the same mode of reciprocal regulation we encountered with glycogen phosphorylase and glycogen synthase, except this time the two enzyme activities are on the same molecule. In the presence of glucagon, protein kinase A (PKA) is active and it phosphorylates the bifunctional enzyme (Figure 12.27). Thus, glucagon stimulates gluconeogenesis and inhibits glycolysis in liver cells causing glucose levels in the bloodstream to rise. At the same time, epinephrine can stimulate glycogen degradation and inhibit glycogen synthesis in muscle cells. The result is more glucose for muscle cells and more ATP from glycolysis.



◀ **Figure 12.24**
Effects of glucagon on glycogen metabolism. The binding of glucagon to its receptors stimulates glycogen degradation via protein kinase A.



▼ **Figure 12.25**
Effect of insulin on glycogen metabolism. Insulin simulates the phosphatase activity of phosphoprotein phosphatase-1, leading to inactivation of glycogen phosphorylase and activation of glycogen synthase.

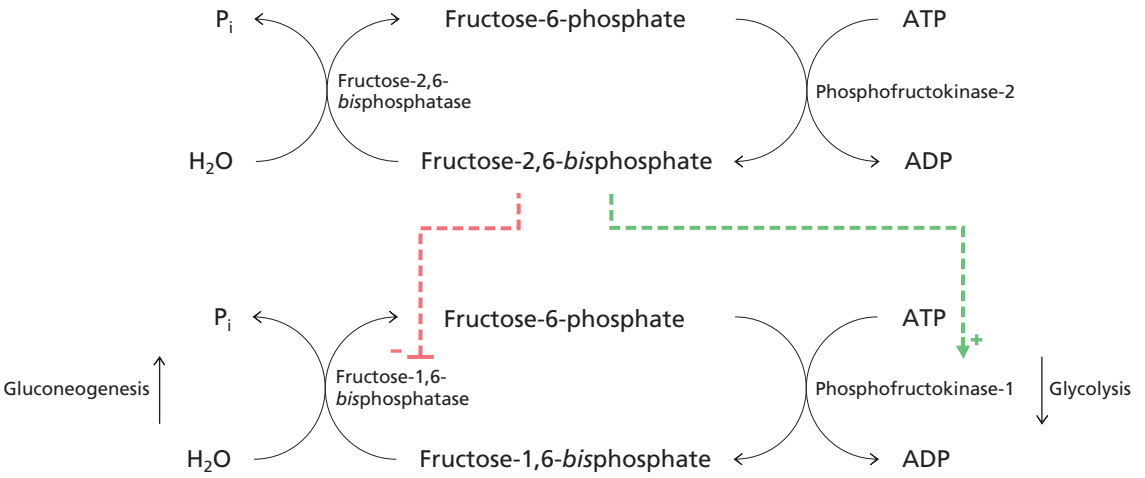


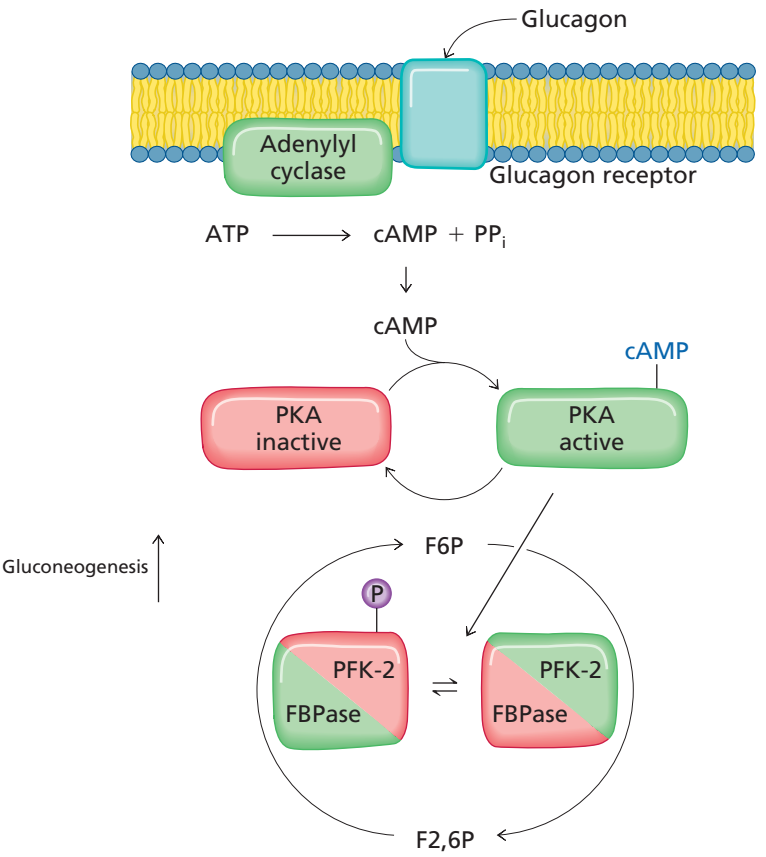
Figure 12.26 ▲
The role of fructose 2,6-*bis*phosphate in regulating glycolysis and gluconeogenesis.

12.7 Maintenance of Glucose Levels in Mammals

Mammals maintain blood glucose levels within strict limits by regulating both the synthesis and degradation of glucose. Glucose is the major metabolic fuel in the body. Some tissues, such as brain, rely almost entirely on glucose for their energy needs. The concentration of glucose in the blood seldom drops below 3 mM or exceeds 10 mM. When the concentration of glucose in the blood falls below 2.5 mM, glucose uptake into the brain is compromised, with severe consequences. Conversely, when blood glucose levels are very high, glucose is filtered out of the blood by the kidneys accompanied by osmotic loss of water and electrolytes.

The liver plays a unique role in energy metabolism participating in the interconversions of all types of metabolic fuels: carbohydrates, amino acids, and fatty acids.

Figure 12.27 ►
Effect of glucagon on gluconeogenesis. Glucagon binds to its receptor, causing activation of adenylate cyclase. Increased levels of cAMP activate protein kinase A, which phosphorylates the bifunctional enzyme leading to activation of fructose 2,6-*bis*phosphatase activity. In the absence of the effector fructose 2,6-*bis*phosphate, fructose 1,6-*bis*phosphatase is activated and this increases flux in the gluconeogenesis pathway.

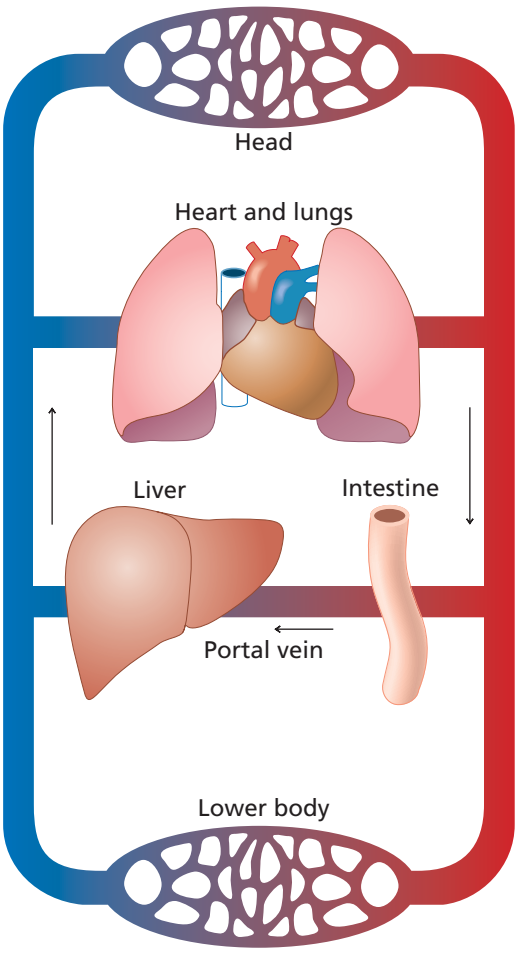


Anatomically, the liver is centrally located in the circulatory system (Figure 12.28). Most tissues are perfused in parallel with the arterial system supplying oxygenated blood and the venous circulation returning blood to the lungs for oxygenation. The liver, however, is perfused in series with the visceral tissues (gastrointestinal tract, pancreas, spleen, and adipose tissue); blood from these tissues drains into the portal vein and then flows to the liver. This means that after the products of digestion are absorbed by the intestine, they pass immediately to the liver. Using its specialized complement of enzymes, the liver regulates the distribution of dietary fuels and supplies fuel from its own reserves when dietary supplies are exhausted.

The consumption of glucose by tissues removes dietary glucose from the blood. When glucose levels fall, liver glycogen and gluconeogenesis become the sources of glucose. However, since these sources are limited, hormones act to restrict the use of glucose to those cells and tissues that absolutely depend on glycolysis for generating ATP (kidney medulla, retina, red blood cells, and parts of the brain). Other tissues can generate ATP by oxidizing fatty acids mobilized from adipose tissue (Sections 16.1C and 16.2).

The complexity of carbohydrate metabolism in mammals is evident from the changes that occur on feeding and starvation. In the 1960s, George Cahill examined the glucose utilization of obese patients as they underwent therapeutic starvation. After an initial feeding of glucose, the subjects received only water, vitamins, and minerals. Cahill noted that glucose homeostasis (maintenance of constant levels in the circulation) proceeds through five phases. Figure 12.29, based on Cahill's observations, summarizes the metabolic changes in the five phases.

1. During the initial absorptive phase (the first four hours), dietary glucose enters the liver via the portal vein and most tissues use glucose as the primary fuel. Under these conditions, the pancreas secretes insulin, which stimulates glucose uptake by muscle and adipose tissue via GLUT4. The glucose taken up by these tissues is



◀ **Figure 12.28**
Placement of the liver in the circulatory system. Most tissues are perfused in parallel. However, the liver is perfused in series with visceral tissues. Blood that drains from the intestine and other visceral tissues flows to the liver via the portal vein. The liver is therefore ideally placed to regulate the passage of fuels to other tissues.

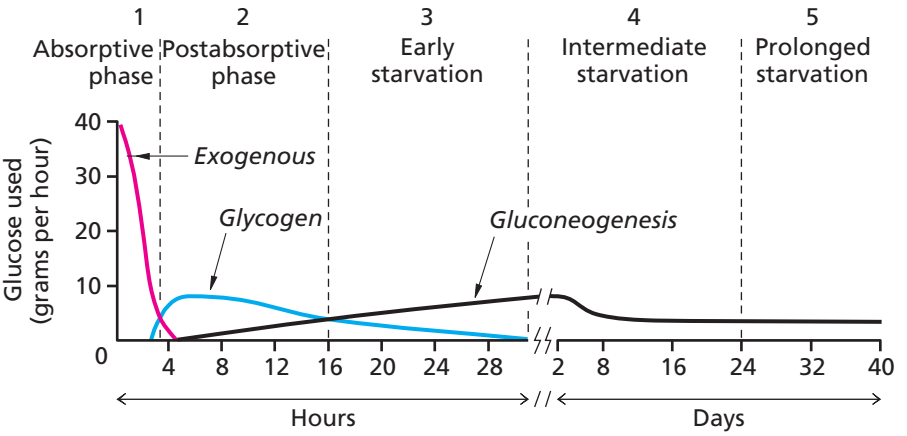


Figure 12.29 ►
Five phases of glucose homeostasis. The graph, based on observations of a number of individuals, illustrates glucose utilization in a 70 kg man who consumed 100 g of glucose and then fasted for 40 days.

phosphorylated to glucose 6-phosphate, which cannot diffuse out of the cells. Liver cells also absorb glucose and convert it to glucose 6-phosphate. Excess glucose is stored as glycogen in liver and muscle cells.

2. When the dietary glucose is consumed, the body mobilizes liver glycogen to maintain circulating glucose levels. In the liver, glucose 6-phosphatase catalyzes the hydrolysis of glucose 6-phosphate to glucose, which exits the liver via glucose transporters. Glycogen in muscle (which lacks glucose 6-phosphatase) is metabolized to lactate to produce ATP for contraction; the lactate is used by other tissues as a fuel or by the liver for gluconeogenesis.
3. After about 24 hours, liver glycogen is depleted, and the only source of circulating glucose is gluconeogenesis in the liver, using lactate, glycerol, and alanine as precursors. Fatty acids mobilized from adipose tissue become an alternate fuel for most tissues. The obligatory glycolytic tissues continue to use glucose and produce lactate, which is converted to glucose in the liver by the Cori cycle; this cycle makes energy, not carbon, from fatty acid oxidation in the liver available to other tissues.
4. Gluconeogenesis in the liver continues at a high rate for a few days, then decreases. As starvation progresses, gluconeogenesis in the kidney becomes proportionately more significant. Proteins in peripheral tissues are broken down to provide gluconeogenic precursors. In this phase, the body adapts to several alternate fuels.
5. In prolonged starvation, there is less gluconeogenesis and lipid stores are depleted. If refeeding does not occur, death will follow. On refeeding, metabolism is quickly restored to the conditions of the fed state.

We have seen how glucose, a major fuel, can be stored in polysaccharide form and mobilized as needed. Glucose can also be synthesized from noncarbohydrate precursors by the reactions of gluconeogenesis. We have seen that glucose can be oxidized by the pentose phosphate pathway to produce NADPH or transformed by glycolysis into pyruvate.

The effect of insulin and diabetes on the production of ketone bodies is described in Section 16.11 (Box 16.6).

Diabetes mellitus (DM) is a metabolic disease that results from improper regulation of carbohydrate and lipid metabolism. Despite an ample supply of glucose, the body behaves as though starved and glucose is overproduced by the liver and underused by other tissues. As a result, the concentration of glucose in the blood is extremely high. The levels of glucose in the blood often exceed the capacity of the kidney to reabsorb glucose so some of it spills into the urine. The high concentration of glucose in urine draws water osmotically from the body.

There are two types of diabetes both of which arise from faulty control of fuel metabolism by the hormone insulin. In Type 1 diabetes mellitus (also called insulin-dependent diabetes mellitus, or IDDM) damage to the β cells of the pancreas, where insulin is synthesized, results in diminished or absent secretion of insulin. This autoimmune disease is characterized by early onset (usually before age 15). Patients are thin and exhibit hyperglycemia (high blood glucose levels), dehydration, excessive urination, hunger, and thirst. In Type 2 (also called non-insulin-dependent diabetes, or NIDDM), chronic hyperglycemia results from insulin resistance—decreased

sensitivity to insulin possibly caused by a shortage or decreased activity of insulin receptors. Insulin secretion may be normal and circulating levels of insulin may even be elevated. This type is also known as adult-onset diabetes (although its incidence is increasing among children) and it is usually associated with obesity. Type 2 diabetes affects about 5% of the population and Type 1 affects about 1%. In addition, about 2% to 5% of pregnant women develop a form of diabetes. Most women who exhibit gestational diabetes return to normal after giving birth but are at risk for developing Type 2 diabetes.

To understand diabetes, we must consider the functions of insulin. Insulin stimulates the synthesis of glycogen, triacylglycerols, and proteins and inhibits the breakdown of these compounds. Insulin also stimulates glucose transport into muscle cells and adipocytes. When insulin levels are low in IDDM, glycogen is broken down in the liver and gluconeogenesis occurs regardless of the glucose supply. In addition, glucose uptake and its use in peripheral tissues are restricted.

12.8 Glycogen Storage Diseases

Several metabolic diseases are related to the storage of glycogen. The general rule about metabolic diseases is that they usually affect the activity of nonessential genes and enzymes. Defects in essential genes are usually lethal and don't show up as metabolic diseases.

Many metabolic enzymes in humans are encoded by gene families. Different versions are expressed in different tissues. In the case of enzymes involved in glycogen metabolism, the most common versions are found in liver and muscle. A deficiency in one of these enzymes will produce severe symptoms but may not be lethal. There are nine types of glycogen storage diseases resulting from defects in glycogen metabolism.

Type 0: In type 0a, the activity of liver glycogen synthase is affected. The gene for this enzyme is on the short arm of chromosome 12 at locus 12p12.2 (MIM = 240600). This is a severe disease causing early death in cases where the activity is very low. Type 0b affects the muscle version of glycogen synthase whose gene is on the long arm of chromosome 19 at 19q13.3 (MIM = 611556). Patients have no muscle glycogen and are unable to engage in strenuous physical activity.

Type I: The most common glycogen storage disease is called von Gierke disease. It is caused by a deficiency in glucose 6-phosphatase (Type 1a, MIM = 23220) whose gene is on chromosome 17 (17q21). Defects in the complex that transports glucose across the endoplasmic reticulum (Section 21.1D) also cause von Gierke's disease. Type 1b affects the glucose 6-phosphate transporter (chromosome 11 (11q23), MIM = 232220) and type 1c affects the phosphate transporter (chromosome 6 (6p21.3), MIM = 232240). Patients are unable to secrete glucose leading to accumulation of glycogen in the liver and kidneys.

Type II: Patients suffering from type II disease, known as Pompe's disease, suffer from reduced activity of α-1,4-glucosidase, or acid maltase, an enzyme required for glycogen breakdown in lysozomes (MIM = 232300). The gene is on chromosome 17 (17q25.2). The defect causes glycogen to accumulate in lysosomes leading to problems with muscle tissue, especially in the heart. In the most severe forms, children die within the first few years of life.

Type III: Type III is Cori disease, characterized by defects in the gene encoding the glycogen debranching enzyme in liver and muscle (chromosome 1 (1p21), MIM = 232400). People suffering from this disease have weakened muscles because they are unable to mobilize all of the stored glycogen. Some defects have very mild symptoms.

Type IV: Often called Anderson's disease, the mutations occur in the gene for liver branching enzyme found on chromosome 3 (3p12, MIM = 232500). Long-chain polysaccharides accumulate in patients with these mutations, resulting in death within a few years from heart failure or liver failure.

MIM numbers refer to the Online Mendelian Inheritance in Man (OMIM) database at: *ncbi.nlm.nih.gov/omim*

Type V: McArdle's disease (type V glycogen storage disease) is caused by a deficiency of muscle glycogen phosphorylase (MIM = 232600). The gene is on chromosome 11 (11q13). Individuals having this genetic disease cannot perform strenuous exercise and suffer painful muscle cramps.

Type VI: Hers' disease (type VI) is a mild form of glycogen storage disease due to a deficiency in liver glycogen phosphorylase (MIM = 232700). Several mutant alleles interfere with proper splicing of the primary transcript from the gene on chromosome 14 (14q21).

Type VII: Mutations in the gene for muscle phosphofructokinase-1 cause Tarui's disease, characterized by inability to exercise and muscle cramps (MIM = 232800). The gene for this isozyme is on chromosome 12 (12q13.3).

Type VIII: Now recognized as a subtype of type IX.

Type IX: This form of glycogen storage disease manifests as muscle weakness and/or muscle cramps. The symptoms are usually mild. All subtypes are due to mutations in the genes for the various subunits of glycogen phosphorylase kinase. Type IXa: liver α subunit gene on the X chromosome at Xp20 (MIM = 300798). Type IXb: β subunit gene at 16q12 (MIM = 172490). Type IXc: liver γ subunit gene at 16p12 (MIM = 172471). Type IXd: muscle α subunit gene on the X chromosome at Xq13 (MIM = 311870).

Summary

1. Gluconeogenesis is the pathway for glucose synthesis from noncarbohydrate precursors. The seven near-equilibrium reactions of glycolysis proceed in the reverse direction in gluconeogenesis. Four enzymes specific to gluconeogenesis catalyze reactions that bypass the three metabolically irreversible reactions of glycolysis.
2. Noncarbohydrate precursors of glucose include pyruvate, lactate, alanine, and glycerol.
3. Gluconeogenesis is regulated by glucagon, allosteric modulators, and the concentrations of its substrates.
4. The pentose phosphate pathway metabolizes glucose 6-phosphate to generate NADPH and ribose 5-phosphate. The oxidative stage of the pathway generates two molecules of NADPH per molecule of glucose 6-phosphate converted to ribulose 5-phosphate and CO_2. The nonoxidative stage includes isomerization of ribulose 5-phosphate to ribose 5-phosphate. Further metabolism of pentose phosphate molecules can convert them to glycolytic intermediates.The combined activities of transketolase and transaldolase convert pentose phosphates to triose phosphates and hexose phosphates.
5. Glycogen synthesis is catalyzed by glycogen synthase, using a glycogen primer and UDP-glucose.
6. Glucose residues are mobilized from glycogen by the action of glycogen phosphorylase. Glucose 1-phosphate is then converted to glucose 6-phosphate.
7. Glycogen degradation and glycogen synthesis are reciprocally regulated by hormones. Kinases and phosphatases control the activities of the interconvertible enzymes glycogen phosphorylase and glycogen synthase.
8. Mammals maintain a nearly constant concentration of glucose in the blood. The liver regulates the amount of glucose supplied by the diet, glycogenolysis, and other fuels.
9. Glycogen storage diseases result from defects in genes required for glycogen metabolism.

Problems

1. Write a balanced equation for the synthesis of glucose from pyruvate. Assuming that the oxidation of NADH is equal to 2.5 ATP equivalents (Section 14.11), how many ATP equivalents are required in this pathway? Convert this to kJ mol^{-1} and explain how this value compares to the total energy required to synthesize glucose from CO_2 and H_2O.
2. What important products of the citric acid cycle are required for gluconeogenesis from pyruvate?
3. Epinephrine promotes the utilization of stored glycogen for glycolysis and ATP production in muscles. How does epinephrine promote the use of liver glycogen stores for generating the energy needed by contracting muscles?
4. (a) In muscle cells, insulin stimulates a protein kinase that catalyzes phosphorylation of protein phosphatase-1, thereby activating it. How does this affect glycogen synthesis and degradation in muscle cells?
   (b) Why does glucagon selectively regulate enzymes in the liver but not in other tissues?
   (c) How does glucose regulate the synthesis and degradation of liver glycogen via protein phosphatase-1?
5. The polypeptide hormone glucagon is released from the pancreas in response to low blood glucose levels. In liver cells, glucagon plays a major role in regulating the rates of the opposing glycolysis

and gluconeogenesis pathways by influencing the concentrations of fructose 2,6-*bis*phosphate (F2,6 BP). If glucagon causes a decrease in the concentrations of F2,6 BP, how does this result in an increase in blood glucose levels?

6. When the concentration of glucagon rises in the blood, which of the following enzyme activities is decreased? Explain.

Adenylyl cyclase
Protein kinase A
PFK-2 (kinase activity)
Fructose 1, 6-*bis*phosphatase

7. (a) Is the energy required to synthesize glycogen from glucose 6-phosphate greater than the energy obtained when glycogen is degraded to glucose 6-phosphate?

(b) During exercise, glycogen in both muscle and liver cells can be converted to glucose metabolites for ATP generation in the muscles. Do liver glycogen and muscle glycogen supply the same amount of ATP to the muscles?

8. Individuals with a total deficiency of muscle glycogen phosphorylase (McArdle's disease) cannot exercise strenuously due to muscular cramping. Exertion in these patients leads to a much greater than normal increase in cellular ADP and P_i. Furthermore, lactic acid does not accumulate in the muscles of these patients, as it does in normal individuals. Explain the chemical imbalances in McArdle's disease.

9. Compare the number of ATP equivalents generated in the breakdown of one molecule of glucose 1-phosphate into two molecules of lactate with the number of ATP equivalents required for the synthesis of one molecule of glucose 1-phosphate from two molecules of lactate. (Assume anaerobic conditions.)

10. (a) How does the glucose–alanine cycle allow muscle pyruvate to be used for liver gluconeogenesis and subsequently returned to muscles as glucose?

(b) Does the glucose–alanine cycle ultimately provide more energy for muscles than the Cori cycle does?

11. Among other effects, insulin is a positive modulator of the enzyme glucokinase in liver cells. If patients with diabetes mellitus produce insufficient insulin, explain why these patients cannot properly respond to increases in blood glucose.

12. Glycogen storage diseases (GSDs) due to specific enzyme deficiencies can affect the balance between glycogen stores and blood glucose. Given the following diseases, predict the effects of each on (1) the amount of liver glycogen stored and (2) blood glucose levels.

(a) Von Gierke disease (GSD-1a), defective enzyme: glucose 6-phosphatase.
(b) Cori's disease (GSD III), defective enzyme: amylo-1,6 glucosidase (debranching enzyme).
(c) Hers' disease (GSD VI), defective enzyme: liver phosphorylase

13. The pentose phosphate pathway and the glycolytic pathway are interdependent, since they have in common several metabolites whose concentrations affect the rates of enzymes in both pathways. Which metabolites are common to both pathways?

14. In many tissues, one of the earliest responses to cellular injury is a rapid increase in the levels of enzymes in the pentose phosphate pathway. Ten days after an injury, heart tissue has levels of glucose 6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase that are 20 to 30 times higher than normal, whereas the levels of glycolytic enzymes are only 10% to 20% of normal. Suggest an explanation for this phenomenon.

15. (a) Draw the structures of the reactants and products for the second reaction catalyzed by transketolase in the pentose phosphate pathway. Show which carbons are transferred.

(b) When 2-[^{14}C]-glucose 6-phosphate enters the pathway, which atom of fructose 6-phosphate produced by the reaction in Part (a) is labeled?

Selected Readings

Gluconeogenesis

Hanson, R. W., and Hakimi, P. (2008). Born to run. *Biochimie.* 90:838–842.

Hanson, R. W., and Reshef, L. (1997). Regulation of phosphenolpyruvate carboxykinase (GTP) gene expression. *Annu. Rev. Biochem.* 66:581–611. Describes the metabolic control of gene expression.

Hines, J. K., Chen, X., Nix, J. C., Fromm, H. J., and Honzatko, R. B. (2007). Structures of mammalian and bacterial fructose-1,6-bisphosphatase reveal the basis for synergism in AMP/fructose 2,6-bisphosphate inhibition. *J. Biol. Chem.* 282:36121–36131.

Jitrapakdee, S., and Wallace, J. C. (1999). Structure, function and regulation of pyruvate carboxylase. *Biochem. J.* 340:1–16.

Kemp, R. G. and Gunasekera, D. (2002). Evolution of the allosteric ligand sites of mammalian phosphofructo-1-kinase. *Biochem.* Biochemistry 41:9426–9430.

Ou, X., Ji, C., Han, X., Zhao, X., Li, X., Mao, Y., Wong, L-L., Bartlam, M., and Rao, Z. (2006). Crystal structure of human glycerol 3-phosphate dehydrogenase (GPD1). *J. Mol. Biol.* 357:858–869.

Pilkis, S. J., and Granner, D. K. (1992). Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annu. Rev. Physiol.* 57:885–909.

Rothman, D. L., Magnusson, I., Katz, L. D., Shulman, R. G., and Shulman, G. I. (1991). Quantitation of hepatic glycogenolysis and gluconeogenesis in fasting humans with ^{13}C NMR. *Science.* 254:573–576. Describes the continuous operation of the pathway of gluconeogenesis in humans.

Sullivan, S. M., and Holyoak (2008). Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proc. Natl. Acad. Sci. (USA)* 105:13829–13834.

van de Werve, G., Lange, A., Newgard, C., Méchin, M.-C., Li, Y., and Berteloot, A. (2000). New lessons in the regulation of glucose metabolism taught by the glucose 6-phosphatase system. *Eur. J. Biochem.* 267:1533–1549. Explains why there is still much to learn about the catalytic site and the transporter associated with this enzyme.

Xue, Y., Huang, S., Liang, J. Y., Zhang, Y., and Lipscomb, W. N. (1994). Crystal structure of fructose-1,6-bisphosphatase complexed with fructose 2,6-bisphosphate, AMP, and Zn2+ at 2.0-A resolution: aspects of synergism between inhibitors. *Proc. Natl. Acad. Sci.* (USA) 91:12482–12486.

Pentose Phosphate Pathway

Au, S.W.N., Gover, S., Lam, V.M.S., and Adams, M.J. (2000) Human glucose-6-phospate dehydrogenase: the crystal structure reveals a structural $NADP^+$ molecule and provides insights into enzyme deficiency. Structure 8:293–303.

Wood, T. (1985). *The Pentose Phosphate Pathway*. (Orlando: Academic Press).

Wood, T. (1986). Physiological functions of the pentose phosphate pathway. *Cell Biochem. Func.* 4:241–247.

Glycogen Metabolism

Barford, D. Hu, S-H., and Johnson, L. N. (1991). Structural mechanisms for glycogen phosphorylase control by phosphorylation and AMP. *J. Mol. Biol.* 218:233–260.

Chou, J. Y., Matern, D., Mansfield, B. C., and Chen, Y. T. (2002). Type I glycogen storage diseases: disorders of the glucose 6-phosphate complex. *Curr. Mol. Med.* 2:121–143.

Cohen, P., Alessi, D. R., and Cross, D. A. E. (1997). PDK1, one of the missing links in insulin signal transduction? *FEBS Lett.* 410:3–10.

Fischer, E. (2010). Memories of Ed Krebs. *J. Biol. Chem.* 285:4267.

Johnson, L. N. (2009). Novartis Medal Lecture: The regulation of protein phosphorylation. *Biochem. Soc. Trans.* 37:627–641.

Johnson, L. N., and Barford, D. (1990). Glycogen phosphorylase: the structural basis of the allosteric response and comparison with other allosteric proteins. *J. Biol. Chem.* 265:2409–2412.

Johnson, L. N., Lowe, E. D., Noble, M. E. M., and Owen, D. J. (1998). The structural basis for substrate recognition and control by protein kinases. *FEBS Lett.* 430:1–11.

Larner, J. (1990). Insulin and the stimulation of glycogen synthesis: the road from glycogen synthase to cyclic AMP-dependent protein kinase to insulin mediators. *Adv. Enzymol. Mol. Biol.* 63:173–231.

Meléndez-Hevia, E., Waddell, T. G., and Shelton, E. D. (1993). Optimization of molecular design in the evolution of metabolism: the glycogen molecule. *Biochem. J.* 295:477–483.

Murray, R. K., Bender, D. A., Kennelly, P. J., Rodwell, V. W., and Weil. P. A. (2009). *Harper's Illustrated Biochemistry*, 28th ed. (New York: McGraw-Hill).

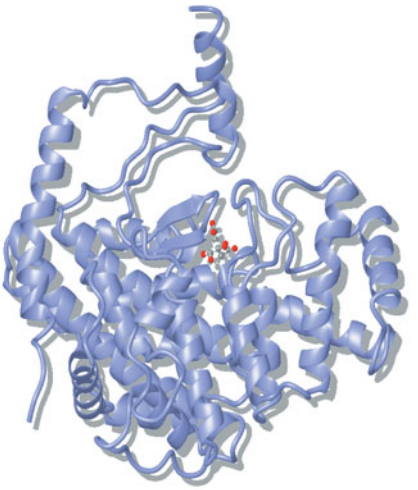
Pinotsis, N., Leonidas, D. D., Chrysina, E. D., Oikonomakos, N. G., and Mavridis, I. M. (2003). The binding of β- and γ-cyclodextrins to glycogen phosphorylase b: kinetic and crystallographic studies. *Prot. Sci.* 12:1914–1924.

Shepherd, P. R., Withers, D. J., and Siddle, K. (1998). Phosphoinositide 3-kinase: the key switch mechanism in insulin signalling. *Biochem. J.* 333:471–490.

Smythe, C., and Cohen, P. (1991). The discovery of glycogenin and the priming mechanism for glycogen biosynthesis. *Eur. J. Biochem.* 200:625–631.

Villar-Palasi, C., and Guinovart, J. J. (1997). The role of glucose 6-phosphate in the control of glycogen synthase. *FASEB J.* 11:544–558.





The Citric Acid Cycle

Since citric acid reacts catalytically in the tissue it is probable that it is removed by a primary reaction but regenerated by a subsequent reaction. In the balance sheet no citrate disappears and no intermediate products accumulate.

—H. A. Krebs and W. A. Johnson (1937)

In the last two chapters we were mainly concerned with the synthesis and degradation of complex carbohydrates such as glucose. We saw that the biosynthetic pathway leading to glucose began with pyruvate and oxaloacetate and that pyruvate was the end product of glycolysis. In this chapter we will describe pathways that interconvert a number of simple organic acids. Several of these compounds are essential precursors for the biosynthesis of amino acids, fatty acids, and porphyrins.

Acetyl CoA is one of the key intermediates in the interconversion of small organic acids. Acetyl CoA is formed by the oxidative decarboxylation of pyruvate with the release of CO_2. This reaction is catalyzed by pyruvate dehydrogenase, an enzyme that we briefly encountered in Section 11.3 when we discussed the fate of pyruvate. We begin this chapter with a more detailed description of this important enzyme.

The acetyl group (a two-carbon organic acid) from acetyl CoA can be transferred to the four-carbon dicarboxylic acid, oxaloacetate, to form a new six-carbon tricarboxylic acid known as citrate (citric acid). Citrate can then be oxidized in a seven-step pathway to regenerate oxaloacetate and release two molecules of CO_2. Oxaloacetate can recombine with another molecule of acetyl CoA and the citrate oxidation reactions are repeated. The net effect of this eight-enzyme cyclic pathway is the complete oxidation of an acetyl group to CO_2 and the transfer of electrons to several cofactors to form reducing equivalents. The pathway is known as the **citric acid cycle**, the tricarboxylic acid cycle (TCA cycle), or the Krebs cycle, after Hans Krebs who discovered it in the 1930s.

The citric acid cycle lies at the hub of energy metabolism in eukaryotic cells—especially in animals. The energy released in the oxidations of the citric acid cycle is largely conserved as reducing power when the coenzymes $NAD^{\oplus}$ and ubiquinone (Q) are reduced to form NADH and QH_2. This energy is ultimately derived from pyruvate (via acetyl CoA). Since pyruvate is the end product of glycolysis, we can think of the citric acid cycle as a series of reactions that complete the oxidation of glucose. NADH and QH_2 are substrates in the reactions of membrane-associated electron transport leading to the formation of a proton gradient that drives the synthesis of ATP (Chapter 14).

Top: Citrate synthase with its product citrate in the active site. This enzyme catalyzes the first step of the citric acid cycle. [PDB 1CTS]

Hans Krebs and W. A. Johnson proposed the citric acid cycle in 1937 in order to explain several puzzling observations. They were interested in understanding how the oxidation of glucose in muscle cells was coupled to the uptake of oxygen. Albert Szent-Györgyi had previously discovered that adding a four-carbon dicarboxylic acid—succinate, fumarate, or oxaloacetate—to a suspension of minced muscle stimulated the consumption of O_2. The substrate of the oxidation was carbohydrate, either glucose or glycogen. Especially intriguing was the observation that adding small amounts of four-carbon dicarboxylic acids caused larger amounts of oxygen to be consumed than were required for their own oxidation. This indicated that these four-carbon organic acids had catalytic effects.

Krebs and Johnson observed that citrate, a six-carbon tricarboxylic acid, and α-ketoglutarate, a five-carbon compound, also had a catalytic effect on the uptake of O_2. They proposed that citrate was formed from a four-carbon intermediate and an unknown two-carbon derivative of glucose (later shown to be acetyl CoA). The cyclic nature of the pathway explained how its intermediates could act catalytically without being consumed. Albert Szent-Györgyi received the Nobel Prize in Physiology or Medicine in 1937 for his work on respiration, including the catalytic role of fumarate in biological combustion processes. Hans Krebs was awarded the Nobel Prize in Physiology or Medicine in 1953 for discovering the citric acid cycle.

In muscle cells, the intermediates in the citric acid cycle are almost exclusively used in the cyclic pathway of energy metabolism. In these cells, the metabolic machinery is mainly devoted to extracting energy from glucose in the form of ATP. This is why it was possible to recognize the cyclic nature of the pathway by carrying out experiments on muscle extracts. In other cells, the intermediates of the citric acid cycle are the starting points for many biosynthetic pathways. Thus, the enzymes of the citric acid cycle play a key role in both anabolic and catabolic reactions.

Many of these same enzymes are found in prokaryotes although few bacteria possess a complete citric acid cycle. In this chapter, we examine the reactions of the citric acid cycle as they occur in eukaryotic cells. We will explore how these enzymes are regulated. Next we will introduce the various biosynthetic pathways that require citric acid cycle intermediates and examine the relationship of these pathways to the main reactions of the cyclic pathway in eukaryotes and the partial pathways in bacteria. We will also look at pathways involving glyoxylate, specifically the glyoxylate shunt and the glyoxylate cycle. These are pathways that are closely related to the citric acid cycle. Finally, we will discuss the evolution of the citric acid cycle enzymes.

BOX 13.1 AN EGREGIOUS ERROR

In 1937, Krebs and Johnson submitted a paper to *Nature* outlining their discovery of citric acid as a catalyst in the oxidation of glucose by muscle tissue. The journal declined to publish the paper on the grounds that it had too many papers in press. Krebs writes in his memoirs, "This was the first time in my career, after having published more than fifty papers, that I experienced a rejection or semi-rejection."

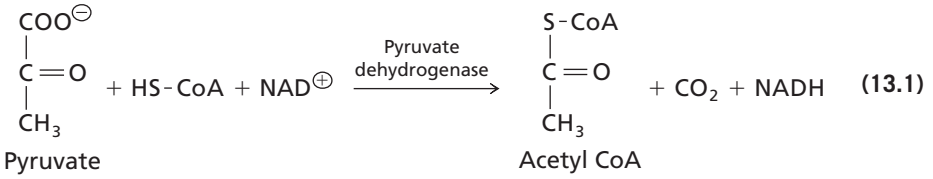
Krebs and Johnson published the paper in the journal *Enzymologia* and Krebs went on to win the Nobel Prize based largely on this paper. It took *Nature* 51 years to publically recognize the mistake it made. An editor wrote in the October 28, 1988 issue, "An editor's nightmare is to reject a Nobel-prizewinning paper. . . . Rejection of Hans Krebs' discovery of the tricarboxylic (or Krebs') cycle, a pivot of biochemical metabolism, remains *Nature's* most egregious error (as far as we know)."



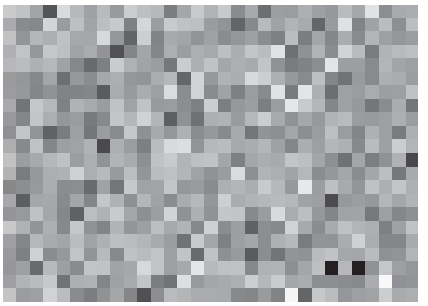
▶ **Hans Krebs (1900–1981).** Krebs was awarded the Nobel Prize in Physiology or Medicine in 1953 for his discovery of the citric acid cycle. He is shown here beside a Warburg apparatus for measuring oxygen consumption in metabolizing tissue. Krebs worked with Otto Warburg in the 1920s.

13.1 Conversion of Pyruvate to Acetyl CoA

Pyruvate is a key substrate in a number of reactions, as described in Section 11.3. In this chapter we are concerned with the conversion of pyruvate to acetyl CoA since acetyl CoA is the main substrate of the citric acid cycle. The reaction is catalyzed by a large complex of enzymes and cofactors known as the pyruvate dehydrogenase complex (Figure 13.1). The stoichiometry of the complete reaction is



(13.1)



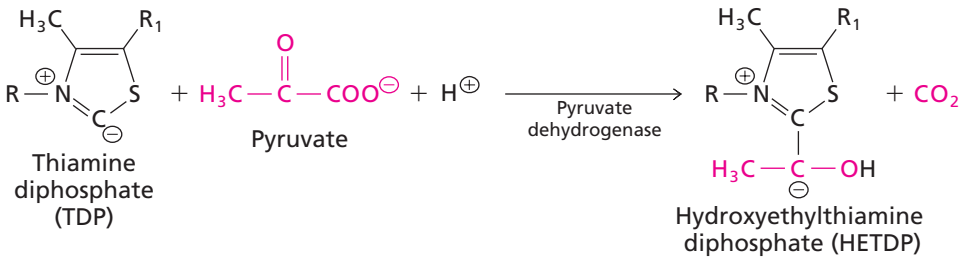
▲ **Figure 13.1**
Electron micrograph of pyruvate dehydrogenase complexes from *E. coli*.

where HS-CoA is coenzyme A. This is the first step in the oxidation of pyruvate and the products of the reaction are acetyl CoA, one molecule of carbon dioxide, and one molecule of reducing equivalent (NADH). The pyruvate dehydrogenase reaction is an oxidation–reduction reaction. In this case, the oxidation of pyruvate to CO_2 is coupled to the reduction of $NAD^{\oplus}$ to NADH. The net result is the transfer of two electrons from pyruvate to NADH.

The pyruvate dehydrogenase complex is a multienzyme complex containing multiple copies of three distinct enzymatic activities: pyruvate dehydrogenase (E_1 subunits), dihydrolipoamide acetyltransferase (E_2 subunits), and dihydrolipoamide dehydrogenase (E_3 subunits). The oxidative decarboxylation of pyruvate can be broken down into five steps. (In each step of the following reactions the fates of the atoms from pyruvate are shown in red.)

The systematic names of the enzymes in the complex are: pyruvate lipoamide 2-oxidoreductase (E_1); acetyl CoA:dihydrolipoamide *S*-acetyltransferase (E_2); and dihydrolipoamide:$NAD^{\oplus}$ oxidoreductase (E_3).

1. The E_1 component contains the prosthetic group thiamine diphosphate (TDP). As we saw in Chapter 7, TDP (vitamin B_1) plays a catalytic role in a number of decarboxylase reactions. The initial reaction results in the formation of a hydroxyethyl–TDP intermediate and the release of CO_2.

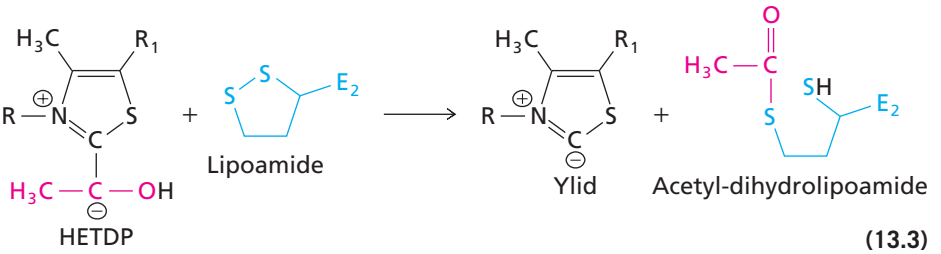


(13.2)

Note that the reactive form of TDP is the carbanion or ylid form. The carbanion form is relatively stable because of the unique environment of the coenzyme bound to the protein (Section 7.6). The product of the first step is the carbanion form of hydroxyethyl–TDP. The mechanism is similar to the pyruvate decarboxylase mechanism (Section 7.7).

2. In the second step, the two-carbon hydroxylethyl group is transferred to the lipoamide group of E_2. The lipoamide group consists of lipoic acid covalently bound by an amide linkage to a lysine residue of an E_2 subunit (Figure 7.29). This particular coenzyme is only found in pyruvate dehydrogenase and related enzymes.

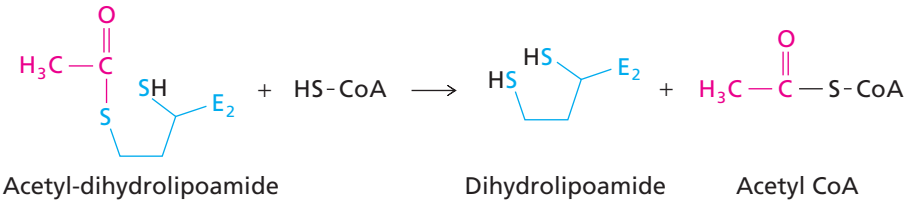
The transfer reaction is catalyzed by the E_1 component of the pyruvate dehydrogenase complex.



(13.3)

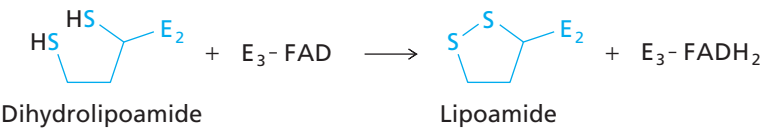
In this reaction, the oxidation of hydroxyethyl–TDP is coupled to the reduction of the disulfide of lipoamide and the acetyl group is transferred to one of the sulfhydryl groups of the coenzyme regenerating the ylid form of TDP.

3. The third step involves the transfer of the acetyl group to HS-CoA, forming acetyl CoA and leaving the lipoamide in the reduced dithiol form. This reaction is catalyzed by the E_2 component of the complex.



(13.4)

4. The reduced lipoamide of E_2 must be reoxidized in order to regenerate the prosthetic group for additional reactions. This is accomplished in step 4 by transferring two protons and two electrons from the dithiol form of lipoamide to FAD. FAD is the prosthetic group of E_3 and the redox reaction produces the reduced coenzyme ($FADH_2$). (Recall from Section 7.5 that $FADH_2$ carries two electrons and two protons that are usually acquired as a single proton and a hydride ion.)



(13.5)

5. In the final step, E_3–$FADH_2$ is reoxidized to FAD. This reaction is coupled to the reduction of $NAD^{\oplus}$.

$$E_3\text{—}FADH_2 + NAD^{\oplus} \longrightarrow E_3\text{—}FAD + NADH + H^{\oplus} \qquad \textbf{(13.6)}$$

The oxidation of E_3–$FADH_2$ regenerates the original pyruvate dehydrogenase complex, completing the catalytic cycle. Step 5 produces NADH and $H^{\oplus}$. Note that one proton is released in step 5 and one proton is taken up in step 1 so that the overall stoichiometry of the pyruvate dehydrogenase reaction shows no net gain or loss of protons (Reaction 13.1).

The interplay of five coenzymes in the pyruvate dehydrogenase complex illustrates the importance of coenzymes in metabolic reactions. Two of the coenzymes are cosubstrates (HS-CoA and $NAD^{\oplus}$), and three are prosthetic groups (TDP, lipoamide, and FAD—one

cofactor is bound to each type of subunit). The lipoamide groups bound to E_2 are primarily responsible for transferring reactants from one active site in the complex to another. A lipoamide picks up a two-carbon unit from hydroxyethyl–TDP in step 2 to form the acetyl–dihydrolipoamide intermediate. This intermediate is repositioned in the active site of dihydrolipoamide acetyltransferase where the two-carbon group is transferred to coenzyme A in step 3. The reduced lipoamide produced in that reaction is then moved to the active site of dihydrolipoamide dehydrogenase in E_3. Lipoamide is reoxidized in step 4 and the regenerated coenzyme is repositioned in the active site of E_1 where it is ready to receive a new two-carbon group. In these reactions, the lipoamide prosthetic group acts as a swinging arm that visits the three active sites in the pyruvate dehydrogenase complex (Figure 13.2). The swinging arm portion of the E_2 subunit consists of a flexible polypeptide chain that includes the lysine residue to which lipoamide is covalently bound.

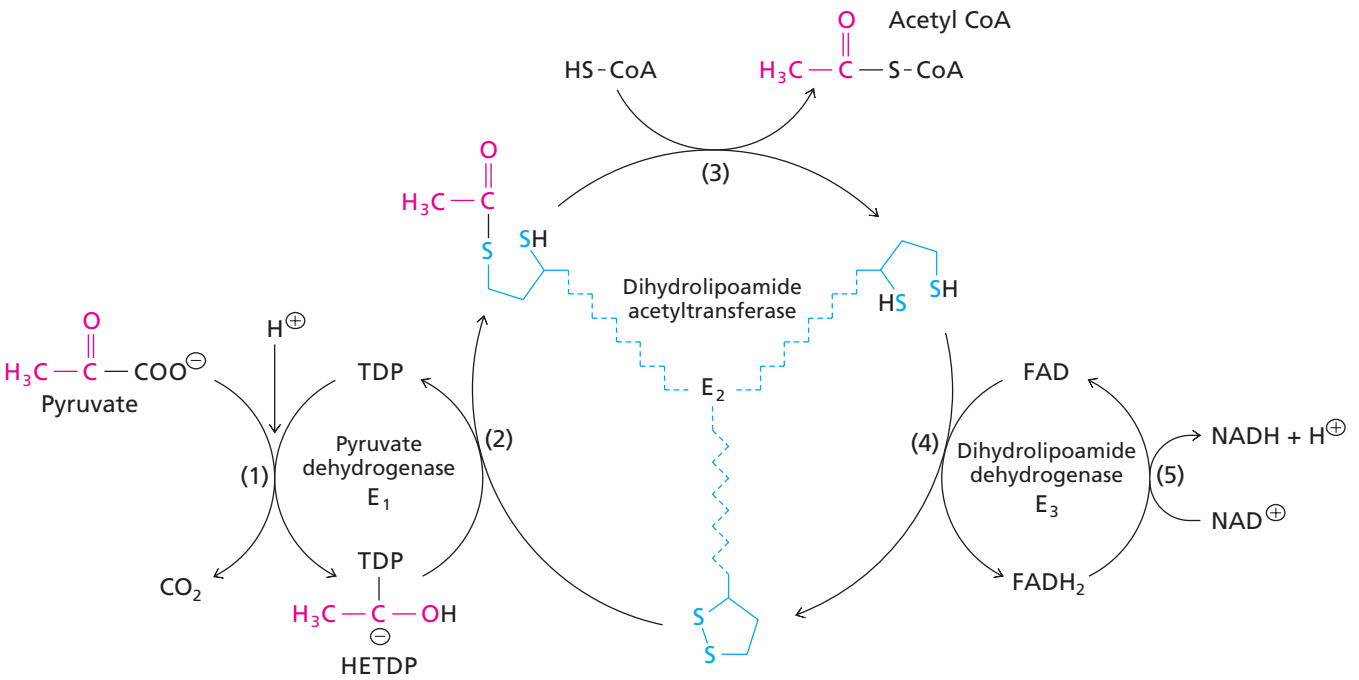
The various subunits of the complex are arranged in a way that facilitates the swinging arm mechanism of lipoamide. The mechanism ensures that the product of one reaction does not diffuse into the medium but is immediately acted on by the next component of the system. This is a form of channeling where the product of one reaction becomes the substrate of a second reaction but it differs from other examples because, in this case, the two-carbon intermediate is covalently bound to the flexible lipoamide group of E_2.

Channeling and multienzyme complexes were discussed in Section 5.11.

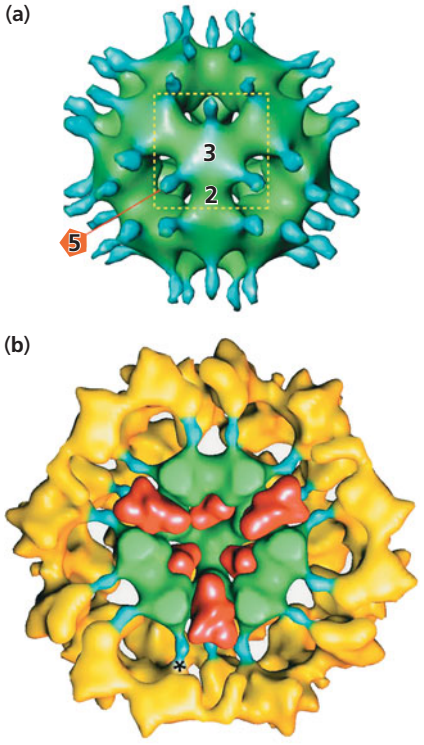
The entire pyruvate dehydrogenase reaction is a series of coupled oxidation–reduction reactions in which electrons are transported from the initial substrate (pyruvate) to the final oxidizing agent ($NAD^{\oplus}$). The four half reactions are. . . .

	$E^{\circ\prime}$
1. acetyl CoA + CO_2 + $H^{\oplus}$ + $2e^{\ominus}$ ⟶ pyruvate + CoA	–0.48
2. E_2—lipoamide + $2H^{\oplus}$ + $2e^{\ominus}$ ⟶ E_2—dihydrolipoamide	–0.29
3. E_3—FAD + $2H^{\oplus}$ + $2e^{\ominus}$ ⟶ E_3—$FADH_2$	–0.34
4. $NAD^{\oplus}$ + $2H^{\oplus}$ + $2e^{\ominus}$ ⟶ NADH + $H^{\oplus}$	–0.32

(13.7)

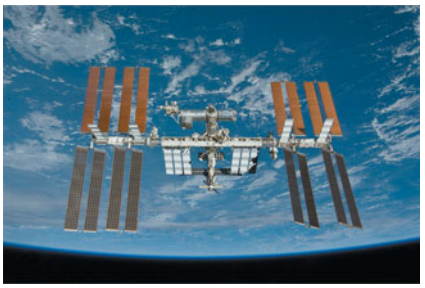


▲ **Figure 13.2**
Reactions of the pyruvate dehydrogenase complex. The lipoamide prosthetic group (blue) is attached by an amide linkage between lipoic acid and the side chain of a lysine residue of E_2. This prosthetic group is a swinging arm that carries the two-carbon unit from the pyruvate dehydrogenase active site to the dihydrolipoamide acetyltransferase active site. The arm then carries hydrogen to the dihydrolipoamide dehydrogenase active site.



▲ **Figure 13.3**
Structural model of the pyruvate dehydrogenase complex. (a) The inner core consists of 60 E_2 enzymes arranged in the shape of a pentagonal dodecahedron with one E_2 trimer at each of the 20 vertices. A single trimer is outlined by a yellow box. The center of the pentagon shape is indicated by the orange pentagon. Note the linker regions projecting upward from the surface of the core structure. **(b)** Cutaway view of the complete complex showing the outer E_1 enzymes (yellow) and the BP–E_3 enzymes (red) located in the space between the E_2 enzymes of the inner core.

[From Zhou, H. Z. et al. (2001). The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes. *Proc. Natl. Acad. Sci.* (USA) 98:14082–14087.]



▲ **A biochemistry laboratory.**

Each half-reaction has a characteristic standard reduction potential (Table 10.4) that provides some indication of the direction of electron flow. (Recall from Section 10.9 that the actual reduction potentials depend on the concentrations of reducing agents and oxidizing agents.) Electron transport begins with pyruvate, which gives up two electrons in the reverse of half-reaction 1. These electrons are taken up by E_2–lipoamide. Subsequent electron flow is from E_2–lipoamide to E_3–FAD to $NAD^{\oplus}$. The final product is NADH, which carries a pair of electrons. There are many examples of metabolic pathway enzymes with simple electron transport systems such as this one. They should not be confused with the much more complex membrane-associated electron transport system covered in the next chapter.

The pyruvate dehydrogenase complex is enormous. It is several times bigger than a ribosome. In bacteria these complexes are located in the cytosol, and in eukaryotic cells they are found in the mitochondrial matrix. Pyruvate dehydrogenase complexes are also present in chloroplasts.

The eukaryotic pyruvate dehydrogenase complex is the largest multienzyme complex known. The core of the complex is formed from 60 E_2 subunits arranged in the shape of a pentagonal dodecahedron (12 pentagons joined at their edges to form a ball). This shape has 20 vertices and each vertex is occupied by an E_2 trimer (Figure 13.3A). Each of the E_2 subunits has a linker region projecting upward from the surface. This linker contacts an outer ring of E_1 subunits that surround the inner core (Figure 13.3B). The linker region contains the lipoamide swinging arm.

The outer shell has 60 E_1 subunits. Each E_1 enzyme contacts one of the underlying E_2 enzymes and makes additional contacts with its neighbors. The E_1 enzyme consists of two α subunits and two β subunits ($\alpha_2\beta_2$), so it is considerably larger than the E_2 enzyme of the core. The E_3 enzyme (an α_2 dimer) lies in the center of the pentagon formed by the core E_2 enzymes. There are 12 E_3 enzymes in the complete complex, corresponding to the 12 pentagons in the pentagonal dodecahedron shape. In eukaryotes, the E_3 enzymes are associated with a small binding protein (BP) that's part of the complex.

The model shown in Figure 13.3 has been constructed from high resolution electron microscopy images of pyruvate dehydrogenase complexes at low temperature (cryo-EM) (Figure 13.1). In this technique, a large number of individual images are combined and a three-dimensional image is built with the help of a computer.

Sample Calculation 13.1

Q. Calculate the standard Gibbs free energy change for the pyruvate dehydrogenase reaction.

A. From Equation 10.26, the overall change in standard reduction potential is

$$\begin{aligned}\Delta E^{\circ\prime} &= \Delta E^{\circ\prime}_{\text{electron acceptor}} - \Delta E^{\circ\prime}_{\text{electron donor}} \\ &= -0.32 - (-0.48) = 0.16\ \text{V}\end{aligned}$$

from Equation 10.25,

$$\begin{aligned}\Delta G^{\circ\prime} &= -nF\Delta E^{\circ\prime} \\ &= -(2)(96.5)(0.16) \\ &= -31\ \text{kJ mol}^{-1}\end{aligned}$$

The model is then matched with the structures of any of the individual subunits that have been solved by X-ray crystallography or NMR. So far, it has not been possible to grow large crystals of the entire pyruvate dehydrogenase complex on Earth and experiments to grow crystals on the International Space Station in the absence of gravity were also unsuccessful.

A similar pyruvate dehydrogenase complex is present in many species of bacteria although some, such as gram-negative bacteria, have a smaller version where there are only 24 E_2 enzymes in the core. In these bacteria, the core enzymes are arranged as a cube with one trimer at each of the eight vertices. The E_2 subunits of the two different bacterial enzymes and the eukaryotic mitochondrial and chloroplast versions are all closely related. However, the gram-negative bacterial enzymes contain E_1 enzymes that are unrelated to the eukaryotic versions.

Pyruvate dehydrogenase is a member of a family of multienzyme complexes known as the 2-oxo acid dehydrogenase family. (Pyruvate is the smallest 2-oxo organic acid.) We will encounter two other 2-oxo (or α-keto) acid dehydrogenases that closely resemble pyruvate dehydrogenase in structure and function. One is a citric acid cycle enzyme, α-ketoglutarate dehydrogenase (Section 13.3#4), and the other is branched chain α-keto acid dehydrogenase, used in amino acid metabolism (Section 17.10E). All members of the family catalyze essentially irreversible reactions in which an organic acid is oxidized to CO_2 and a coenzyme A derivative is formed.

The reverse reactions are catalyzed in some bacteria by entirely different enzymes. These reactions form part of a pathway for fixing carbon dioxide in anaerobic bacteria. Some bacteria and some anaerobic eukaryotes convert pyruvate to acetyl CoA and CO_2 using pyruvate:ferredoxin 2-oxidoreductase, an enzyme that is unrelated to pyruvate dehydrogenase.

$$\text{Pyruvate} + \text{CoA} + 2\ \text{Fd}_{ox} \rightarrow \text{acetyl CoA} + 2\ \text{Fd}_{red} + 2\ \text{H}^{\oplus} \qquad \textbf{(13.8)}$$

The terminal electron carrier in this case is reduced ferredoxin (Fd_{red}) and not NADH, as with pyruvate dehydrogenase. The pyruvate:ferredoxin oxidoreductase reaction is reversible and may be used to fix CO_2 by reductive carboxylation. Bacterial species that have diverged very early in the history of life often contain pyruvate:ferredoxin oxidoreductase and not pyruvate dehydrogenase suggesting that the former enzyme is more primitive and pyruvate dehydrogenase evolved later.

13.2 The Citric Acid Cycle Oxidizes Acetyl CoA

Acetyl CoA formed from pyruvate or other compounds (such as fatty acids or some amino acids) can be oxidized by the citric acid cycle. The eight reactions of the citric acid cycle are listed in Table 13.1. Before examining each of the reactions individually, we should consider two general features of the pathway; the flow of carbon and the production of "high energy" molecules.

The fates of the carbon atoms are depicted in Figure 13.4. In the first reaction of the citric acid cycle, the two-carbon acetyl group of acetyl CoA is transferred to the four-carbon dicarboxylic acid oxaloacetate to form citrate, a six-carbon tricarboxylic acid. The cycle proceeds with oxidative decarboxylation of a six-carbon acid and a five-carbon acid. This releases two molecules of CO_2 and produces succinate, a four-carbon dicarboxylic acid. The remaining steps of the cycle convert succinate to oxaloacetate, the original reactant that began the cycle.

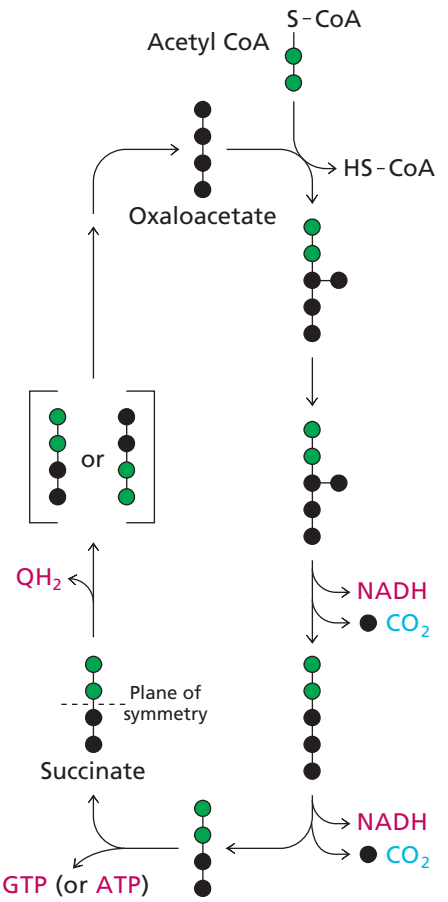
The complete reactions are shown in Figure 13.5 where the two carbons of the acetyl group are also colored green so their fate can be followed. Note that the two carbon atoms entering the cycle as the acetyl group on acetyl CoA are not the same carbon atoms that are lost as CO_2. However, the carbon balance in the overall reaction pathway is such that for each two-carbon group from acetyl CoA that enters the cycle, two carbon atoms are released during one complete turn of the cycle. The two carbon atoms of acetyl CoA become half of the symmetric four-carbon dicarboxylic acid (succinate) in the fifth step of the cycle. The two halves of this symmetric molecule are chemically equivalent so carbons arising from acetyl CoA become evenly distributed in molecules formed from succinate.

Acetyl CoA is a "high energy" molecule (Section 10.8). The thioester linkage conserves some of the energy gained from the decarboxylation of pyruvate by the pyruvate dehydrogenase complex. The net equation of the citric acid cycle (Table 13.1) tends to

The regulation of pyruvate dehydrogenase is examined in Section 13.5.

KEY CONCEPT

Large multienzyme complexes improve efficiency by channeling substrates and products.

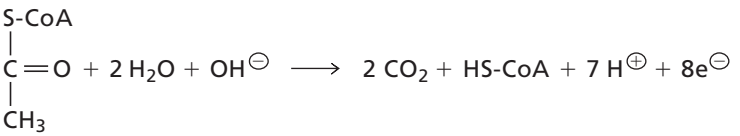


▲ **Figure 13.4**
Fates of the carbon atoms from oxaloacetate and acetyl CoA during one turn of the citric acid cycle. The plane of symmetry of succinate means that the two halves of the molecule are chemically equivalent; thus, carbon atoms from acetyl CoA (green) are uniformly distributed in the four-carbon intermediates leading to oxaloacetate. Carbon atoms from acetyl CoA that enter in one turn of the cycle are thus lost as CO_2 only in the second and subsequent turns. Energy is conserved in the reduced coenzymes NADH and QH_2 and in one GTP (or ATP) produced by substrate level phosphorylation.

Table 13.1 The enzymatic reactions of the citric acid cycle

Reaction	Enzyme
1. Acetyl CoA + Oxaloacetate + H_2O ⟶ Citrate + HS−CoA + $H^{\oplus}$	Citrate synthase
2. Citrate ⇌ Isocitrate	Aconitase (Aconitate hydratase)
3. Isocitrate + $NAD^{\oplus}$ ⟶ α-Ketoglutarate + NADH + CO_2	Isocitrate dehydrogenase
4. α-Ketoglutarate + HS-CoA + $NAD^{\oplus}$ ⟶ Succinyl CoA + NADH + CO_2	α-Ketoglutarate dehydrogenase complex
5. Succinyl CoA + GDP (or ADP) + P_i ⇌ Succinate + GTP(or ATP) + HS−CoA	Succinyl-CoA synthetase
6. Succinate + Q ⇌ Fumarate + QH_2	Succinate dehydrogenase complex
7. Fumarate + H_2O ⇌ L-Malate	Fumarase (Fumarate hydratase)
8. L-Malate + $NAD^{\oplus}$ ⇌ Oxaloacetate + NADH + $H^{\oplus}$	Malate dehydrogenase
Net equation:	
Acetyl CoA + 3 $NAD^{\oplus}$ + Q + GDP (or ADP) + P_i + 2 H_2O ⟶ HS−CoA + 3 NADH + QH_2 + GTP (or ATP) + 2 CO_2 + 2 $H^{\oplus}$	

obscure the fact that the citric acid cycle is equivalent to the oxidation of an acetyl CoA molecule with release of electrons. The overall reaction sequence can be simplified to



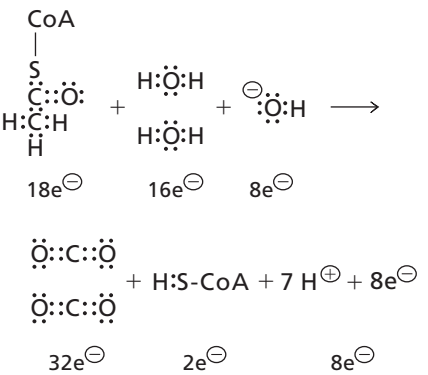
(13.9)

where the hydroxyl group is donated by inorganic phosphate in Reaction 5 and some of the products are shown as free protons and free electrons. This form of the net equation reveals that eight electrons are released during the oxidation. (Recall that oxidation reactions release electrons and reduction reactions take up electrons.) Six of the electrons are transferred to three molecules of $NAD^{\oplus}$ along with three of the protons depicted in Reaction 13.9. The remaining two electrons are transferred to one molecule of ubiquinone (Q) along with two of the protons. Two free protons are produced in each turn of the cycle. (Keep in mind that the carbon dioxide molecules released during the citric acid cycle do not actually come directly from acetyl CoA. Reaction 13.9 is a simplified version that emphasizes the net oxidation.)

BOX 13.2 WHERE DO THE ELECTRONS COME FROM?

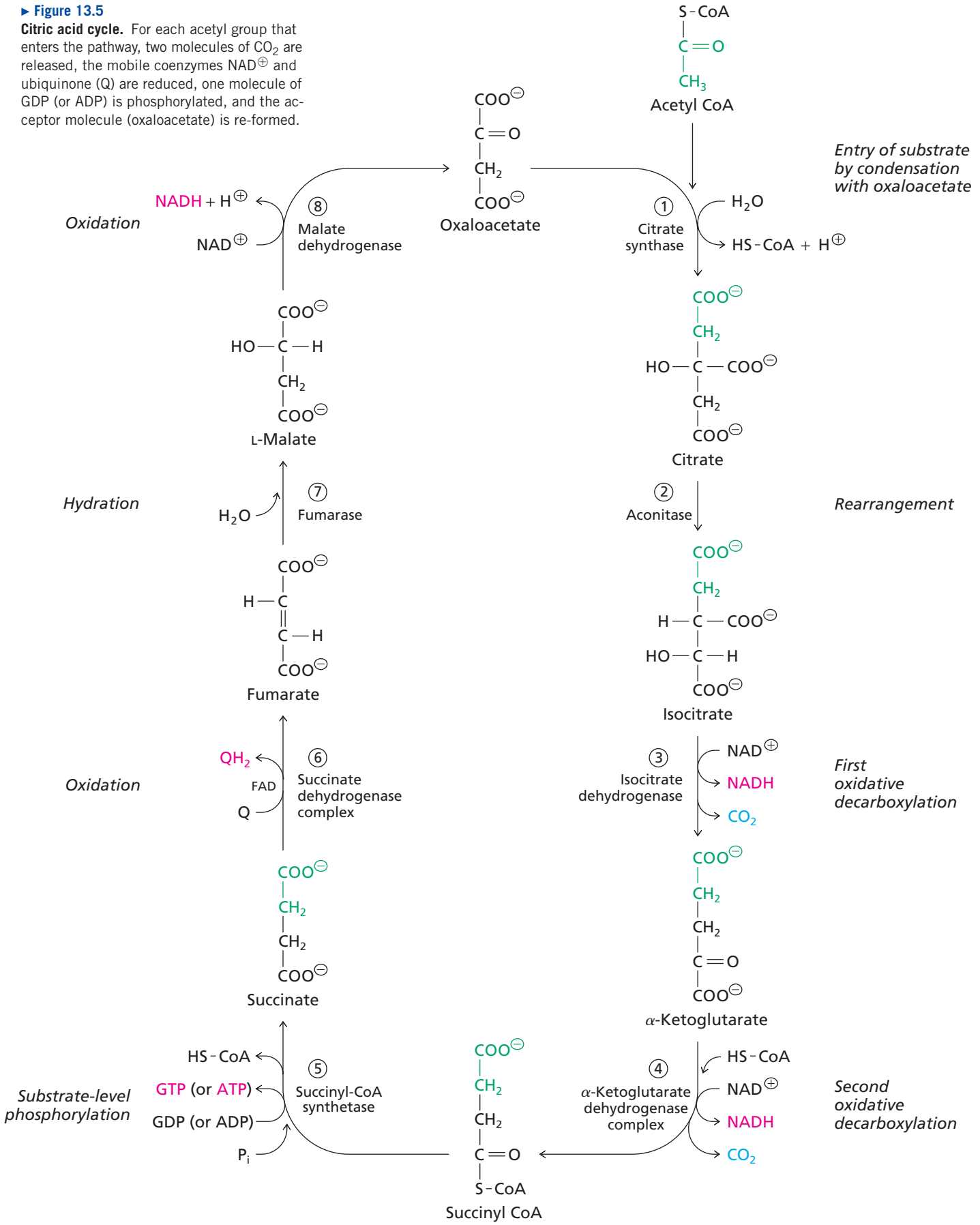
Chemical reaction equations, such as Reaction 13.9, aren't very helpful in understanding where electrons are released and taken up. In order to see the electron balance in such reactions it's often useful to redraw the structures with the valence electrons replacing the lines that represent the chemical bonds in most drawings. Each covalent bond involves a shared pair of electrons and each of the standard atoms (C, O, N, S) requires eight valence electrons. Covalently bonded hydrogen atoms have only a single pair of electrons in their single shell.

The oxidation of acetyl CoA from Equation 13.8 is shown in this form in the figure. Note that only the electrons in the outer shells of the atoms are shown. These are the ones removed by oxidations or added in reduction reactions. There are 42 electrons (21 pairs) in the reactants and 34 electrons (17 pairs) in the products: CO_2 and Coenzyme A. Thus, 8 electrons are released in the oxidation. Most of the time, electrons are released when double bonds are formed (as in carbon dioxide) since this results in the sharing of an extra electron pair.



▲ **The oxidation of an acetyl CoA equivalent by the citric acid cycle showing the valence electrons in the reactants and products.**

▶ **Figure 13.5**
Citric acid cycle. For each acetyl group that enters the pathway, two molecules of CO_2 are released, the mobile coenzymes $NAD^{\oplus}$ and ubiquinone (Q) are reduced, one molecule of GDP (or ADP) is phosphorylated, and the acceptor molecule (oxaloacetate) is re-formed.



Most of the energy released in the citric acid cycle reactions is conserved in the form of electrons transferred from organic acids to generate the reduced coenzymes NADH and QH_2 (Figure 13.5). NADH is formed by the reduction of $NAD^{\oplus}$ at three oxidation–reduction steps—two of these are oxidative decarboxylations. QH_2 is formed when succinate is oxidized to fumarate. Subsequent oxidation of the reduced coenzymes by membrane-associated electron transport leads to the transfer of electrons from NADH and QH_2 to a terminal electron acceptor. In the case of most eukaryotes (and many prokaryotes), this terminal electron acceptor is oxygen, which is reduced to water. Membrane-associated electron transport is coupled to the production of ATP from ADP and P_i. The entire process (electron transport + phosphorylation of ADP) is often referred to as oxidative phosphorylation when oxygen is present (Chapter 14). In addition to the formation of reducing equivalents, the citric acid cycle produces a nucleotide triphosphate directly by substrate level phosphorylation. The product can be either ATP or GTP, depending on the cell type or species.

KEY CONCEPT

The citric acid cycle is a mechanism for the oxidation of the acetyl group of acetyl CoA.

13.3 The Citric Acid Cycle Enzymes

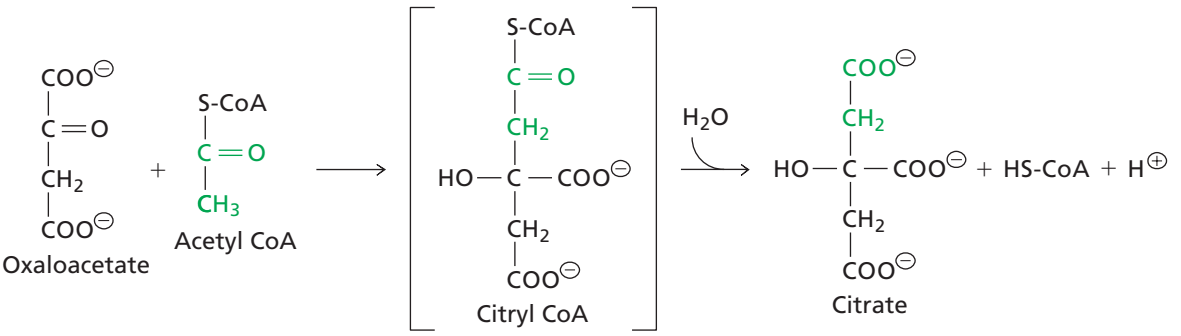
The citric acid cycle can be viewed as a multistep catalytic reaction returning to its original state after an acetyl CoA molecule is oxidized. This view is based on the fact that when the reactions operate as a cycle the original reactant, oxaloacetate, is regenerated. By definition, a catalyst increases the rate of a reaction without itself undergoing net transformation. All enzymatic reactions, in fact all catalytic reactions, can be represented as cycles. An enzyme goes through a cyclic series of conversions and finally returns to the form in which it began. In this sense, the citric acid cycle fits the description of a catalyst.

Taken as a whole, the citric acid cycle is a mechanism for oxidizing the acetyl group of acetyl CoA to CO_2 by $NAD^{\oplus}$ and ubiquinone. When the citric acid cycle operates in isolation its intermediates are re-formed with each full turn of the cycle. As a result, the citric acid cycle doesn't appear to be a pathway for net synthesis or degradation of any of the intermediates in the pathway unlike, for example, the gluconeogenesis pathway or the glycolysis pathway. However, we will see later on (Section 13.6) that the citric acid pathway doesn't always operate in isolation and appearances can be deceiving. Some of the intermediates are shared with other pathways. Let's first examine the catalytic aspect of the citric acid cycle by examining each of the eight enzymatic steps.

1. Citrate Synthase

In the first reaction of the citric acid cycle, acetyl CoA reacts with oxaloacetate and water to form citrate, HS-CoA, and a proton. This reaction is catalyzed by citrate synthase and results in the formation of an enzyme-bound intermediate called citryl CoA (Figure 13.6).

Citrate is the first of two tricarboxylic acids that are part of the cycle. The standard Gibbs free energy change for the citrate synthase reaction is $-31.5\ kJ{\cdot}mol^{-1}$ ($\Delta G^{\circ\prime} = -31.5\ kJ{\cdot}mol^{-1}$) due to the hydrolysis of the high energy thioester bond in the citryl CoA intermediate. Normally you might expect that such a large negative Gibbs free energy change



▼ Figure 13.6
Reaction catalyzed by citrate synthase. In the first step, acetyl CoA combines with oxaloacetate to form an enzyme-bound intermediate, citryl CoA. The thioester is hydrolyzed to release the products, citrate and HS-CoA.

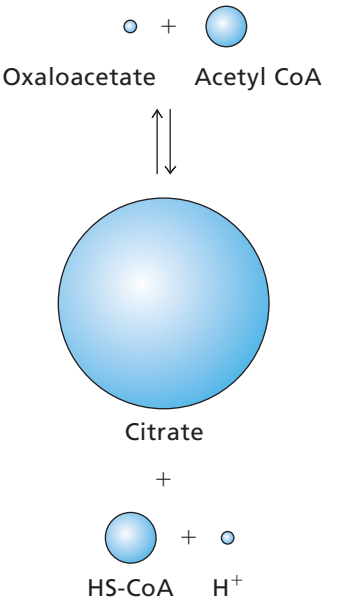
would be coupled to synthesis of ATP—keeping in mind that the actual Gibbs free energy change inside the cell might be very different. Indeed, the hydrolysis of the similar thioester bond in succinyl CoA (Reaction 5 of the citric acid cycle) is coupled to synthesis of GTP (or ATP). However, in the case of the citrate synthase reaction, the available energy is used for a different purpose. It ensures that the reaction proceeds in the direction of citrate synthesis when the concentration of oxaloacetate is very low (Figure 13.7). This appears to be the normal situation when the citric acid cycle is operating. In the presence of only small (catalytic) amounts of oxaloacetate the equilibrium of the reaction depicted in Figure 13.6 still favors citrate synthesis. In other words, the actual Gibbs free energy change inside the cell is close to zero. The reaction is a near-equilibrium reaction. The thermodynamics ensures that the citric acid cycle operates in the direction of acetyl CoA oxidation even under conditions where the concentration of oxaloacetate is very low.

Citrate synthase is a transferase—one of the six categories of enzymes described in Section 5.1. Transferases catalyze transfer reactions, in this case transfer of an acetyl group. The term "synthase" is used for transferases that do not use ATP as a cofactor. "Synthetases," on the other hand, are members of the ligase category of enzymes (Section 5.1). The reactions catalyzed by synthetases must be coupled to ATP (or GTP) hydrolysis. It's important to remember the difference between synthases and synthetases since the words look very similar and since the citric acid cycle contains an example of each type of enzyme. (For some reason, it's easier to pronounce "synthetase" and it's tempting to throw in the extra syllable when you should be saying "synthase.")

In gram-positive bacteria, archaebacteria, and eukaryotes, citrate synthase is a dimeric protein composed of two identical subunits. In gram-negative bacteria, the enzymes are hexameric complexes of identical subunits.

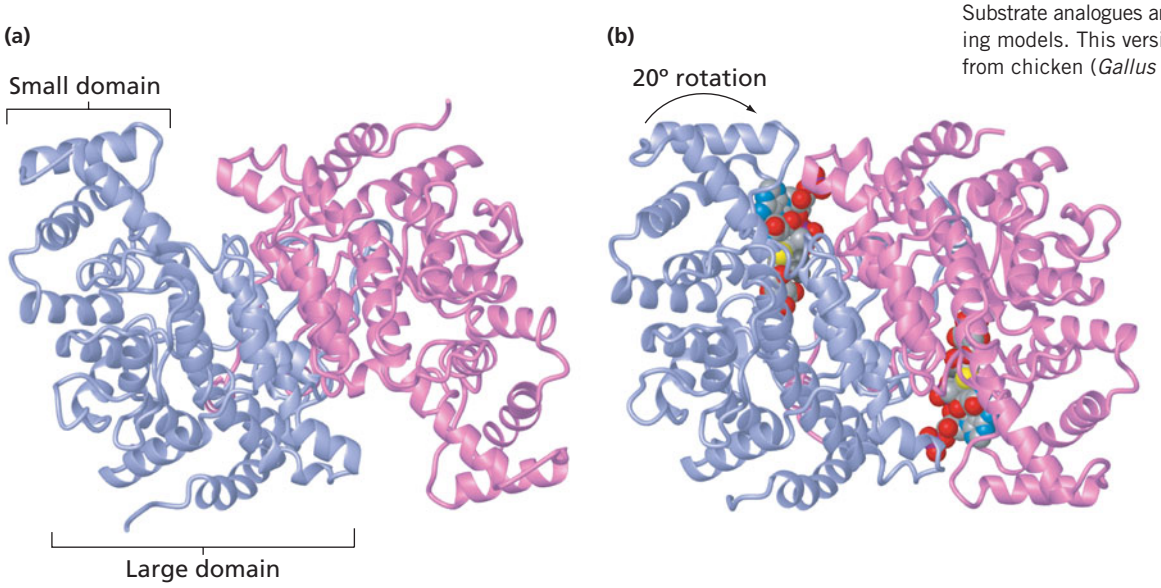
In animals each subunit of the enzyme has two distinct domains: a small flexible domain on the outer surface and a larger domain that forms the core of the protein (Figure 13.8). The two subunits associate by interactions between four α helices in each of the large domains to form an α helix sandwich. Citrate synthase undergoes a large conformational change on binding oxaloacetate as shown in Figure 13.8. The binding site lies at the base of a deep cleft between the small domain of one subunit and the large domain of the other subunit. When oxaloacetate is bound, the small domain rotates by 20° relative to the large domain. This closure creates the binding site for acetyl CoA—a site which is formed by amino acid side chains from both large and small domains. When the reaction is complete coenzyme A is released. The enzyme then reverts to the open conformation when citrate is released.

The structure of the enzyme requires that oxaloacetate and acetyl CoA bind sequentially. This reduces the chance of binding acetyl CoA in the absence of oxaloacetate and



▲ **Figure 13.7**
Representation of the relative ratios of products and reactants in the citrate synthase reaction. The equilibrium constant (K_{eq}) for the citrate synthase reaction can be calculated from standard Gibbs free energy change according to Equation 1.12, $K_{eq} = 2.7 \times 10^5$, meaning that, at equilibrium, the concentrations of products are more than 200,000 times that of the reactants. [Not to scale.]

▼ **Figure 13.8**
Citrate synthase induced fit mechanism. The two identical subunits are colored blue and purple. Each is composed of a small and a large domain. **(a)** Open conformation. The substrate binding site is located in the deep cleft between the small domain of one subunit and the large domain of the other. [PDB 5CSC] **(b)** Closed conformation. The small domain has shifted relative to the large domain in order to close off the large binding cleft seen in the open conformation. Substrate analogues are shown as space-filling models. This version of the enzyme is from chicken (*Gallus gallus*). [PDB 6CSC]



BOX 13.3 CITRIC ACID

The discovery of citric acid is usually attributed to Abu Musa Jābir ibn Hayyān (~721–~815), known as Geber in Europe. He worked in Kufa in modern-day Iraq and is recognized as the father of modern chemistry. Jābir identified citric acid as a major component of citrus fruits such as lemons and limes. We know now that the level of citric acid in these fruits is related to its ability to act as a preservative and a reservoir of carbon. This is unrelated to the role of citrate in the citric acid cycle.

Citric acid is a weak organic acid (pK_{a1}= 3.2, pK_{a2}= 4.8, pK_{a3}= 6.4). The sodium salt is sometimes used as a buffer in biochemistry labs and in drugs but its most important application is as a food additive, especially in soft drinks.

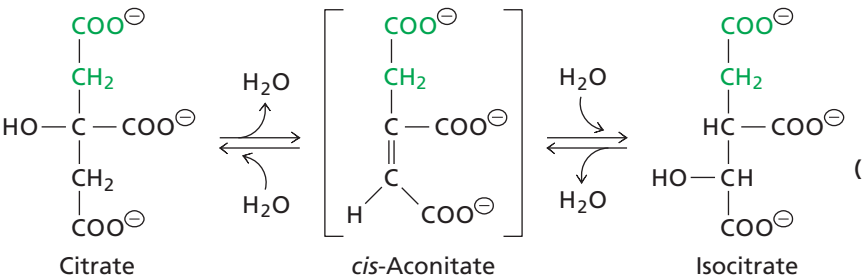


▲ Citric acid is an important natural preservative in citrus fruits.

the possibility of catalyzing hydrolysis of the thioester bond of acetyl CoA in a wasteful reaction. This potential side reaction is a very real danger since the thioester bond of acetyl CoA is near the active site for hydrolysis of the citryl CoA thioester and since the concentration of oxaloacetate may be very low relative to that of acetyl CoA. Our previous examples of an induced fit mechanism involved protecting ATP from inappropriate hydrolysis but the same principle applies here. We will encounter several other examples of important structure–function relationships in this chapter and the next one.

2. Aconitase

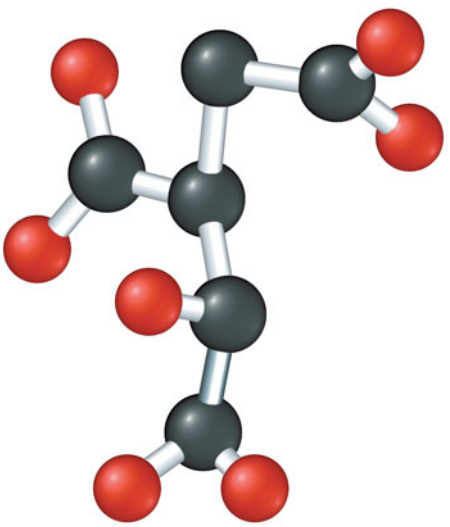
Aconitase (systematic name: aconitate hydratase) catalyzes a near-equilibrium conversion of citrate to isocitrate. Citrate is a tertiary alcohol and thus cannot be oxidized directly to a keto acid. The formation of a keto acid intermediate is required for the oxidative decarboxylation reaction that occurs in step 3 of the citric acid cycle. The step catalyzed by aconitase creates a secondary alcohol in preparation for step 3. The name of the enzyme is derived from *cis*-aconitate, an enzyme-bound intermediate of the reaction. The reaction proceeds by the elimination of water from citrate to form a carbon–carbon double bond. This is followed by stereospecifc addition of water to form isocitrate.



(13.10)

KEY CONCEPT

Stereospecific reactions occur because substrates bind to enzymes in specific orientations.



▲ Figure 13.9
Structure of 2R,3S-isocitrate.

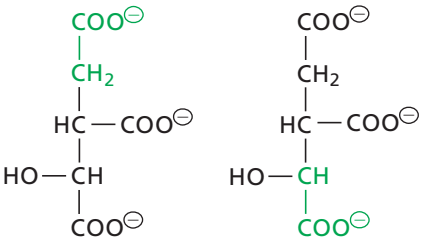
The aconitase gene is a member of a complex gene family. The family encodes distinct mitochondrial and cytoplasmic versions of aconitase, a regulatory protein with no catalytic activity, and an enzyme involved in the synthesis of amino acids (Sections 13.8 and 17.3C). Bacteria contain two distantly related enzymes, aconitase A and aconitase B. All family members contain a characteristic [4 Fe–4 S] iron–sulfur cluster. In the next chapter we will encounter many oxidation–reduction enzymes with iron–sulfur clusters. In most of these oxidation–reduction enzymes, the iron–sulfur clusters participate in electron transport but members of the aconitase family are unusual because the role of the iron–sulfur cluster is to aid in the binding of citrate. The aconitase reaction is an isomerization reaction and not an oxidation-reduction reaction.

Note that citrate is not a chiral molecule because none of the carbon atoms is bonded to four different groups. However, the product of the reaction, isocitrate, has two chiral centers, C2 and C3. Each of these carbon atoms has four different constituents

BOX 13.4 THREE POINT ATTACHMENT OF PROCHIRAL SUBSTRATES TO ENZYMES

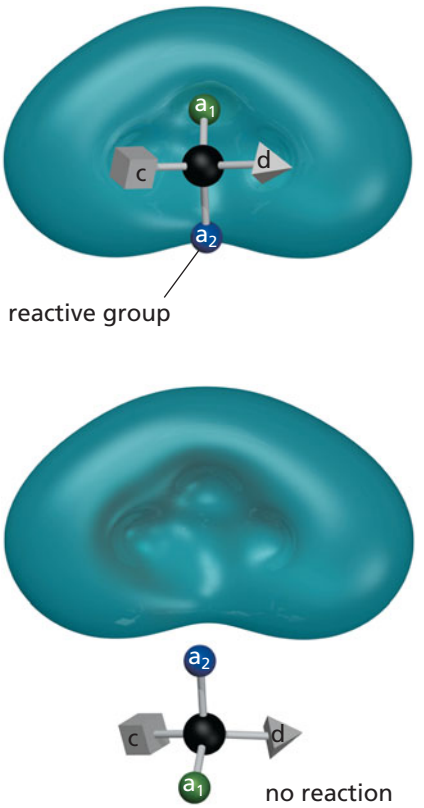
When the citric acid cycle was first proposed by Krebs, the inclusion of the citrate-to-isocitrate reaction was a major barrier to its acceptance because labeling studies indicated that only one of the two possible forms of 2R,3S-isocitrate was produced in cells. The "problem" was not that a chiral molecule was produced from a non-chiral molecule—this is easily understood. The difficulty was in understanding why formation of the double bond of *cis*-aconitate, and subsequent addition of water to form isocitrate, occurred only in the moiety contributed originally by oxaloacetate and not in the group derived from acetyl CoA. When isotopically labeled acetate was added to cells the ^{14}C-labeled carbon atoms appeared in citrate as shown in green in Reaction 13.10. Since citrate is a symmetric molecule, the labeled carbon atoms were expected to show up equally in the two versions of isocitrate shown in the figure on the right.

Instead, only the left-hand form was produced. At the time, conversion of a non-chiral molecule to a single form of chiral isomer was unknown but in 1948, Alexander Ogston showed how the active site of an enzyme could distinguish between chemically equivalent groups on the citrate molecule. Ogston envisioned citrate binding in a manner he called three point attachment, with nonidentical groups involved in the enzyme–substrate binding (see figure). Once citrate is correctly bound to the asymmetric binding site, the two —CH_2—$COO^{\ominus}$ groups of citrate have specific orientations and thus are no longer equivalent. Formation of the carbon–carbon double bond can only take place in the group contributed by oxaloacetate.



▲ **Two forms of isocitrate.** The green carbon atoms represent the group originally derived from acetyl CoA. The reaction catalyzed by aconitase was expected to yield two forms of isocitrate in equal quantities because the substrate (citrate) is symmetric. Only the left-hand form was produced.

Citrate is a *prochiral* molecule because it can react asymmetrically in spite of the fact that it is chemically symmetric. There are now many examples of such reactions in metabolic pathways.



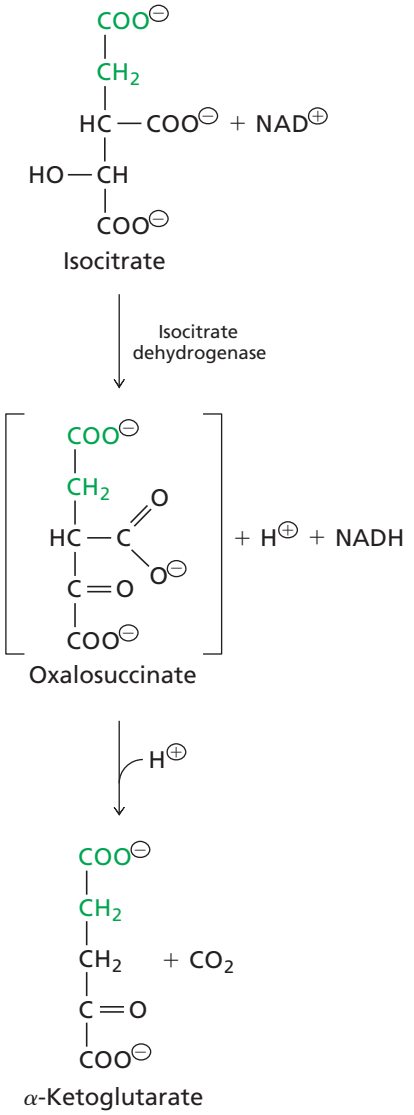
▲ **Binding of citrate to the active site of aconitase.** The central carbon atom of the citrate molecule is shown with four attached groups: the hydroxyl group (—OH) is represented by a square; the carboxyl group (—COOH) by a triangle; the two —CH2—COO— groups are shown as spheres. The two —CH2—COO— groups are chemically indistinguishable, but the one derived from acetyl CoA is shown as a green sphere and the one derived from oxaolacetate is colored blue. A cartoon of aconitase is depicted as an asymmetric molecule with three-point attachments sites for the hydroxyl group, the carboxyl group, and one of the —CH2—COO— groups. When citrate is oriented as shown in the top figure, it can bind to aconitase and the reaction takes place in the moiety derived from oxaloacetate. The other orientation (bottom) cannot bind to the enzyme and the reaction cannot take place in the group derived from acetyl CoA.

and in each case the four groups can be arranged in two different orientations. There are four different stereoisomers of isocitrate but only one of them is produced in the reaction catalyzed by aconitase. The formal name of this product is 2R,3S-isocitrate (Figure 13.9) using the RS nomenclature described in Box 3.2. This is one of the few times when this nomenclature is useful in introductory biochemistry.

3. Isocitrate Dehydrogenase

Isocitrate dehydrogenase catalyzes the oxidative decarboxylation of isocitrate to form α-ketoglutarate (Figure 13.10). This reaction is the first of four oxidation–reduction reactions in the citric acid cycle. The reaction is coupled to the reduction of $NAD^{\oplus}$ and occurs in two steps involving an enzyme-bound oxalosuccinate intermediate.

The regulation of isocitrate dehydrogenase in prokaryotes is described in Section 13.8.



▲ **Figure 13.10**
Isocitrate dehydrogenase reaction. The enzyme catalyzes an oxidation–reduction reaction using $NAD^{\oplus}$ as the electron acceptor. Oxalosuccinate is an unstable intermediate that is rapidly decarboxylated to CO_2 and α-ketoglutarate. This is the first decarboxylation step in the citric acid cycle.

KEY CONCEPT

The important "pay off" reactions of the citric acid cycle are those that produce reducing equivalents such as NADH and QH_2.

In the first step, the alcohol group of isocitrate is oxidized by removal of two hydrogens to form a —C═O double bond. This is a typical dehydrogenase reaction. One of the hydrogens (the one bound to the carbon atom) is transferred to $NAD^{\oplus}$ as a hydride ion carrying two electrons and the other (the one on the —OH group) is incorporated into the final product. This is the first of the reactions that result in the loss of electrons (i.e., oxidation of an organic acid).

Oxalosuccinate, an unstable keto acid, is the product of the first step in the overall reaction catalyzed by α-ketoglutarate dehydrogenase. Before it is released from the enzyme, the intermediate undergoes decarboxylation to form α-ketoglutarate in the second step of the reaction. The decarboxylation reaction is associated with the release of CO_2 and uptake of a proton. The overall stoichiometry of the reaction is

$$\text{Isocitrate} + NAD^{\oplus} \longrightarrow \alpha\text{-Ketoglutarate} + NADH + CO_2 \qquad \textbf{(13.11)}$$

There are several different versions of isocitrate dehydrogenase. Bacteria contain both an $NAD^{\oplus}$-dependent enzyme and an $NADP^{\oplus}$-dependent enzyme. Eukaryotes also have both types but, in addition, the $NADP^{\oplus}$-dependent enzymes form several subclasses. In general, the $NAD^{\oplus}$-dependent enzyme is localized to the mitochondria and plays the major role in the citric acid cycle. The $NADP^{\oplus}$-dependent enzymes are found in the cytoplasm, chloroplasts, and other membrane compartments. All forms of the enzymes are homologous by sequence similarity and they share a common ancestor with an enzyme in the leucine biosynthesis pathway (Section 13.9, Section 17.3C).

4. The α-Ketoglutarate Dehydrogenase Complex

Oxidative decarboxylation of α-ketoglutarate is analogous to the reaction catalyzed by pyruvate dehydrogenase. In both cases, the reactants are an α-keto acid and HS-CoA and the products are CO_2 and a "high energy" thioester compound. Step 4 of the citric acid cycle is catalyzed by α-ketoglutarate dehydrogenase (also known as 2-oxoglutarate dehydrogenase) (Figure 13.11)

α-Ketoglutarate dehydrogenase is a large complex that resembles pyruvate dehydrogenase in both structure and function. The same coenzymes are involved and the reaction mechanism is the same. The three component enzymes of the α-ketoglutarate dehydrogenase complex are α-ketoglutarate dehydrogenase (E_1, containing TDP), dihydrolipoamide succinyl transferase (E_2, containing a lipoamide swinging arm), and dihydrolipoamide dehydrogenase (E_3, the same flavoprotein found in the pyruvate dehydrogenase complex). The overall reaction is the second of the two CO_2 producing reactions in the citric acid cycle and the second reaction that generates reducing equivalents. In the four remaining reactions of the cycle, the four-carbon succinyl group of succinyl CoA is converted back to oxaloacetate.

Eukaryotic cells have a single mitochondrial α-ketoglutarate dehydrogenase. Archaebacteria, and some other species of bacteria, do not have α-ketoglutarate dehydrogenase. Instead, they convert α-ketoglutarate to succinyl CoA using an entirely different enzyme called 2-oxoglutarate:ferredoxin oxidoreductase.

5. Succinyl CoA Synthetase

The conversion of succinyl CoA to succinate is catalyzed by succinyl CoA synthetase, sometimes called succinate thiokinase. The reaction couples hydrolysis of the thioester linkage in succinyl CoA to formation of a nucleoside triphosphate—either GTP or ATP, depending on the species. The complicated IUPAC names of these two related enzymes are: succinate-CoA ligase, ADP-forming (E.C. 6.2.1.5); and succinate-CoA ligase, GDP-forming (E.C. 6.2.1.4).

Inorganic phosphate is one of the reactants and the reaction takes place in three steps (Figure 13.12).

The first step generates succinyl phosphate as an intermediate and releases coenzyme A. In the second step, the phosphoryl group is transferred to a histidine side chain in the active site of the enzyme to form a stable phosphoenzyme intermediate. The third step transfers the phosphoryl group to GDP to form GTP. This reaction is the only example of *substrate level phosphorylation* in the citric acid cycle. (Recall from Section 10.8

that the standard Gibbs free energy change for hydrolysis of the thioester linkage in succinyl CoA is approximately equivalent to that of ATP hydrolysis.) The overall stoichiometry of the succinyl CoA synthetase reaction is

$$\text{Succinyl CoA} + P_i + \text{GDP} \longrightarrow \text{Succinate} + \text{HS-CoA} + \text{GTP} \qquad \textbf{(13.12)}$$

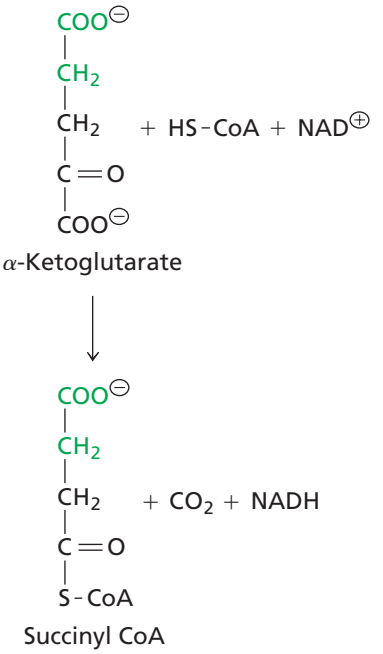
Inorganic phosphate contributes the phosphoryl group to GDP, plus an oxygen to form succinate and a hydrogen to form HS-CoA. Note that the enzyme is named for the reverse reaction where succinyl CoA is synthesized from succinate at the expense of GTP or ATP. It is called a synthetase because the reaction combines two molecules and it is coupled to hydrolysis of nucleoside triphosphate.

The enzyme is composed of two α and two β subunits ($\alpha_2\beta_2$). The β subunits contain the binding site for the nucleotide. Bacterial versions use ATP while animals often have two versions of the enzyme—one that uses GTP and one that uses ATP. They differ in their β subunits. The GTP-dependent versions clearly have evolved from the ATP-dependent versions. It's not clear why animal mitochondria have two versions of succinyl CoA synthetase in their mitochondria but one possibility is that the ATP-dependent version is used in the citric acid cycle and the GTP-dependent version primarily catalyzes the reverse reaction in some cells. Archaebacteria, and some other bacteria, do not have succinyl CoA synthetase. They carry out a similar reaction using an entirely different enzyme.

6. Succinate Dehydrogenase Complex

Succinate dehydrogenase catalyzes the oxidation of succinate to fumarate forming a carbon–carbon double bond with the loss of two protons and two electrons (Figure 13.13). The protons and electrons are passed to a quinone, which is reduced to QH_2. (Ubiquinone is the preferred substrate in almost all cases but some bacteria use menaquinone.) The enzyme is present in all species and FAD is an essential bound cofactor.

One important feature of this reaction is the scrambling of the original acetyl carbon atoms. They can no longer be specifically identified (i.e., green) in the symmetrical



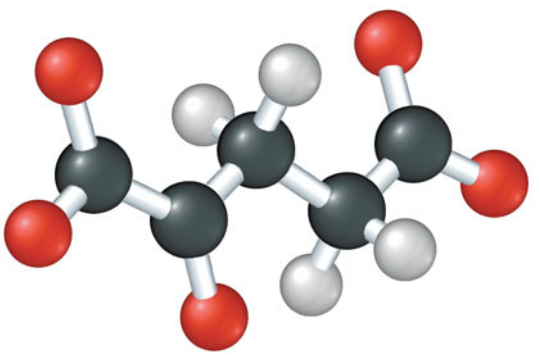
▲ **Figure 13.11**
Reaction catalyzed by α-ketoglutarate dehydrogenase. This is similar to the reaction catalyzed by pyruvate dehydrogenase.

The structure of menaquinone is shown in Figure 14.21.

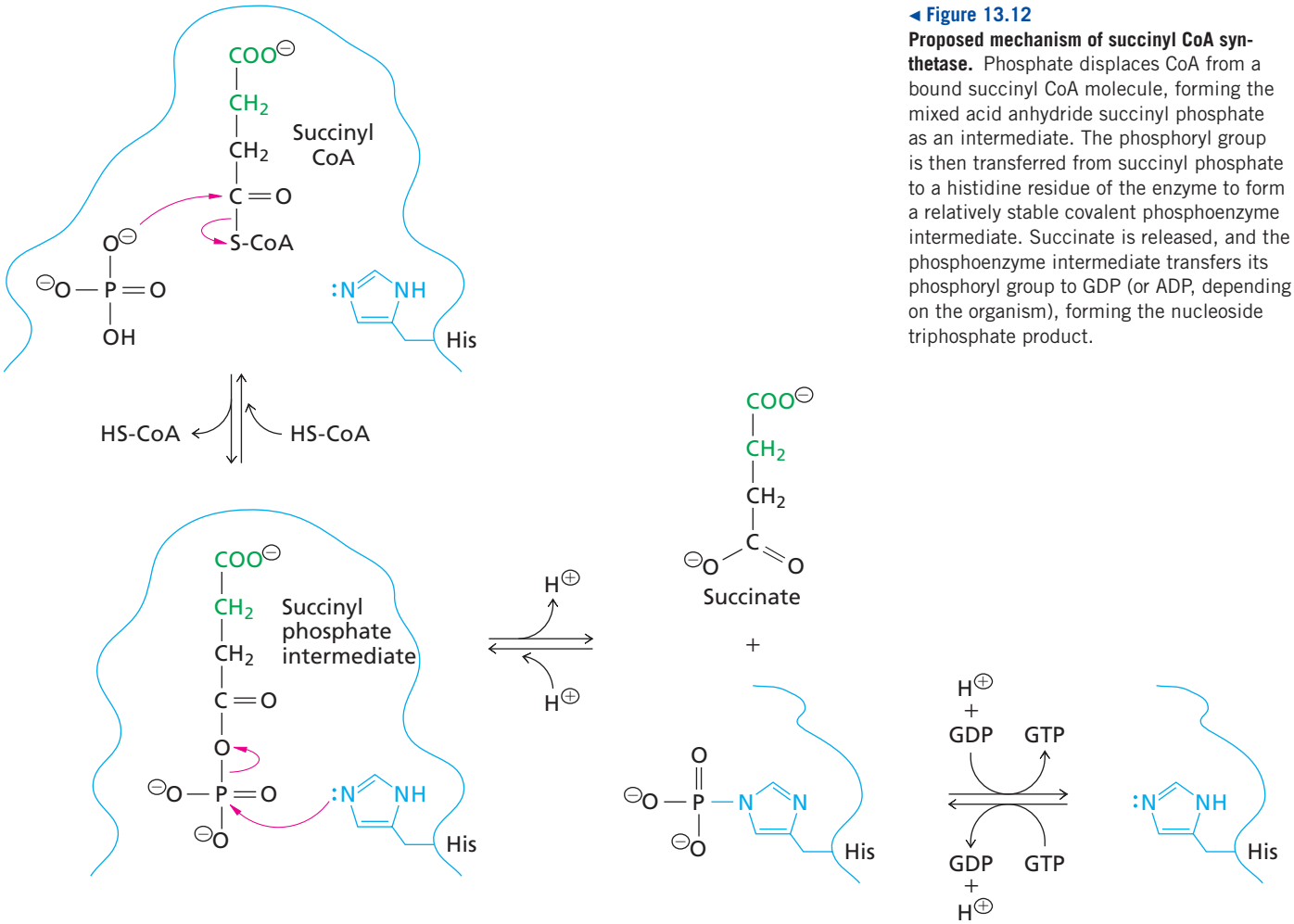
BOX 13.5 WHAT'S IN A NAME?

α-Ketoglutarate is clearly named after the five-carbon dicarboxylic acid glutarate ($^{\ominus}OOC—CH_2—CH_2—CH_2—COO^{\ominus}$). The keto group is on the α carbon or the first carbon after one of the carboxyl groups. This naming convention is similar to the one we encountered in naming α-amino acids (Section 3.1). As is the case with amino acids, the correct chemical name, or systematic name, for α-ketoglutarate could be "2-ketoglutarate." However, the formal name is actually 2-oxoglutarate since according to the IUPAC/IUBMB rules of nomenclature the term "keto" should now be avoided.

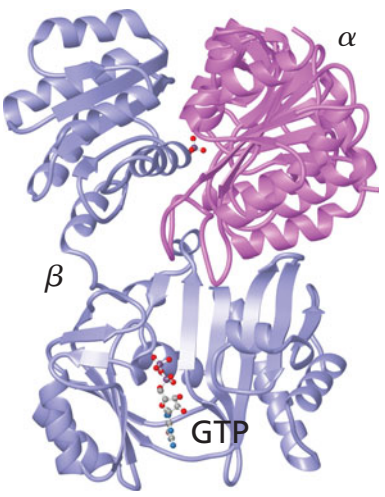
It is perfectly acceptable to refer to organic molecules by their common (trivial) names if these common names are well known. For example, if you look back to step 1 of the citric acid cycle you can see that the systematic name for oxaloacetate is 2-oxosuccinate since it is a derivative of the four-carbon dicarboxylic acid, succinate. "Oxaloacetate" is the well-known and accepted common name for this compound and it would be confusing to use any other name. When it comes to the correct name for α-ketoglutarate, the situation is more complicated because α-ketoglutarate is the old-fashioned systematic name of the molecule and the new rules say that the systematic name should be 2-oxoglutarate. The new name is becoming more and more popular in the scientific literature. Here, we continue to use the well-known name α-ketoglutarate on the grounds that it has become an acceptable common name for this compound. It's very likely that this will change in future editions.



► 2-oxoglutarate, aka α-ketoglutarate.



◂ Figure 13.12
Proposed mechanism of succinyl CoA synthetase. Phosphate displaces CoA from a bound succinyl CoA molecule, forming the mixed acid anhydride succinyl phosphate as an intermediate. The phosphoryl group is then transferred from succinyl phosphate to a histidine residue of the enzyme to form a relatively stable covalent phosphoenzyme intermediate. Succinate is released, and the phosphoenzyme intermediate transfers its phosphoryl group to GDP (or ADP, depending on the organism), forming the nucleoside triphosphate product.



▲ GTP-dependent succinyl CoA synthetase. The structure of one unit of the dimer is shown with the α and β subunits in different colors. A molecule of GTP is bound at the active site within the β subunit. This is the pig (*Sus scrofa*) version of the enzyme. [PDB 2FPG]

reactant, succinate, or in the product, fumarate. This has interesting consequences (see Problem #6).

The active site of the enzyme is formed from two different subunits. One subunit contains iron–sulfur clusters and the other is a flavoprotein with covalently bound FAD. The succinate dehydrogenase dimer is bound to two membrane polypeptides to form a larger complex. The membrane components consist of a cytochrome *b*, with its associated heme group, and a quinone binding site. The electron transport cofactors participate in the transfer of electrons from succinate to FAD to several iron–sulfur clusters to heme to the quinone.

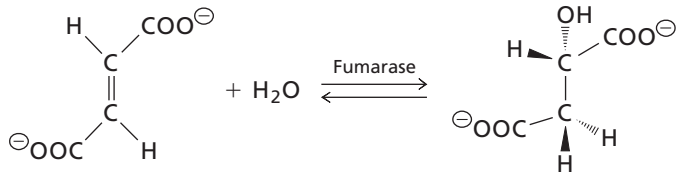
Recall that $FADH_2$ in subunit E_3 of pyruvate dehydrogenase is reoxidized by $NAD^{\oplus}$ to complete the catalytic cycle of that enzyme. In the succinate dehydrogenase reaction, $FADH_2$ is reoxidized by Q to regenerate FAD. In the past, it was very common to show $FADH_2$ as the redox product of this reaction but since FAD is covalently bound to the enzyme, the catalytic cycle is not completed until bound $FADH_2$ is reoxidized and the mobile product QH_2 is released.

The succinate dehydrogenase reaction is unusual for a dehydrogenase because it uses ubiquinone as an electron acceptor (oxidizing agent) instead of $NAD^{\oplus}$. It is also unusual in many other ways, as we will see in the next chapter. The succinate dehydrogenase complex is part of the electron transport system located in the plasma membrane of prokaryotes and in the inner mitochondrial membrane in eukaryotic cells. We will discuss this enzyme in more detail in Section 14.6 and examine its structure (Figure 14.9). In bacteria, the bulk of the enzyme complex projects into the cytoplasm where it can bind succinate and release fumarate as part of the citric acid cycle. In mitochondria, the active site is on the matrix side of the membrane where the other citric acid cycle enzymes are located.

The substrate analog malonate is a competitive inhibitor of the succinate dehydrogenase complex as described in Section 5.7A. Malonate, like succinate, is a dicarboxylate that binds to cationic amino acid residues in the active site of the succinate dehydrogenase complex. However, malonate cannot undergo oxidation because it lacks the —CH_2—CH_2— group necessary for dehydrogenation. In experiments with isolated mitochondria or cell homogenates, the presence of malonate caused succinate, α-ketoglutarate, and citrate to accumulate. Such experiments provided some of the original evidence for the sequence of reactions in the citric acid cycle.

7. Fumarase

Fumarase (systematic name: fumarate hydratase) catalyzes the near-equilibrium conversion of fumarate to malate through the stereospecific *trans* addition of water to the double bond of fumarate.



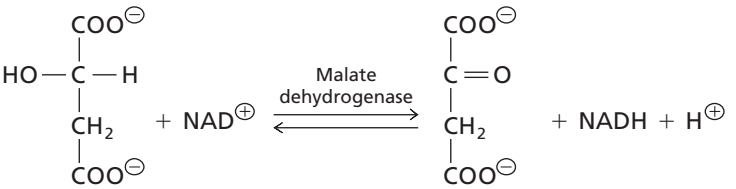
Fumarate L-Malate (13.13)

Fumarate is a prochiral molecule. When fumarate is positioned in the active site of fumarase, the double bond of the substrate can be attacked from only one direction. The product of the reaction is exclusively the L stereoisomer of the hydroxy acid malate.

There are two unrelated fumarases that can catalyze the same reaction. The class I enzyme is found in most bacteria. The class II enzyme is present in some bacteria and all eukaryotes. Some bacteria, such as *E. coli*, have both forms of the enzyme. One form is active in the normal citric acid cycle pathway and the other usually specializes in the reverse reaction to convert malate to fumarate.

8. Malate Deydrogenase

The last step in the citric acid cycle is the oxidation of malate to regenerate oxaloacetate, with formation of a molecule of NADH.



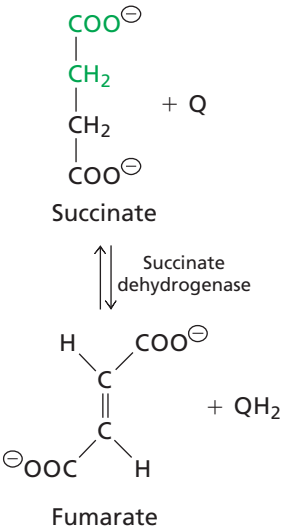
L-Malate Oxaloacetate (13.14)

BOX 13.6 ON THE ACCURACY OF THE WORLD WIDE WEB

There's lots of good stuff on the web but everyone should be cautious about the quality of some webpages. The citric acid cycle is a fun test case for accuracy. Most sites get the basics correct but students are often challenged to find a website that accurately depicts every reaction of the pathway with no errors—including balancing every equation. Can you find such a website? The most common errors are leaving out protons and QH_2.

The one site you can rely on is the IUBMB Enzyme Nomenclature site that lists the correct reactions for each enzyme in the citric acid cycle: **www.chem.qmul.ac.uk/iubmb/enzyme/**

Some instructors have been known to give extra marks to students who can find a completely accurate website. Some students have been known to create their own webpages.



▲ **Figure 13.13**
The succinate dehydrogenase reaction.



▲ **Green (unripe) apples.** The sour taste of unripe apples is mostly due to the presence of malate. Malic acid was first isolated from apple juice and it was named after the Latin word for apple (*malum*).

We consider the transfer of reducing equivalents to Q again in Chapter 14, where we will see the role of the succinate dehydrogenase complex in membrane-associated electron transport.

The evolutionary origin of fumarase and the significance of the reverse reaction in bacteria are described in Section 13.8.

This reaction is catalyzed by $NAD^{\oplus}$-dependent malate dehydrogenase. The near-equilibrium interconversion of the α-hydroxy acid L-malate and the keto acid oxaloacetate is analogous to the reversible reaction catalyzed by lactate dehydrogenase (Sections 7.3 and 11.3B). This is not surprising since lactate dehydrogenase and malate dehydrogenase are homologous—they share a common ancestor.

The standard Gibbs free energy change for this reaction is +30 kJ mol^{-1} ($\Delta G^{\circ\prime}$ = 30 kJ mol^{-1}). Since this is a near-equilibrium reaction it means that under the conditions found inside the cell, the concentration of malate is very much higher than that of oxaloacetate. We've seen in the case of the citrate synthase reaction that the low concentration of oxaloacetate explains the Gibbs free energy change of that reaction. In the next section we'll see how the low concentration of oxaloacetate relative to that of malate explains some transport pathways.

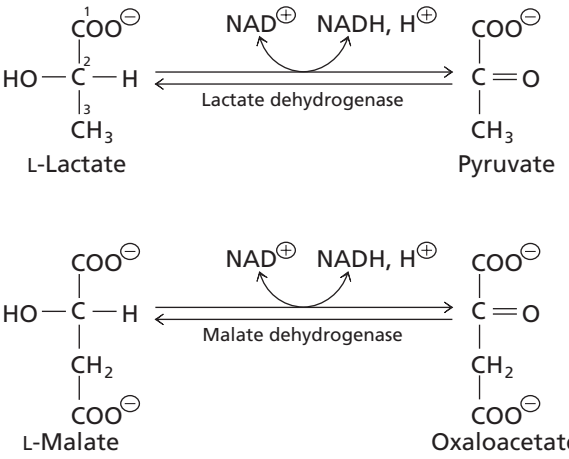
The structures of malate dehydrogenase and lactate dehydrogenase are compared in Figure 4.22.

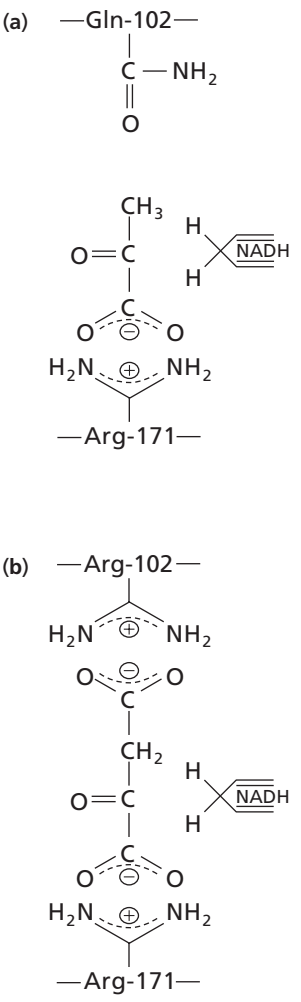
13.4 Entry of Pyruvate Into Mitochondria

In bacterial cells, pyruvate is converted to acetyl CoA in the cytosol but in eukaryotic cells the pyruvate dehydrogenase complex is located in mitochondria (and in chloroplasts). Since glycolysis takes place in the cytoplasm, pyruvate must first be imported into the mitochondria (or chloroplasts) so that it can serve as a substrate in the reaction. The mitochondrion is enclosed by a double membrane. Small molecules such as

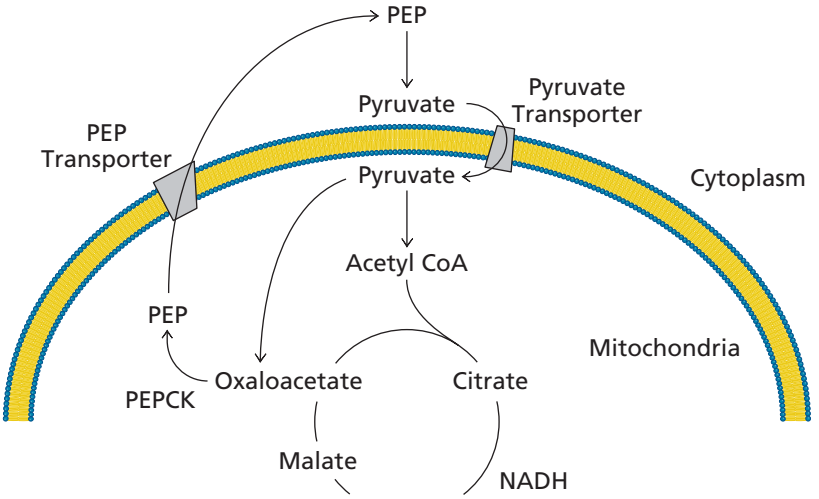
BOX 13.7 CONVERTING ONE ENZYME INTO ANOTHER

Despite having a low sequence identity, lactate dehydrogenase and malate dehydrogenase are closely related in three-dimensional structure and they clearly have evolved from a common ancestor. These enzymes catalyze reversible oxidation of 2-hydroxy acids that differ by only one carbon (malate has an additional carboxylate attached to C-3 of lactate). Both enzymes are highly specific for their own substrates. However, site specific mutation of a single amino acid residue of the lactate dehydrogenase of *Bacillus stearothermophilus* changes this enzyme to a malate dehydrogenase (see figure). Conversion of Gln-102 to Arg-102 completely reverses the specificity of the dehydrogenase. The positively charged side chain of the arginine forms an ion pair with the 4-carboxylate group of malate, and the mutant enzyme becomes inactive with lactate.





▲ **Orientation of the substrate molecule in the active site of lactate dehydrogenase from *Bacillus stearothermophilus*. (a)** The three-carbon substrate pyruvate bound to the native enzyme. Neither oxaloacetate nor malate can bind at this site. **(b)** The four-carbon substrate oxaloacetate bound to the Gln-to-Arg mutant (position 102).



◀ **Figure 13.14**
Import of pyruvate and export of PEP. Pyruvate is imported into mitochondria from the cytoplasm via a pyruvate transporter located in the inner mitochondrial membrane. Phosphoenolpyruvate (PEP) is exported to the cytoplasm via a PEP transporter.

pyruvate pass through the outer membrane via aqueous channels formed by transmembrane proteins called porins (Section 9.11A). These channels allow free diffusion of molecules with molecular weights less than 10,000. However, in order to pass through the inner membrane a specific transport protein is required for most metabolites. Pyruvate translocase specifically transports pyruvate in symport with $H^{\oplus}$. Once inside the mitochondrion, pyruvate can be converted to acetyl CoA and CO_2. In eukaryotic cells the enzymes of the citric acid cycle are also located in the mitochondria (Figure 13.14).

Recall that one of the intermediates in the citric acid cycle is oxaloacetate and it can also be a substrate for gluconeogenesis. Since gluconeogenesis is a cytoplasmic pathway, it's necessary to move oxaloacetate, or its equivalent, from the mitochondria to the cytoplasm. In mammals this is accomplished using a mitochondrial version of phosphoenolpyruvate carboxykinase (PEPCK), that converts oxaloacetate to phosphoenolpyruvate (PEP). Mitochondria possess a PEP transporter that moves PEP to the cytoplasm (Figure 13.14). It would be very inefficient to transport oxaloacetate directly because its concentration in the mitochondria is very low compared to its concentration in the cytoplasm. (Deficiencies in the human mitochondrial PEPCK lead to death within the first two years of life.)

There are two other problems associated with the compartmentation of the citric acid cycle in mitochondria. Acetyl CoA is required for fatty acid synthesis in the cytoplasm, so there has to be a mechanism for transporting acetyl CoA from the mitochondria to the cytoplasm. This is accomplished using a tricarboxylic acid transporter that exports citrate. Once in the cytoplasm, citrate has to be reconverted to oxaloacetate and acetyl CoA and this is accomplished by a cytoplasmic enzyme called ATP-citrate lyase (Figure 13.15). ATP-citrate lyase doesn't just catalyze the reverse of the citrate synthase reaction. The enzyme has to be coupled to hydrolysis of ATP in order to drive the synthesis of "high energy" acetyl CoA in the cytoplasm. The mitochondrial enzyme can catalyze the same reaction (reversing the citric acid cycle reaction) because the concentration of citrate is so high relative to oxaloacetate (see Figure 13.7). In the cytoplasm, on the other hand, the steady state concentrations of citrate and oxaloacetate are comparable, so coupling to ATP hydrolysis is necessary.

Some species don't have a mitochondrial version of PEPCK so they have to use an alternative method of exporting oxaloacetate. The malate–aspartate shuttle is a common transport system, present even in species that have a mitochondrial PEPCK. A simplifed version of this shuttle is shown in Figure 13.16. We will describe it in more detail in Section 14.12.

Oxaloacetate is converted to malate by the reaction catalyzed by malate dehydrogenase. This is the same enzyme used in the citric acid cycle. Recall that the equilibrium concentrations of reactants and products in this reaction result in a very much higher concentration of malate than oxaloacetate. Thus, a malate transporter is much more efficient than an oxaloacetate transporter could be.

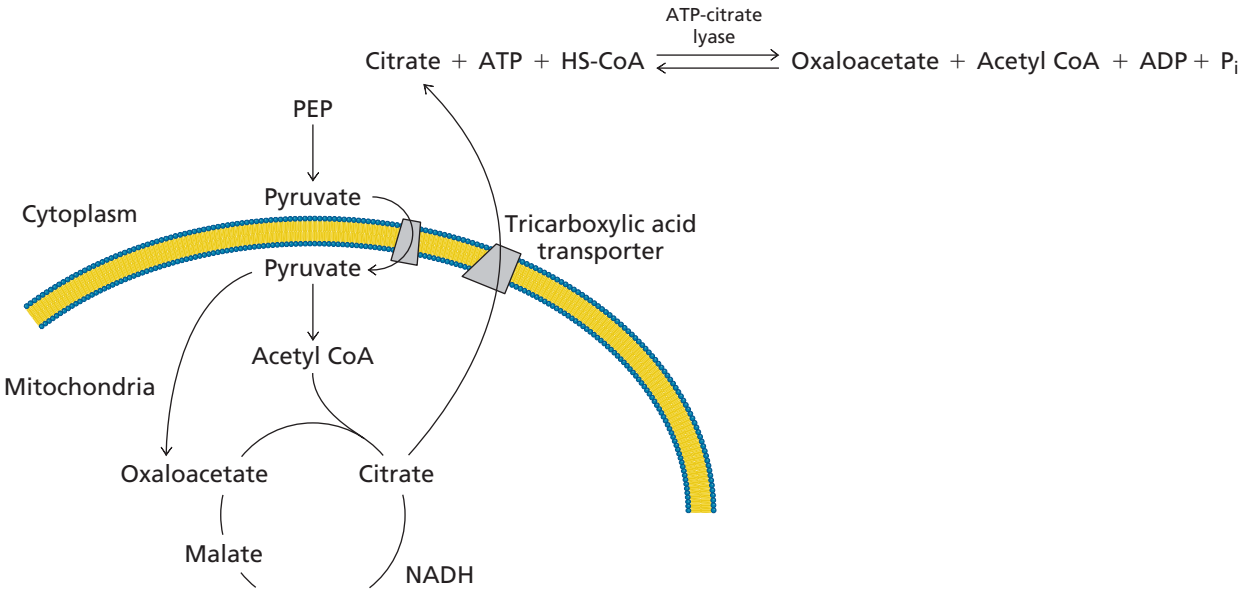
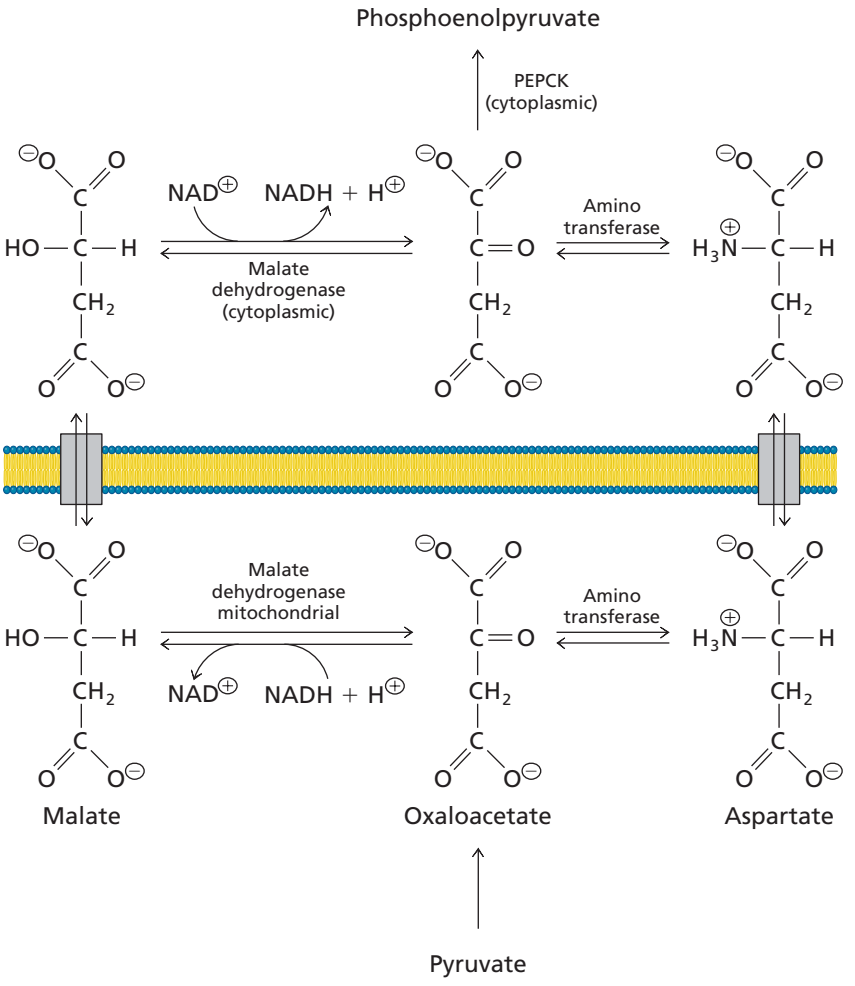


Figure 13.15 ▲
Export of acetyl CoA from mitochondria. Citrate is exported via the tricarboxylic acid transporter. Citrate is subsequently converted to acetyl CoA by cytoplasmic ATP-citrate lyase.

Malate is converted back to oxaloacetate by a cytoplasmic version of malate dehydrogenase. The net effect is that oxaloacetate from mitochondria can serve as a substrate for gluconeogenesis as described in the previous chapter.

The other part of the shuttle achieves the same goal by using a mitochondrial aminotransferase to convert oxaloacetate to aspartate. Aspartate is transported across the mitochondrial membrane by an aspartate transporter. In the cytoplasm, oxaloacetate can

Figure 13.16 ▶
Transport of oxaloacetate via the malate–aspartate shuttle.



be re-formed by the action of a cytoplasmic aminotransferase. As you might guess, this pathway normally operates in the opposite direction, since the low concentration of oxaloacetate in the mitochondria means that the conversion of oxaloacetate to aspartate is unlikely.

13.5 Reduced Coenzymes Can Fuel Production of ATP

In the net reaction of the citric acid cycle, three molecules of NADH, one molecule of QH_2, and one molecule of GTP or ATP are produced for each molecule of acetyl CoA entering the pathway.

$$\begin{aligned}&\text{Acetyl CoA} + 3\ \text{NAD}^{\oplus} + \text{Q} + \text{GDP (or ADP)} + \text{P}_i + 2\ \text{H}_2\text{O} \longrightarrow \\ &\text{HS-CoA} + 3\ \text{NADH} + \text{QH}_2 + \text{GTP (or ATP)} + 2\ \text{CO}_2 + 2\ \text{H}^{\oplus}\end{aligned} \qquad \textbf{(13.15)}$$

As mentioned earlier, NADH and QH_2 can be oxidized by the membrane-asscociated electron transport chain that is coupled to the the production of ATP. As we will see when we examine these reactions in Chapter 14, approximately 2.5 molecules of ATP are generated for each molecule of NADH oxidized to $NAD^{\oplus}$, and up to 1.5 molecules of ATP are produced for each molecule of QH_2 oxidized to Q. The complete oxidation of one molecule of acetyl CoA by the citric acid cycle and subsequent reactions is therefore associated with the production of approximately ten ATP equivalents (Table 13.2).

The citric acid cycle is the final stage in the catabolism of many major nutrients. It is the pathway for oxidation of all acetyl CoA molecules produced by the degradation of carbohydrates, lipids, and amino acids. Having covered glycolysis in Chapter 11, we can now give a complete accounting of the ATP produced from the degradation of one molecule of glucose.

Recall that glycolysis converts glucose to two molecules of pyruvate with a net gain of two molecules of ATP. There are two molecules of NADH produced in the reaction catalyzed by glyceraldehyde 3-phosphate dehydrogenase. This corresponds to a combined yield of seven ATP equivalents from glycolysis. The conversion of both pyruvate molecules to acetyl CoA by the pyruvate dehydrogenase complex yields two NADH molecules, which correspond to about five additional molecules of ATP. When these are combined with the ATP equivalents from the citric acid cycle via the oxidation of two molecules of acetyl CoA, the total yield is about 32 molecules of ATP per molecule of glucose (Figure 13.17).

In bacteria, the two molecules of NADH produced by glycolysis in the cytosol can be directly reoxidized by the membrane-associated electron transport system in the plasma membrane. Thus, the theoretical maximum yield from complete oxidation of glucose (32 ATP equivalents) is achieved in bacteria cells.

In eukaryotic cells, glycolysis produces NADH in the cytosol but the membrane-associated electron transport complex is located in mitochondria membranes. The reducing equivalents from cytosolic NADH can be transported into the mitochondrion by shuttle

Table 13.2 Energy production in the citric acid cycle

Reaction	Energy-yielding product	ATP equivalents
Isocitrate dehydrogenase	NADH	2.5
α-Ketoglutarate dehydrogenase complex	NADH	2.5
Succinyl-CoA synthetase	GTP or ATP	1.0
Succinate dehydrogenase complex	QH_2	1.5
Malate dehydrogenase	NADH	2.5
Total		10.0

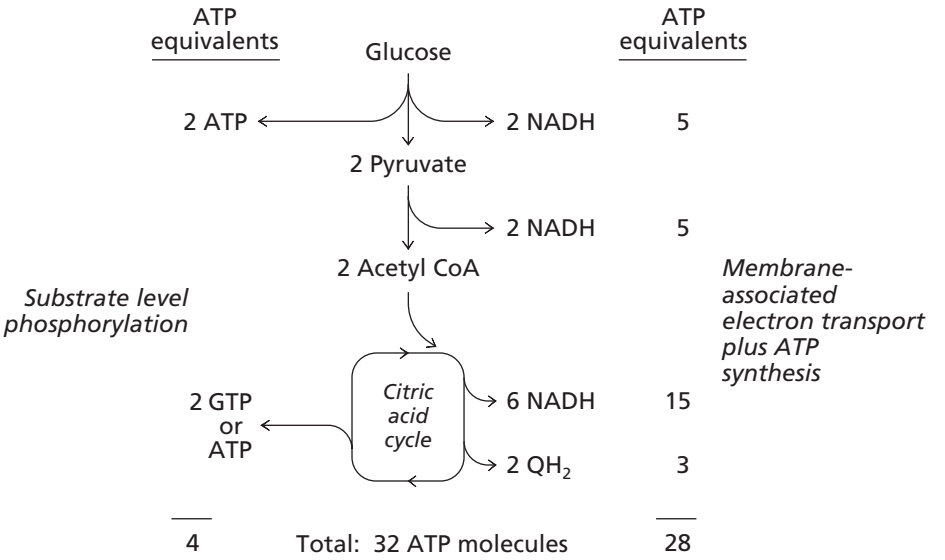


Figure 13.17 ►
ATP production from the catabolism of one molecule of glucose by glycolysis, the citric acid cycle, and reoxidation of NADH and QH_2. The complete oxidation of glucose leads to the formation of up to 32 molecules of ATP.

mechanisms such as the malate–aspartate shuttle described in Section 13.4. The transport of reducing equivalents of NADH will be described in more detail in Section 14.12.

It's interesting to compare this pathway (Figure 13.17) for complete oxidation of glucose to the pentose phosphate cycle described in Section 12.4. That pathway also results in the complete oxidation of one molecule of glucose. The result is production of 12 NADPH molecules that are equal to 30 ATP equivalents.

13.6 Regulation of the Citric Acid Cycle

Because the citric acid cycle occupies a central position in cellular metabolism, it's not surprising to find that the pathway is controlled. Regulation is mediated by allosteric modulators and by covalent modification of the citric acid cycle enzymes. Flux through the pathway is further controlled by the supply of acetyl CoA.

As noted earlier, acetyl CoA arises from several sources, including pathways for the degradation of carbohydrates, lipids, and amino acids. The activity of the pyruvate dehydrogenase complex controls the supply of acetyl CoA produced from pyruvate and hence from the degradation of carbohydrates. In general, substrates of the pyruvate dehydrogenase complex activate the complex and products inhibit it. In most species, the activities of the E_2 and E_3 components of the pyruvate dehydrogenase complex (dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase, respectively) are controlled by simple mass action effects when their products accumulate. The activity of the acetyltransferase (E_2) is inhibited when the concentration of acetyl CoA is high, whereas the dehydrogenase (E_3) is inhibited by a high NADH/$NAD^{\oplus}$ ratio (Figure 13.18). In general, the inhibitors are likely to be present in high concentrations when energy resources are plentiful, and the activators predominate when energy resources are scarce.

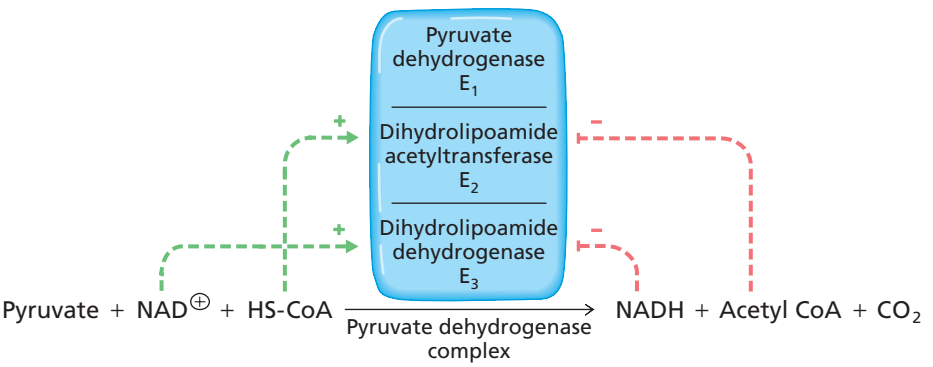
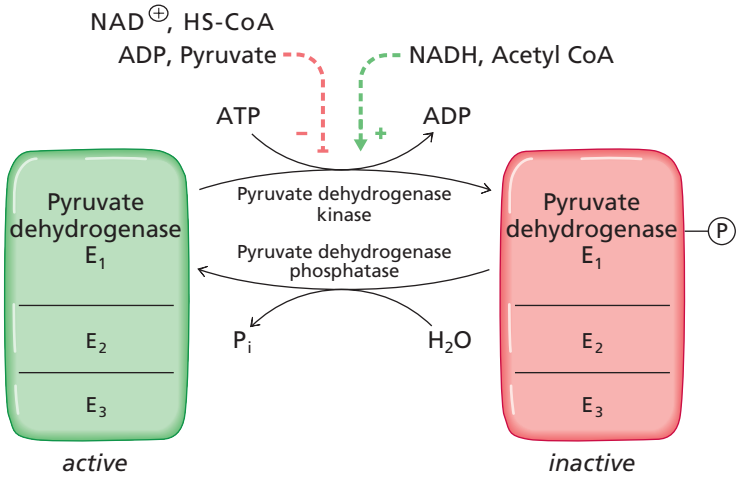


Figure 13.18 ►
Regulation of the the pyruvate dehydrogenase complex. Accumulation of the products acetyl CoA and NADH decreases flux through the reversible reactions catalyzed by E_2 and E_3.



◀ **Figure 13.19**
Regulation of the mammalian pyruvate dehydrogenase complex by phosphorylation of the E_1 component. The regulatory kinase and phosphatase are both components of the mammalian complex. The kinase is activated by NADH and acetyl CoA, products of the reaction catalyzed by the pyruvate dehydrogenase complex, and inhibited by ADP and the substrates pyruvate, $NAD^{\oplus}$, and HS-CoA.

Mammalian (but not prokaryotic) pyruvate dehydrogenase complexes are further regulated by covalent modification. A protein kinase and a protein phosphatase are associated with the mammalian multienzyme complex. Pyruvate dehydrogenase kinase (PDK) catalyzes the phosphorylation of E_1, thereby inactivating the enzyme. Pyruvate dehydrogenase phosphatase (PDP) catalyzes the dephosphorylation and activation of pyruvate dehydrogenase (Figure 13.19). Control of E_1 activity controls the rate of reaction of the entire complex.

Pyruvate dehydrogenase kinase and pyruvate dehydrogenase phosphatase are themselves regulated. The kinase is allosterically activated by NADH and acetyl CoA, products of pyruvate oxidation. The accumulation of NADH and acetyl CoA signals energy availability and leads to an increase in phosphorylation of the pyruvate dehydrogenase subunit and inhibition of the further oxidation of pyruvate. Conversely, pyruvate, $NAD^{\oplus}$, HS-CoA, and ADP inhibit the kinase, leading to activation of the pyruvate dehydrogenase subunit.

Three enzymes of the citric acid cycle are regulated: citrate synthase, isocitrate dehydrogenase, and the α-ketoglutarate dehydrogenase complex. Citrate synthase catalyzes the first reaction of the citric acid cycle. This would seem to be a suitable control point for regulation of the entire cycle. ATP inhibits the enzyme *in vitro*, but significant changes in ATP concentration are unlikely *in vivo*; therefore, ATP may not be a physiological regulator. Some bacterial citrate synthases are activated by α-ketoglutarate and inhibited by NADH.

Mammalian isocitrate dehydrogenase is allosterically activated by Ca^{2+} and ADP and inhibited by NADH. In mammals, the enzyme is not subject to covalent modification. In bacteria, however, isocitrate dehydrogenase is regulated by phosphorylation. We will discuss this in more detail in Section 13.8.

Although the α-ketoglutarate dehydrogenase complex resembles the pyruvate dehydrogenase complex, the enzymes have quite different regulatory features. No kinase or phosphatase is associated with the α-ketoglutarate dehydrogenase complex. Instead, calcium ions bind to E_1 of the complex and decrease the K_m of the enzyme for α-ketoglutarate, thereby increasing the rate of formation of succinyl CoA. NADH and succinyl CoA are inhibitors of the α-ketoglutarate complex *in vitro*, but it has not been established that they have a significant regulatory role in living cells.

13.7 The Citric Acid Cycle Isn't Always a "Cycle"

The citric acid cycle is not exclusively a catabolic pathway for the oxidation of acetyl CoA. It also plays a central role in metabolism at the intersection of several other pathways. Some intermediates of the citric acid cycle are important anabolic precursors in biosynthesis pathways, and some catabolic pathways produce citric acid cycle intermediates. Pathways that are both catabolic and anabolic are said to be *amphibolic* (Section 10.1). The citric acid cycle is an excellent example.

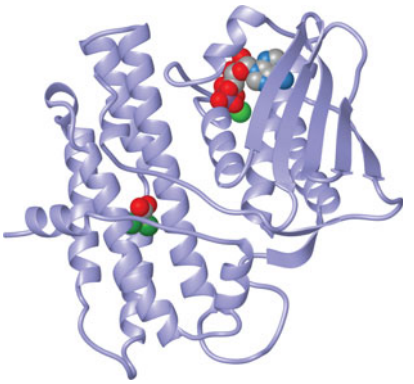
BOX 13.8 A CHEAP CANCER DRUG?

In the absence of oxygen, the glycolytic pathway terminates at lactate and the citric acid cycle is not used in the oxidation of acetyl CoA. Under these conditions, pyruvate dehydrogenase is inactivated by phosphorylation. Many cancer cells grow anaerobically and pyruvate dehydrogenase is not active in these cells.

The activity of pyruvate dehydrogenase phosphorylase kinase (PDHK) can be inhibited by dichloroacetate (DCA). DCA binds to the active site of the enzyme preventing phosphorylation of pyruvate dehydrogenase. The net effect of DCA is activation of pyruvate dehydrogenase and this, in turn, causes major disruptions in cancer cell metabolism leading to death of the cancer cells. The chemical has been effective in a few trial studies with cancer cells *in vitro*. That's a good thing.

Unfortunately, the effectiveness of DCA as a cancer drug has not been demonstrated in clinical trials. Medical researchers are in a difficult position. The biochemistry is sound. It makes sense that cancer cells grow anaerobically (the Warburg effect) and it makes sense that DCA might be an effective cancer drug based on its ability to inhibit PDHK. However most physicians are reluctant to prescribe DCA in the absence of evidence of its effectiveness.

DCA has been around for a long time and it cannot be patented. This has provoked the claim that major drug companies are conspiring to suppress evidence of DCA's effectiveness on the grounds that they cannot make any money by selling DCA. A cottage industry of suppliers has sprung up on the Internet for people who want to treat themselves with this cheap "miracle" drug. The Food and Drug Administration in the United States has been forced to shut down some websites because they were making unsubstantiated claims about its ability to cure cancer. There was also concern about self-medication because high dosages of DCA are toxic. There's bound to be more publicity surrounding this complicated issue in the future. The blog *Respectful Insolence* (scienceblogs.com/insolence) is a good source of scientific and medical information on the controversy.

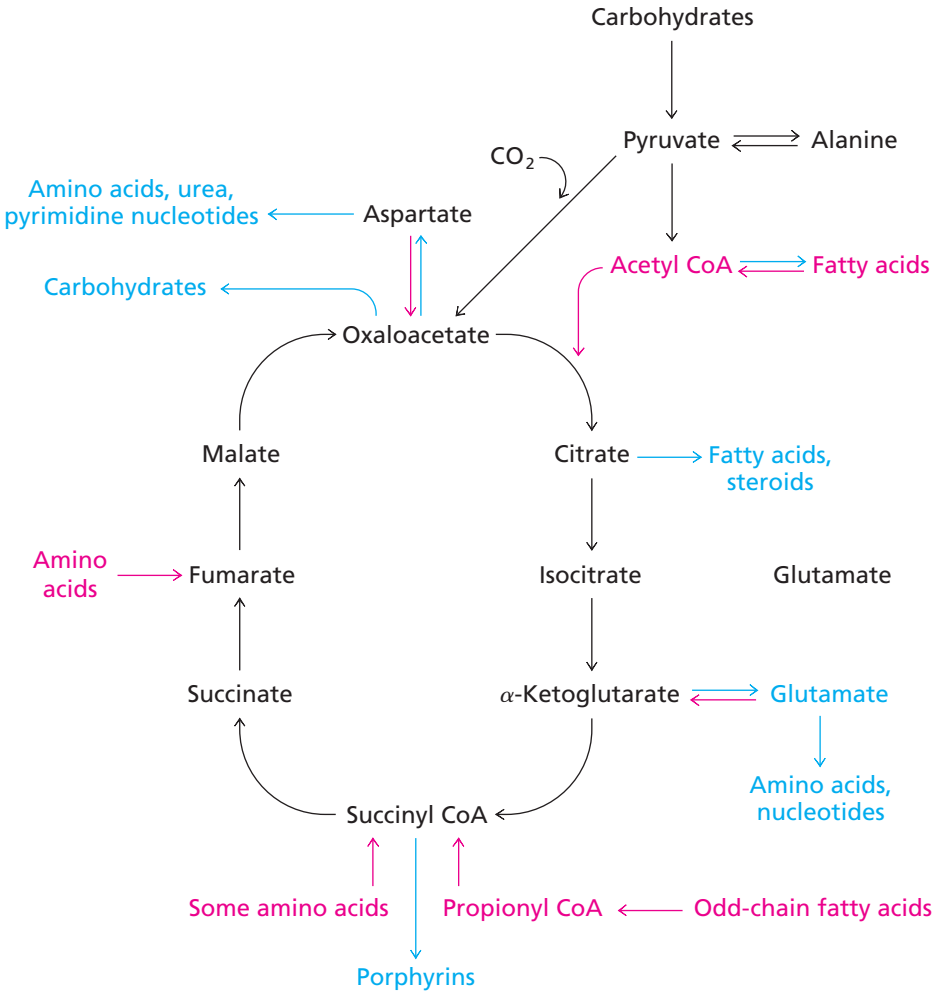


▲ **Pyruvate dehydrogenase kinase with dichloroacetate bound at the active site.** The human (*Homo sapiens*) PDHK is a dimer, only one subunit is shown here. The bound ligands are shown as space-filling molecules. ADP (top) is bound at the allosteric site, and dichloroacetate (left) is bound at the active site. [PDB 2BU8]

As shown in Figure 13.20, citrate, α-ketoglutarate, succinyl CoA, and oxaloacetate all lead to biosynthetic pathways. Citrate is part of a pathway for the formation of fatty acids and steroids. It undergoes cleavage to form acetyl CoA, the precursor of the lipids. In eukaryotes, this reaction takes place in the cytosol, and citrate must be transported from the mitochondria to the cytosol to support fatty acid biosynthesis. One major metabolic fate of α-ketoglutarate is reversible conversion to glutamate, which can then be incorporated into proteins or used for the synthesis of other amino acids or nucleotides. We will see in Chapter 17 that α-ketoglutarate pools are important in nitrogen metabolism. Succinyl CoA can condense with glycine to initiate the biosynthesis of porphyrins such as the heme groups of cytochromes. As we saw in the previous chapter, oxaloacetate is a precursor of carbohydrates formed by gluconeogenesis. Oxaloacetate also interconverts with aspartate, which can be used in the synthesis of urea, amino acids, and pyrimidine nucleotides.

When the citric acid cycle functions as a multistep catalyst, only small amounts of each intermediate are needed to convert large quantities of acetyl CoA to products. Therefore, the rate at which the citric acid cycle metabolizes acetyl CoA is extremely sensitive to changes in the concentrations of its intermediates. Thus, citric acid cycle intermediates that are removed by entry into biosynthetic pathways must be replenished by **anaplerotic** (Greek, "filling up") reactions. Because the pathway is cyclic, replenishing any of the cycle intermediates results in a greater concentration of all intermediates. Depletion of citric acid cycle intermediates is an example of a **cataplerotic** reaction. It's just as important as the filling up reactions.

The production of oxaloacetate by pyruvate carboxylase is an important anaplerotic reaction (Figure 13.20). This reaction is also part of the gluconeogenesis pathway (Section 12.1A). Pyruvate carboxylase is allosterically activated by acetyl CoA. The accumulation of acetyl CoA indicates a low concentration of oxaloacetate and a need for



◀ **Figure 13.20**
Routes leading to and from the citric acid cycle. Intermediates of the citric acid cycle are precursors of carbohydrates, lipids, and amino acids, as well as nucleotides and porphyrins. Reactions feeding into the cycle replenish the pool of cycle intermediates. Anabolic pathways are colored blue and catabolic pathways are colored red.

more citric acid cycle intermediates. The activation of pyruvate carboxylase supplies oxaloacetate for the cycle.

Many species use a variety of different reactions to keep the intake and output of citric acid cycle intermediates in a delicate balance. For example, many plants and some bacteria supply oxaloacetate to the citric acid cycle via a reaction catalyzed by phosphoenolpyruvate carboxylase.

$$\text{Phosphoenolpyruvate} + HCO_3^{\ominus} \rightleftharpoons \text{Oxaloacetate} + P_i \qquad \textbf{(13.16)}$$

Pathways for degrading some amino acids and fatty acids can contribute succinyl CoA to the citric acid cycle. The interconversion of oxaloacetate and aspartate and of α-ketoglutarate and glutamate can either supply or remove intermediates of the cycle.

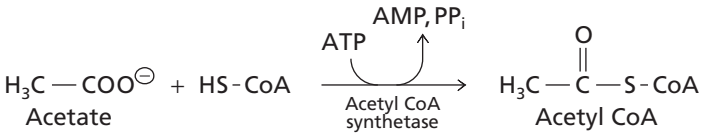
The interplay of all these reactions—the entry of acetyl CoA from glycolysis and other sources, the entry of intermediates from catabolic pathways and anaplerotic reactions, and the exit of intermediates to anabolic pathways—means that the citric acid cycle doesn't always operate as a simple cycle devoted to oxidizing acetyl CoA. In fact, most bacteria don't have all of the classic enzymes of the citric acid cycle so there is no "cycle" in these species. Instead, the enzymes that are present are used mostly in biosynthesis pathways where the intermediates become precursors for the synthesis of amino acids and porphyrins (Section 13.9).

13.8 The Glyoxylate Pathway

The glyoxylate pathway is a route that bypasses some of the reactions of the citric acid cycle. The pathway is named after the two-carbon molecule glyoxylate, an essential

intermediate in the pathway. There are only two reactions. In the first reaction, a six-carbon tricarboxylic acid (isocitrate) is split into a two-carbon molecule (glyoxylate) and a four-carbon dicarboxylic acid (succinate). This reaction is catalyzed by isocitrate lyase (Figure 13.21). In the second reaction, the two-carbon glyoxylate molecule combines with a two-carbon acetyl CoA molecule to make a four-carbon dicarboxylic acid (malate). The enzyme for the second reaction is malate synthase.

The glyoxylate pathway was first discovered in bacteria. Subsequently it was found in plants and later in fungi, protists, and some animals. The pathway is often called the glyoxylate shunt, the glyoxylate bypass, or the glyoxylate cycle. The glyoxylate pathway provides an anabolic alternative for the metabolism of acetyl CoA, leading to the formation of glucose from acetyl CoA via four-carbon compounds. Cells that contain glyoxylate pathway enzymes can synthesize all their required carbohydrates from any substrate that is a precursor of acetyl CoA. For example, yeast can grow on ethanol because yeast cells can oxidize ethanol to form acetyl CoA, which can be metabolized via the glyoxylate pathway to form malate. Similarly, many bacteria use the glyoxylate pathway to sustain growth on acetate, which can be incorporated into acetyl CoA in a reaction catalyzed by acetyl CoA synthetase.



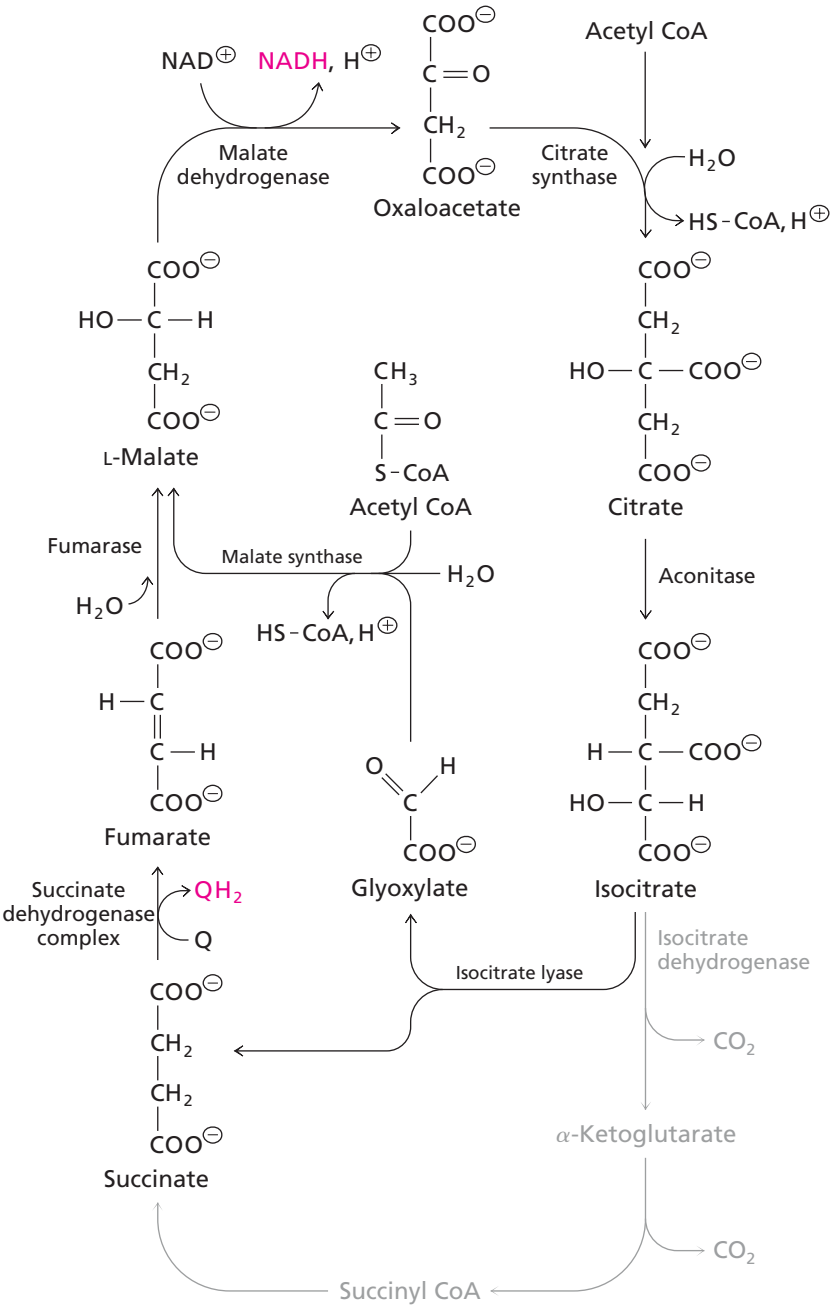
(13.17)

The glyoxylate pathway is a fundamental metabolic pathway in bacteria, protists, fungi, and plants. It is especially active in oily seed plants. In these plants, stored seed oils (triacylglycerols) are converted to carbohydrates that provide fuel during germination. In contrast, genes for the two enzymes of the pathway are present in most animals but the pathway is not actively used. Consequently, in humans acetyl CoA does not serve as the precursor for the net formation of either pyruvate or oxaloacetate; therefore, acetyl CoA is not a carbon source for the net production of glucose. (The carbon atoms of acetyl CoA are incorporated into oxaloacetate by the reactions of the citric acid cycle, but for every two carbon atoms incorporated, two other carbon atoms are released as CO_2.)

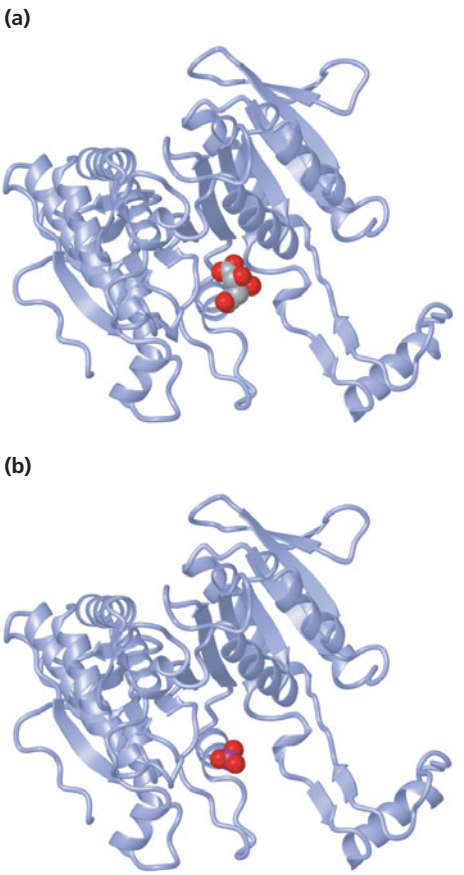
The glyoxylate pathway can be regarded as a shunt within the citric acid cycle, as shown in Figure 13.21. The two reactions provide a bypass around the CO_2-producing reactions of the citric acid cycle. No carbon atoms of the acetyl group of acetyl CoA are released as CO_2 during operation of the glyoxylate shunt, and the net formation of a four-carbon molecule from two molecules of acetyl CoA supplies a precursor that can be converted to glucose by gluconeogenesis. Succinate is oxidized to malate and oxaloacetate by the citric acid cycle to maintain the catalytic amounts of citric acid cycle intermediates. You can think of the glyoxylate shunt as part of a cycle that includes the upper portion of the citric acid cycle. In this case, the net reaction includes the formation of oxaloacetate for gluconeogenesis and the cyclic oxidation of succinate. Two molecules of acetyl CoA are consumed.

$$\begin{aligned} &2\ \text{Acetyl CoA} + 2\ \text{NAD}^{\oplus} + \text{Q} + 3\ \text{H}_2\text{O} \longrightarrow \\ &\qquad \text{Oxaloacetate} + 2\ \text{HS-CoA} + 2\ \text{NADH} + \text{QH}_2 + 4\ \text{H}^{\oplus} \end{aligned} \qquad \textbf{(13.18)}$$

In eukaryotes, the operation of the glyoxylate cycle requires the transfer of metabolites between the mitochondria, where the citric acid cycle enzymes are located, and the cytosol, where isocitrate lyase and malate synthase are found. Thus, the actual pathway is more complicated than the diagram in Figure 13.21. In plants, the glyoxylate pathway enzymes are localized to a special membrane-bound organelle called the glyoxysome. Glyoxysomes contain some special versions of the citric acid cycle enzymes, but some metabolites still have to be transferred between compartments in order for the pathway to operate as a cycle.



◂ **Figure 13.21**
Glyoxylate pathway. Isocitrate lyase and malate synthase are the two enzymes of the pathway. When the pathway is functioning, the acetyl carbon atoms of acetyl CoA are converted to malate rather than oxidized to CO_2. Malate can be converted to oxaloacetate, which is a precursor in gluconeogenesis. The succinate produced in the cleavage of isocitrate is oxidized to oxaloacetate to replace the four-carbon compound consumed in glucose synthesis.



▲ **Figure 13.22**
Phosphorylated and dephosphorylated forms of *E. coli* isocitrate dehydrogenase. **(a)** The dephosphorylated enzyme is active; isocitrate binds to the active site. [PDB5ICD] **(b)** The phosphorylated enzyme is inactive because the negatively charged phosphoryl group (red) electrostatically repels the substrate, preventing it from binding. [PDB4ICD]

In bacteria, the glyoxylate pathway is often used to replenish citric acid cycle metabolites that are diverted into a number of biosynthesis pathways. Since all of the reactions take place in the cytosol in bacteria, it is important to regulate the flow of metabolites. The key regulated enzyme is isocitrate dehydrogenase. Its activity is regulated by covalent modification. Kinase-catalyzed phosphorylation of a serine residue abolishes isocitrate dehydrogenase activity. In the dephosphorylated form of the enzyme, the serine residue forms a hydrogen bond with a carboxylate group of isocitrate. Phosphorylation inhibits enzyme activity by causing electrostatic repulsion of the substrate rather than by causing an R-to-T conformational change (Figure 13.22). The same protein molecule that contains the kinase activity also has a separate domain with phosphatase activity that catalyzes hydrolysis of the phosphoserine residue, reactivating isocitrate dehydrogenase.

The kinase and phosphatase activities are reciprocally regulated; isocitrate, oxaloacetate, pyruvate, and the glycolytic intermediates 3-phosphoglycerate and phosphoenolpyruvate allosterically activate the phosphatase and inhibit the kinase

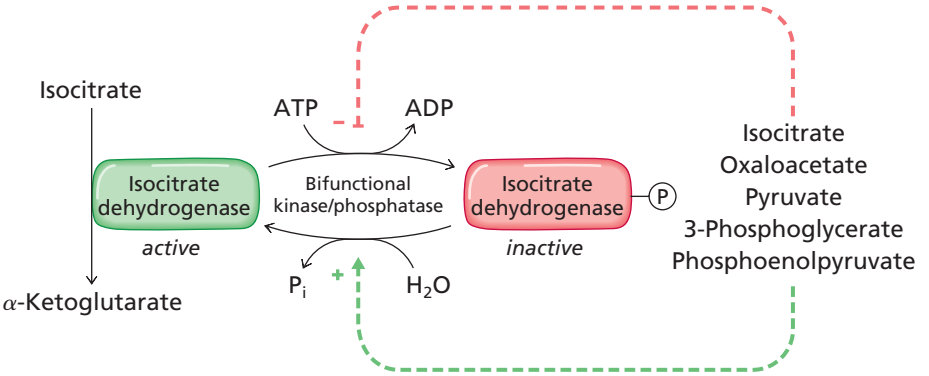


Figure 13.23 ▶
Regulation of *E. coli* isocitrate dehydrogenase by covalent modification. A bifunctional enzyme catalyzes phosphorylation and dephosphorylation of isocitrate dehydrogenase. The two activities of the bifunctional enzyme are reciprocally regulated allosterically by intermediates of glycolysis and the citric acid cycle.

(Figure 13.23). Thus, when the concentrations of glycolytic and citric acid cycle intermediates in *E. coli* are high, isocitrate dehydrogenase is active. When phosphorylation abolishes the activity of isocitrate dehydrogenase, isocitrate is diverted to the glyoxylate pathway.

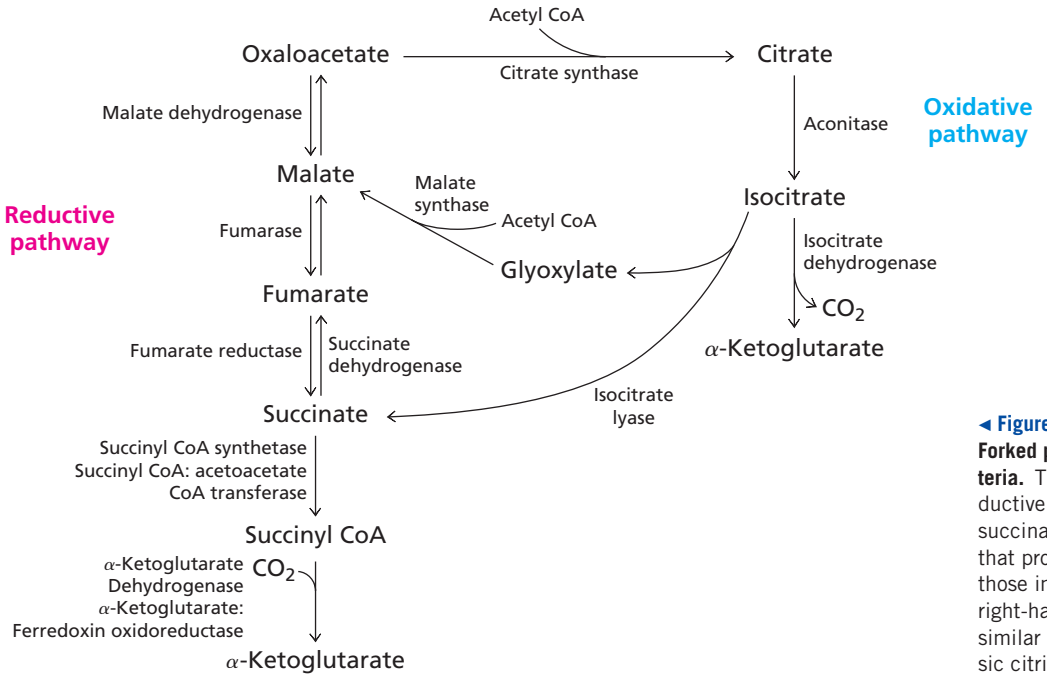
13.9 Evolution of the Citric Acid Cycle

The reactions of the citric acid cycle were first discovered in mammals and many of the key enzymes were purified from liver extracts. As we have seen, the citric acid cycle can be viewed as the end stage of glycolysis because it results in the oxidation of acetyl CoA produced as one of the products of glycolysis. However, there are many organisms that do not encounter glucose as a major carbon source and the production of ATP equivalents via glycolysis and the citric acid cycle is not an important source of metabolic energy in such species.

We need to examine the function of the citric acid cycle enzymes in bacteria in order to understand their role in simple single-celled organisms. These roles might allow us to deduce the pathways that could have existed in the primitive cells that eventually gave rise to complex eukaryotes. Fortunately, the sequences of several hundred prokaryotic genomes are now available as a result of the huge technological advances in recombinant DNA technology and DNA sequencing methods. We can now examine the complete complement of metabolic enzymes in many diverse species of bacteria and ask whether they possess the pathways that we have discussed in this chapter. These analyses are greatly aided by developments in the fields of comparative genomics, molecular evolution, and bioinformatics.

Most species of bacteria do not have a complete citric acid cycle. The most common versions of an incomplete cycle include part of the left-hand side. This short linear pathway leads to production of succinate or succinyl CoA or α-ketoglutarate by a reductive process using oxaloacetate as a starting point. This reductive pathway is the reverse of the traditional cycle that functions in the mitochondria of eukaryotes. In addition, many species of bacteria also have enzymes from part of the right-hand side of the citric acid cycle, especially citrate synthase and aconitase. This allows them to synthesize citrate and isocitrate from oxaloacetate and acetyl CoA. The presence of a forked pathway (Figure 13.24) results in the synthesis of all the precursors of amino acids, porphyrins, and fatty acids.

There are hundreds of diverse species of bacteria that can survive and grow in the complete absence of oxygen. Some of these species are obligate anaerobes—for them, oxygen is a lethal poison! Others are facultative anaerobes—they can survive in oxygen free environments as well as oxygen-rich environments. *E. coli* is one example of a species that can survive in both types of environment. When growing anaerobically, *E. coli* uses a forked version of the pathway to produce the necessary metabolic precursors and avoid the accumulation of reducing equivalents that cannot be reoxidized by the oxygen requiring electron transport system. Bacteria such as *E. coli* can grow in environments where acetate is the only source of organic carbon. In this case, they employ the glyoxylate pathway to convert acetate to malate and oxaloacetate for glucose synthesis.



◀ **Figure 13.24**
Forked pathway found in many species of bacteria. The left-hand side of the fork is a reductive pathway leading to the synthesis of succinate or α-ketoglutarate in reactions that proceed in the reverse direction from those in the classic citric acid cycle. The right-hand branch is an oxidative pathway similar to the first few reactions of the classic citric acid cycle.

The first living cells arose in an oxygen-free environment over three billion years ago. These primitive cells undoubtedly possessed most of the enzymes that interconverted acetate, pyruvate, citrate, and oxaloacetate, since these enzymes are present in most modern bacteria. The development of the main branches of the forked pathway possibly began with the evolution of malate dehydrogenase from a duplication of the lactate dehydrogenase gene. Aconitase and isocitrate dehydrogenase evolved from enzymes that are used in the synthesis of leucine (isopropylmalate dehydratase and isopropylmalate dehydrogenase, respectively). (Note that the leucine biosynthesis pathway is more ubiquitous and more primitive than the citric acid cycle.)

Extension of the reductive branch continued with the evolution of fumarase from aspartase. Aspartase is a common bacterial enzyme that synthesizes fumarate from L-aspartate. L-aspartate, in turn, is synthesized by amination of oxaloacetate in a reaction catalyzed by aspartate transaminase (Section 17.3). It is likely that primitive cells used the pathway oxaloacetate ⟶ aspartate ⟶ fumarate to produce fumarate before the evolution of malate dehydrogenase and fumarase. The reduction of fumarate to succinate is catalyzed by fumarate reductase in many bacteria. The evolutionary origin of this complex enzyme is highly speculative but at least one of the subunits is related to another enzyme of amino acid metabolism. Succinate dehydrogenase, the enzyme that preferentially catalyzes the reverse reaction in the citric acid cycle, is likely to have evolved later on from fumarate reductase via a gene duplication event.

The synthesis of α-ketoglutarate can occur in either branch of the forked pathway. The reductive branch uses α-ketoglutarate:ferredoxin oxidoreductase, an enzyme found in many species of bacteria that don't have a complete citric acid cycle. The reaction catalyzed by this enzyme is not readily reversible. With the evolution of α-ketoglutarate dehydrogenase the two forks can be joined to create a cyclic pathway. It is clear that α-ketoglutarate dehydrogenase and pyruvate dehydrogenase share a common ancestor and it is likely that this was the last enzyme to evolve.

Some bacteria have a complete citric acid cycle but it is used in the reductive direction to fix CO_2 in order to build more complex organic molecules. This could have been one of the selective pressures leading to a complete pathway. The cycle requires a terminal electron acceptor to oxidize NADH and QH_2 when it operates in the more normal oxidative direction seen in eukaryotes. Originally, this terminal electron acceptor was sulfur or various sulfates, and these reactions still occur in many anaerobic bacterial species. Oxygen levels began to rise about 2.5 billion years ago with the evolution of photosynthesis reactions in cyanobacteria. Some bacteria, notably proteobacteria, exploited the availability of

oxygen when the membrane-associated electron transport reactions evolved. One species of proteobacteria entered into a symbiotic relationship with a primitive eukaryotic cell about two billion years ago. This led to the evolution of mitochondria and the modern versions of the citric acid cycle and electron transport in eukaryotes.

The evolution of the citric acid cycle pathway involved several of the pathway evolution mechanisms discussed in Chapter 10. There is evidence for gene duplication, pathway extension, retro-evolution, pathway reversal, and enzyme theft.

Summary

1. The pyruvate dehydrogenase complex catalyzes the oxidation of pyruvate to form acetyl CoA and CO_2.
2. For each molecule of acetyl CoA oxidized via the citric acid cycle, two molecules of CO_2 are produced, three molecules of $NAD^{\oplus}$ are reduced to NADH, one molecule of Q is reduced to QH_2 and one molecule of GTP is generated from GDP + P_i (or ATP from ADP + P_i, depending on the species).
3. The eight enzyme-catalyzed reactions of the citric acid cycle can function as a multistep catalyst.
4. In eukaryotic cells, pyruvate must be imported into the mitochondria by a specific transporter before it can serve as a substrate for the pyruvate dehydrogenase reaction.
5. Oxidation of the reduced coenzymes generated by the citric acid cycle leads to the formation of about 10 ATP molecules per molecule of acetyl CoA entering the pathway, for a total of about 32 ATP molecules per complete oxidation of 1 molecule of glucose.
6. The oxidation of pyruvate is regulated at the steps catalyzed by the pyruvate dehydrogenase complex, isocitrate dehydrogenase, and the α-ketoglutarate dehydrogenase complex.
7. In addition to its role in oxidative catabolism, the citric acid cycle provides precursors for biosynthetic pathways. Anaplerotic reactions replenish cycle intermediates.
8. The glyoxylate cycle, a modification of the citric acid cycle, allows many organisms to use acetyl CoA to generate four-carbon intermediates for gluconeogenesis.
9. The citric acid cycle probably evolved from the more primitive forked pathway found in many modern species of bacteria.

Problems

1. (a) The citric acid cycle converts one molecule of citrate to one molecule of oxaloacetate, which is required for the cycle to continue. If other cycle intermediates are depleted by being used as precursors for amino acid biosynthesis, can a net synthesis of oxaloacetate occur from acetyl CoA via the enzymes of the citric acid cycle?
   (b) How can the cycle continue to function if insufficient oxaloacetate is present?
2. Fluoroacetate, a very toxic molecule that blocks the citric acid cycle, has been used as a rodent poison. It is converted enzymatically *in vivo* to fluoroacetyl CoA, which is then converted by the action of citrate synthase to 2R,3S-fluorocitrate, a potent competitive inhibitor of the next enzyme in the pathway. Predict the effect of fluoroacetate on the concentrations of the intermediates in the citric acid cycle. How can this blockage of the cycle be overcome?
3. Calculate the number of ATP molecules generated by the following net reactions of the citric acid cycle. Assume that all NADH and QH_2 are oxidized to yield ATP, pyruvate is converted to acetyl CoA, and the malate–aspartate shuttle is operating.
   (a) 1 Pyruvate $\longrightarrow$ 3 CO_2
   (b) Citrate $\longrightarrow$ Oxaloacetate + 2 CO_2
4. When one molecule of glucose is completely oxidized to six molecules of CO_2 under the conditions in Problem 3, what percentage of ATP is produced by substrate level phosphorylation?
5. The disease beriberi, which results from a dietary deficiency of vitamin B_1 (thiamine), is characterized by neurologic and cardiac symptoms, as well as increased levels of pyruvate and α-ketoglutarate in the blood. How does a deficiency of thiamine account for the increased levels of pyruvate and α-ketoglutarate
6. In three separate experiments, pyruvate labeled with ^{14}C at C-1, at C-2, or at C-3 is metabolized via the pyruvate dehydrogenase complex and the citric acid cycle. Which labeled pyruvate molecule is the first to yield $^{14}CO_2$? Which is the last to yield $^{14}CO_2$, and how many turns of the cycle are required to release all of the labeled carbon atoms as $^{14}CO_2$?
7. Patients in shock experience decreased delivery of O_2 to tissues, decreased activity of the pyruvate dehydrogenase complex, and increased anaerobic metabolism. Excess pyruvate is converted to lactate, which accumulates in tissues and in the blood, causing lactic acidosis.
   (a) Since O_2 is not a reactant or product of the citric acid cycle, why do low levels of O_2 decrease the activity of the pyruvate dehydrogenase complex?
   (b) To alleviate lactic acidosis, shock patients are sometimes given dichloroacetate, which inhibits pyruvate dehydrogenase kinase. How does this treatment affect the activity of the pyruvate dehydrogenase complex?
8. A deficiency of a citric acid cycle enzyme in both mitochondria and the cytosol of some tissues (e.g., blood lymphocytes) results in severe neurological abnormalities in newborns. The disease is characterized by excretion in the urine of abnormally large amounts of α-ketoglutarate, succinate, and fumarate. What enzyme deficiency would lead to these symptoms?

9. Acetyl CoA inhibits dihydrolipoamide acetyltransferase (E_2 of the pyruvate dehydrogenase complex) but activates the pyruvate dehydrogenase kinase component of the pyruvate dehydrogenase complex. How are these two different actions of acetyl CoA consistent with the overall regulation of the complex?

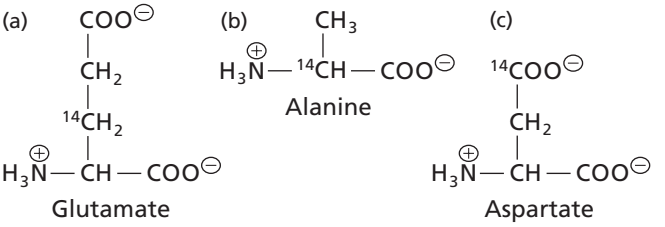
10. Pyruvate dehydrogenase complex deficiency is a disease that results in various metabolic and neurological effects. Pyruvate dehydrogenase complex deficiency can cause lactic acidosis in affected children. Other clinical symptoms include increased concentrations of pyruvate and alanine in the blood. Explain the increase in the levels of pyruvate, lactate, and alanine in individuals with pyruvate dehydrogenase complex deficiency.

11. In response to a signal for contraction and the resulting increased need for ATP in vertebrate muscle, Ca^{2+} is released into the cytosol from storage sites in the endoplasmic reticulum. How does the citric acid cycle respond to the influx of Ca^{2+} in satisfying the increased need for cellular ATP?

12. (a) The degradation of alanine yields pyruvate, and the degradation of leucine yields acetyl CoA. Can the degradation of these amino acids replenish the pool of citric acid cycle intermediates?

(b) Fats (triacylglycerols) stored in adipose tissue are a significant source of energy in animals. Fatty acids are degraded to acetyl CoA, which activates pyruvate carboxylase. How does the activation of this enzyme help recover energy from fatty acids?

13. Amino acids resulting from the degradation of proteins can be further metabolized by conversion to intermediates of the citric acid cycle. If the degradation of a labeled protein leads to the following labeled amino acids, write the structure of the first intermediate of the citric acid cycle into which these amino acids would be converted and identify the labeled carbon in each case.



14. (a) How many molecules of ATP are eventually generated when two molecules of acetyl CoA are converted to four molecules of CO_2 via the citric acid cycle? (Assume NADH 2.5 ATP and QH_2 ~1.5ATP) How many molecules of ATP are generated when two molecules of acetyl CoA are converted to oxaloacetate in the glyoxylate cycle?

(b) How do the yields of ATP relate to the primary functions of the two pathways?

15. The activities of PFK-2 and fructose 2,6-*bis*phosphatase are contained in a bifunctional protein that effects tight control over glycolysis and gluconeogenesis through the action of fructose 2,6-*bis*phosphate. Describe another protein that contains kinase and phosphatase activities in a single protein molecule. What pathways does it control?

Selected Readings

Pyruvate Dehydrogenase Complex

Harris, R. A., Bowker-Kinley, M. M., Huang, B., and Wu, P. (2002). Regulation of the activity of the pyruvate dehydrogenase complex. *Advances in Enzyme Regulation* 42:249–259.

Knoechel, T. R., Tucker, A. D., Robinson, C. M., Phillips, C., Taylor, W., Bungay, P. J., Kasten, S. A., Roche, T. E., and Brown, D. G. (2006). Regulatory roles of the N-terminal domain based on crystal structures of human pyruvate dehydrogenase kinase 2 containing physiological and synthetic ligands. *Biochem.* 45:402–415.

Maeng, C.-Y., Yazdi, M. A., Niu, X.-D., Lee, H. Y., and Reed, L. J. (1994). Expression, purification, and characterization of the dihydrolipoamide dehydrogenase-binding protein of the pyruvate dehydrogenase complex from *Saccharomyces cerevisiae. Biochem.* 33:13801–13807.

Mattevi, A., Obmolova, G., Schulze, E., Kalk, K. H., Westphal, A. H., de Kok, A., and Hol, W. G. J. (1992). Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* 255:1544–1550.

Reed, L. J., and Hackert, M. L. (1990). Structure-function relationships in dihydrolipoamide acyltransferases. *J. Biol. Chem.* 265:8971–8974.

Citric Acid Cycle

Beinert, H., and Kennedy, M. C. (1989). Engineering of protein bound iron–sulfur clusters. *Eur. J. Biochem.* 186:5–15.

Gruer, M. J., Artymiuk, P. J., and Guest, J. R. (1997). The aconitase family: three structural variations on a comon theme. *Trends Biochem. Sci.* 22:3–6.

Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E., Jr., and Stroud, R. M. (1990). Regulation of an enzyme by phosphorylation at the active site. *Science* 249:1012–1016.

Kay, J., and Weitzman, P. D. J., eds. (1987). *Krebs' Citric Acid Cycle—Half a Century and Still Turning* (London: The Biochemical Society).

Krebs, H. A., and Johnson, W. A. (1937). The role of citric acid in intermediate metabolism in animal tissues. *Enzymologia* 4:148–156.

McCormack, J. G., and Denton, R. M. (1988). The regulation of mitochondrial function in mammalian cells by Ca^{2+} ions. *Biochem. Soc. Trans.* 109:523–527.

Remington, S. J. (1992). Mechanisms of citrate synthase and related enzymes (triose phosphate isomerase and mandelate racemase). *Curr. Opin. Struct. Biol.* 2:730–735.

Williamson, J. R., and Cooper, R. H. (1980). Regulation of the citric acid cycle in mammalian systems. *FEBS Lett.* 117 (Suppl.):K73–K85.

Wolodko, W. T., Fraser, M. E., James, M. N. G., and Bridger, W. A. (1994). The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5-Å resolution. *J. Biol. Chem.* 269:10883–10890.

Yankovskaya, V., Horsefield, R., Törnroth, S., Luna-Chavez, C., Miyoshi, H., Léger, C., Byrne, B., Cecchini, G. and Iwata, S. (2003). Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* 299:700–704.

Glyoxylate Cycle

Beevers, H. (1980). The role of the glyoxylate cycle. In *The Biochemistry of Plants: A Comprehensive Treatise*, Vol. 4, P. K. Stumpf and E. E. Conn, eds. (New York: Academic Press), pp. 117–130.





Electron Transport and ATP Synthesis

We now come to one of the most complicated metabolic pathways encountered in biochemistry—the membrane-associated electron transport system coupled to ATP synthesis. The role of this pathway is to convert reducing equivalents into ATP. We usually think of reducing equivalents as products of glycolysis and the citric acid cycle since the oxidation of glucose and acetyl CoA is coupled to the reduction of $NAD^{\oplus}$ and Q. In this chapter we learn that the subsequent reoxidation of NADH and QH_2 results in the passage of electrons through a membrane-associated electron transport system where the energy released can be saved through the phosphorylation of ADP to ATP. The electrons are eventually passed to a terminal electron acceptor. This terminal electron acceptor is usually molecular oxygen (O_2) and this is why the overall process is often called oxidative phosphorylation.

The combined pathway of electron transport and ATP synthesis involves numerous enzymes and coenzymes. It also depends absolutely on the presence of a membrane compartment since one of the key steps in coupling electron transport to ATP synthesis involves the creation of a pH gradient across a membrane. In eukaryotes the membrane is the inner mitochondrial membrane and in prokaryotes it is the plasma membrane.

We begin this chapter with an overview of the thermodynamics of a proton gradient and how it can drive ATP synthesis. We then describe the structure and function of the membrane-associated electron transport complexes and the ATP synthase complex. We conclude with a description of other terminal electron acceptors and a brief discussion of some enzymes involved in oxygen metabolism. Chapter 15 describes the similar membrane-associated electron transport and ATP synthesis pathway that operates during photosynthesis.

According to the chemiosmotic hypothesis of oxidative and photosynthetic phosphorylation proposed by Mitchell, the linkage between electron transport and phosphorylation occurs not because of hypothetical energy-rich chemical intermediates as in the orthodox view, but because oxidoreduction and adenosine triphosphate (ATP) hydrolysis are each separately associated with the net translocation of a certain number of electrons in one direction and the net translocation of the same number of hydrogen atoms in the opposite direction across a relatively ion-, acid-, and base-impermeable coupling membrane.

P. Mitchell, and J. Moyle, (1965)

Top: Sunflowers, cheetahs, and mushrooms all use the same mechanism to make ATP using a proton gradient.

14.1 Overview of Membrane-associated Electron Transport and ATP Synthesis

Membrane-associated electron transport requires several enzyme complexes embedded in a membrane. We will start by examining the pathway that occurs in mitochondria and later we will look at the common features of the prokaryotic and eukaryotic systems. The two processes of membrane-associated electron transport and ATP synthesis are coupled—neither process can occur without the other.

In the common pathway, electrons are passed from NADH to the terminal electron acceptor. There are many different terminal electron acceptors but we are mostly interested in the pathway found in eukaryotic mitochondria where molecular oxygen (O_2) is reduced to form water. As electrons pass along the electron transport chain from NADH to O_2 the energy they release is used to transfer protons from inside the mitochondrion to the intermembrane space between the double membranes. This proton gradient is used to drive ATP synthesis in a reaction catalyzed by ATP synthase (Figure 14.1). A very similar system operates in bacteria.

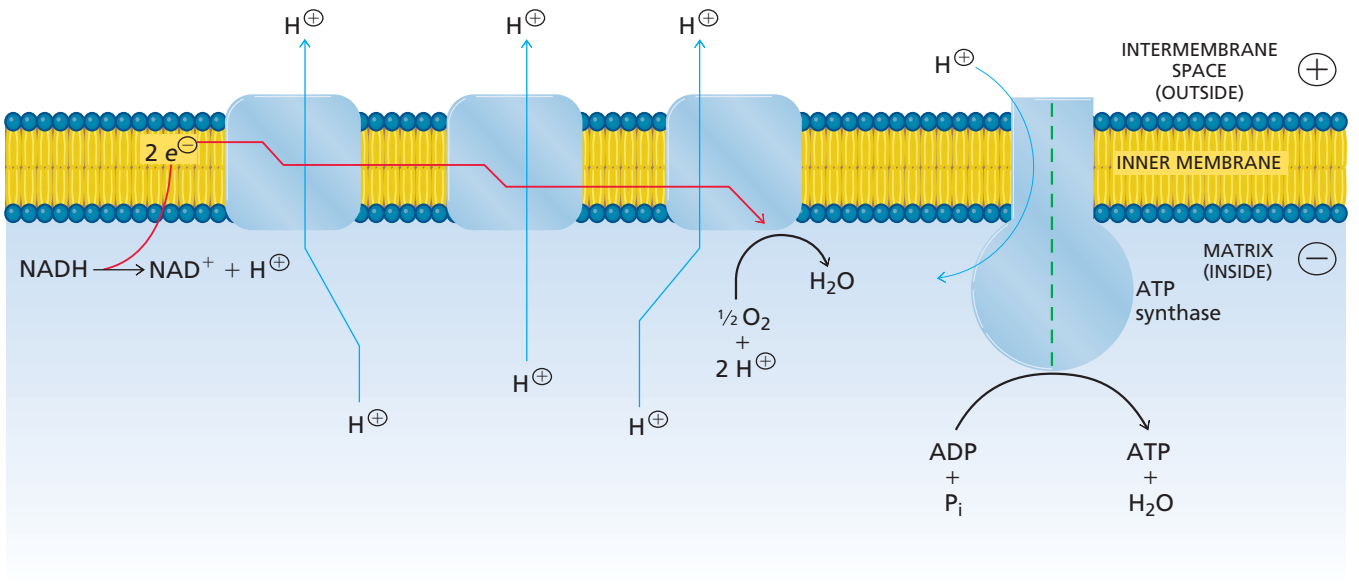
As mentioned above, the entire mitochondrial pathway is often called oxidative phosphorylation because, historically, the biochemical puzzle was to explain the linkage between oxygen uptake and ATP synthesis. You will also see frequent references to "respiration" and "respiratory electron transport." These terms also refer to the pathway that exploits oxygen as the terminal electron acceptor.

The coenzymes mentioned in this chapter are described in detail in Chapter 7: NAD^+, Section 7.4; ubiquinone, Section 7.15; FMN and FAD, Section 7.5; iron–sulfur clusters, Section 7.1; and cytochromes, Section 7.17.

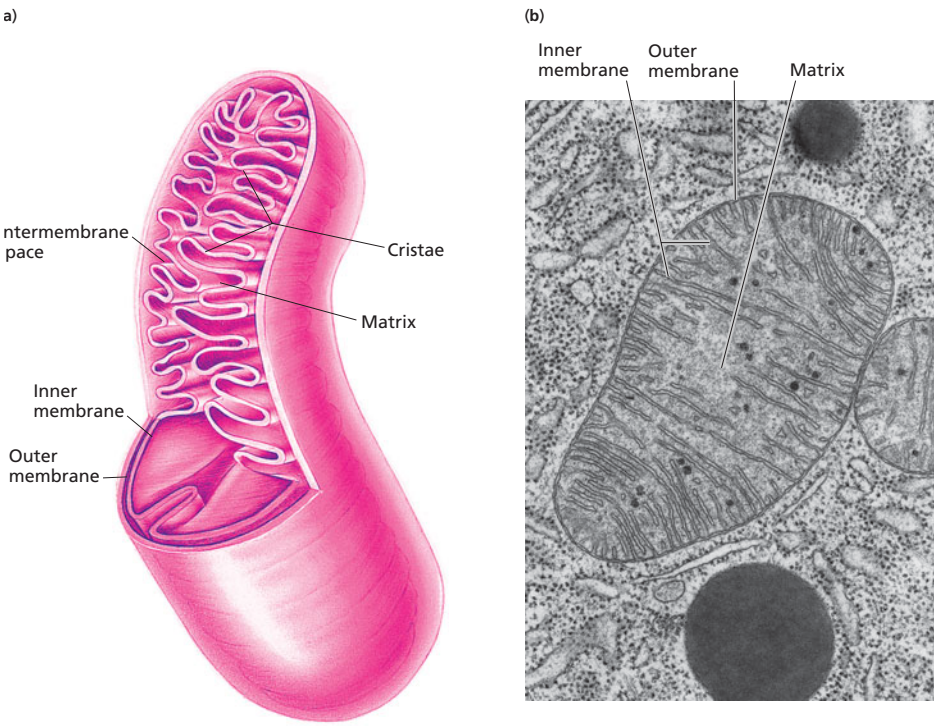
14.2 The Mitochondrion

Much of the aerobic oxidation of biomolecules in eukaryotes takes place in the mitochondrion. This organelle is the site of the citric acid cycle and fatty acid oxidation, both of which generate reduced coenzymes. The reduced coenzymes are oxidized by the electron transport complexes embedded in the mitochondrial membranes. The structure of a typical mitochondrion is shown in Figure 14.2.

The number of mitochondria in cells varies dramatically. Some unicellular algae contain only one mitochondrion whereas the cell of the protozoan *Chaos chaos* contains half a million mitochondria. A mammalian liver cell contains up to 5000 mitochondria. The number of mitochondria is related to the overall energy requirements of the cell. White muscle tissue, for example, relies on anaerobic glycolysis for its energy needs and it contains relatively few mitochondria. The rapidly contracting but swiftly exhausted jaw muscles of the alligator are an extreme example of white muscle. Alligators can snap their jaws with astonishing speed and force but cannot continue this motion beyond a



▼ **Figure 14.1**
Overview of membrane-associated electron transport and ATP synthesis in mitochondria. A proton concentration gradient is produced from reactions catalyzed by the electron transport chain. Protons are translocated across the inner mitochondrial membrane from the matrix to the intermembrane space as electrons from reduced substrates flow through the complexes. The free energy stored in the proton concentration gradient is utilized when protons flow back across the membrane via ATP synthase; their reentry is coupled to the conversion of ADP and P_i to ATP.



◀ **Figure 14.2**
Structure of the mitochondrion. The outer mitochondrial membrane is freely permeable to small molecules but the inner membrane is impermeable to polar and ionic substances. The inner membrane is highly folded and convoluted, forming structures called cristae. The protein complexes that catalyze the reactions of membrane-associated electron transport and ATP synthesis are located in the inner membrane. **(a)** Illustration. **(b)** Electron micrograph: longitudinal section from bat pancreas cell.

very few repetitions. By contrast, red muscle tissue has many mitochondria. The cells of the flight muscles of migratory birds are an example of red muscle cells. These muscles must sustain substantial and steady outputs of power and this power requires prodigious amounts of ATP.

Mitochondria vary greatly in size and shape among different species, in different tissues, and even within a cell. A typical mammalian mitochondrion has a diameter of 0.2 to 0.8 μm and a length of 0.5 to 1.5 μm—this is about the size and shape of an *E. coli* cell. (Recall from Chapter 1 that mitochondria are descendants of bacteria cells that entered into a symbiotic relationship with a primitive eukaryotic cell.)

Mitochondria are separated from the cytoplasm by a double membrane. The two membranes have markedly different properties. The outer mitochondrial membrane has few proteins. One of these proteins is the transmembrane protein porin (Section 9.11A) that forms channels allowing free diffusion of ions and water-soluble metabolites with molecular weights less than 10,000. In contrast, the inner mitochondrial membrane is very rich in protein with a protein-to-lipid ratio of about 4:1 by mass. This membrane is permeable to uncharged molecules such as water, O_2, and CO_2 but it is a barrier to protons and larger polar and ionic substances. These polar substances must be actively transported across the inner membrane using specific transport proteins such as pyruvate translocase (Section 13.4). The entry of anionic metabolites into the negatively charged interior of a mitochondrion is energetically unfavorable. Such metabolites are usually exchanged for other anions from the interior or are accompanied by protons flowing down the concentration gradient that is generated by the electron transport chain.

The inner membrane is often highly folded resulting in a greatly increased surface area. The folds are called cristae. The expansion and folding of the inner membrane also creates a greatly expanded intermembrane space (Figure 14.2a). Since the outer membrane is freely permeable to small molecules, the intermembrane space has about the same composition of ions and metabolites as the cytosol that surrounds the mitochondrion.

The contents of the matrix include the pyruvate dehydrogenase complex, the enzymes of the citric acid cycle (except for the succinate dehydrogenase complex, which is embedded in the inner membrane), and most of the enzymes that catalyze fatty acid oxidation. The protein concentration in the matrix is very high (approaching 500 mg ml^{-1}). Nevertheless, diffusion is only slightly less rapid than in the cytosol (Section 2.3b).



▲ **Alligator jaw muscles.** You're probably safe after this alligator has already snapped at you several times and missed. (If you trust your biochemistry textbook.)



▲ **Canada geese.** If you had more mitochondria in your muscle cells you might be able to fly to a warmer climate for the winter.

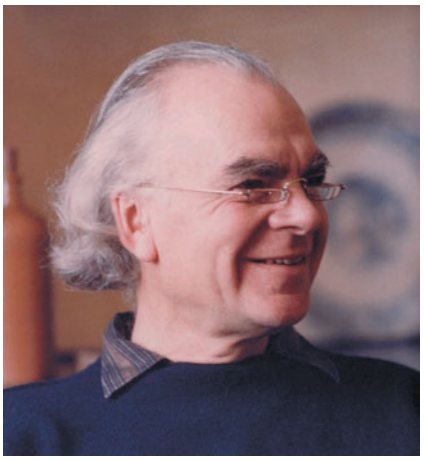
BOX 14.1 AN EXCEPTION TO EVERY RULE

One of the most fascinating things about biology is that there are very few universal rules. We can propose certain general principles that apply in most cases but there are almost always a few examples that don't fit. For example, we can say that eukaryotic cells contain mitochondria as a general rule but we know of some species that don't have mitochondria.

One of the "rules" that seemed valid was that all *animal* cells had mitochondria and they all require oxygen. Now there's even an exception to that rule. Some small microscopic animals of the phylum Loricifera live in deep ocean basins where there is no light and the nearly salt-saturated water is devoid of oxygen. They are incapable of aerobic oxidation and their cells have no mitochondria.



► ***Spinoloricus*** **sp., an anaerobic animal.**



▲ **Peter Mitchell (1920–1992).** Mitchell was awarded the Nobel Prize in Chemistry in 1978 "for his contribution to the understanding of biological energy transfer through the formulation of the chemiosmotic theory." In 1963 Mitchell resigned from his position at Edinburgh University in Scotland and in 1965 he set up a private research institute with his long-time friend and collaborator, Jennifer Moyle. They continued to work on bioenergetics in a laboratory in Mitchell's home, Glynn House, in Cornwall (UK).

KEY CONCEPT

Chemiosmotic theory states that the energy from the oxidation–reduction reactions of electron transport is used to create a proton gradient across the membrane and the resulting protonmotive force is used in the synthesis of ATP.

The matrix also contains metabolites and inorganic ions and a pool of $NAD^{\oplus}$ and $NADP^{\oplus}$ that remains separate from the pyridine nucleotide coenzymes of the cytosol. Mitochondrial DNA and all of the enzymes required for DNA replication, transcription, and translation are located in the matrix. Mitochondrial DNA contains many of the genes that encode the electron transport proteins (see Figure 14.19).

14.3 The Chemiosmotic Theory and Protonmotive Force

Before considering the individual reactions of oxidative phosphorylation we will examine the nature of the energy stored in a proton concentration gradient. The **chemiosmotic theory** is the concept that a proton concentration gradient serves as the energy reservoir that drives ATP formation. The essential elements of this theory were originally formulated by Peter Mitchell in the early 1960s. At the time, the mechanism by which cells carry out oxidative phosphorylation was the subject of intensive research and much controversy. The pathway linking oxidation reactions to the phosphorylation of ADP was not known and many early attempts to identify a "high energy" phosphorylated metabolite that could transfer a phosphoryl group to ADP had ended in failure. Today, thanks to decades of work by many scientists, the formation and dissipation of ion gradients are acknowledged as a central motif in bioenergetics. Mitchell was awarded the Nobel Prize in Chemistry in 1978 for his contribution to our understanding of bioenergetics.

A. Historical Background: The Chemiosmotic Theory

By the time Mitchell proposed the chemiosmotic theory, much information had accumulated on the oxidation of substrates and the cyclic oxidation and reduction of mitochondrial electron carriers. In 1956 Britton Chance and Ronald Williams had shown that when intact isolated mitochondria are suspended in phosphate buffer they oxidize substrates and consume oxygen only when ADP is added to the suspension. In other words, the oxidation of a substrate must be *coupled* to the phosphorylation of ADP. Subsequent experiments showed that respiration proceeds rapidly until all the ADP has been phosphorylated (Figure 14.3a) and that the amount of O_2 consumed depends on the amount of ADP added.

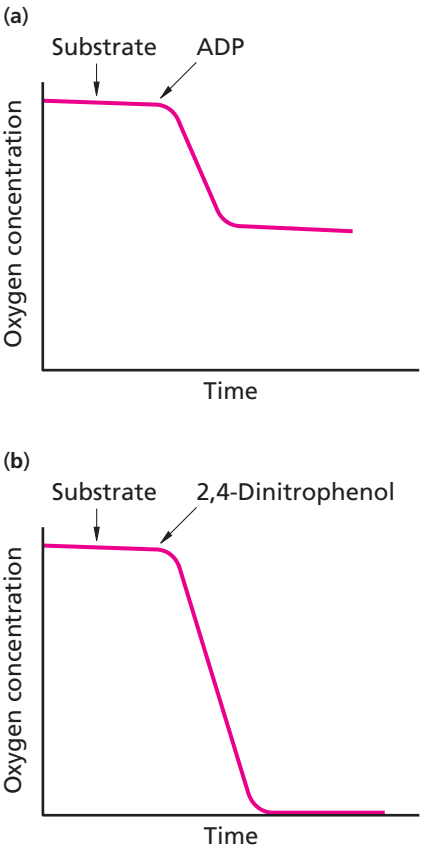
Synthetic compounds called **uncouplers** stimulate the oxidation of substrates in the absence of ADP (Figure 14.3b). The phenomenon of uncoupling helped show how oxidation reactions are linked to ATP formation. In the presence of an uncoupler, oxygen uptake (respiration) proceeds until all the available oxygen is consumed. This rapid oxidation of substrates proceeds with little or no phosphorylation of ADP. In other words, these synthetic compounds uncouple oxidation from phosphorylation. There are many

different kinds of uncouplers and they have little in common chemically except that all are lipid-soluble weak acids. Both their protonated and conjugate base forms can cross the inner mitochondrial membrane—the anionic conjugate base retains lipid solubility because the negative charge is delocalized. The resonance structures of the uncoupler 2,4-dinitrophenol are shown in Figure 14.4.

The effect of uncouplers, and many other experiments, revealed that electron transport (oxygen uptake) and ATP synthesis were normally coupled but the underlying mechanism was unknown. Throughout the 1960s it was commonly believed that there must be several steps in the electron transport process where the Gibbs free energy change was sufficient to drive ATP synthesis. This form of coupling was thought to be analogous to substrate level phosphorylation.

Mitchell proposed that the action of mitochondrial enzyme complexes generates a proton concentration gradient across the inner mitochondrial membrane. He suggested that this gradient provides the energy for ADP phosphorylation via an *indirect* coupling to electron transport. Mitchell's ideas accounted for the effect of the lipid-soluble uncoupling agents—they bind protons in the cytosol, carry them through the inner membrane, and release them in the matrix, thereby dissipating the proton concentration gradient. The proton carriers uncouple electron transport (oxidation) from ATP synthesis because protons enter the matrix without passing through ATP synthase.

ATP synthase activity was first recognized in 1948 as ATPase activity in damaged mitochondria (i.e., damaged mitochondria catalyze hydrolysis of ATP to ADP and P_i). Most workers assumed that mitochondrial ATPase catalyzes the reverse reaction in undamaged mitochondria and this assumption proved to be correct. Efraim Racker and his coworkers isolated and characterized this membrane-bound oligomeric ATPase in the 1960s. The proton driven reversibility of the ATPase reaction was demonstrated by observing the expulsion of protons on hydrolysis of ATP in mitochondria. Further support came from experiments with small membrane vesicles where the enzyme was incorporated into the membrane. When a suitable proton gradient was created across the vesicle membrane, ATP was synthesized from ADP and P_i (Section 14.9).



▲ **Figure 14.3**
Oxygen uptake and ATP synthesis in mitochondria. (a) In the presence of excess P_i and substrate, intact mitochondria consume oxygen rapidly only when ADP is added. Oxygen uptake ceases when all the ADP has been phosphorylated. **(b)** Adding the uncoupler 2,4-dinitrophenol allows oxidation of the substrate to proceed in the absence of phosphorylation of ADP. The arrows indicate the times at which additions were made to the solution of suspended mitochondria.

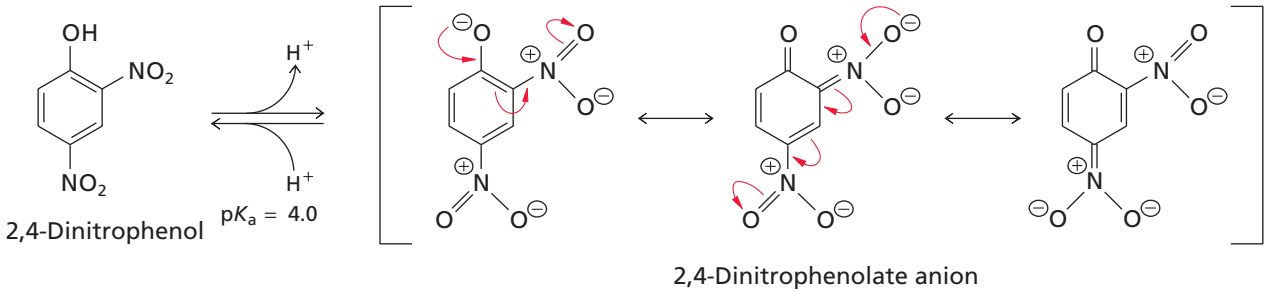
See Box 15.4 for a description of Racker's key experiment.

B. The Protonmotive Force

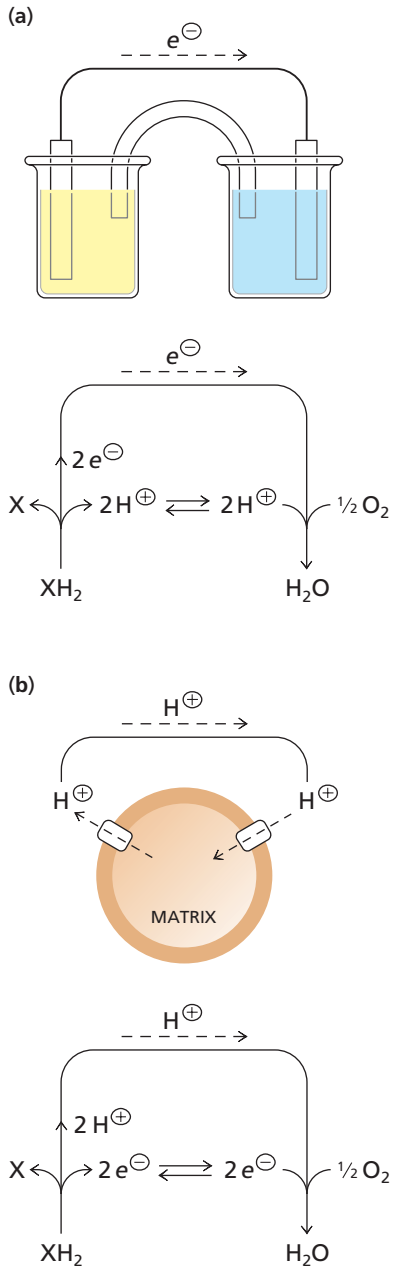
Protons are translocated into the intermembrane space by the membrane-associated electron transport complexes and they flow back into the matrix via ATP synthase. This circular flow forms a circuit that is similar to an electrical circuit. The energy of the proton concentration gradient, called the **protonmotive force**, is analogous to the electromotive force of electrochemistry (Section 10.9A). This analogy is illustrated in Figure 14.5.

Consider a reaction such as the reduction of molecular oxygen by the reducing agent XH_2 in an electrochemical cell.

$$XH_2 + {}^1/_2O_2 \rightleftharpoons X + H_2O \qquad (14.1)$$



▲ **Figure 14.4**
Protonated and conjugate base forms of 2,4-dinitrophenol. The dinitrophenolate anion is resonance stabilized and its negative ionic charge is broadly distributed over the ring structure of the molecule. Because the negative charge is delocalized, both the acid and base forms of dinitrophenol are sufficiently hydrophobic to dissolve in the membrane.



▲ **Figure 14.5**
Electromotive and protonmotive force. (a) In an electrochemical cell, electrons pass from the reducing agent XH_2 to the oxidizing agent O_2 through a wire connecting the two electrodes. The measured electrical potential between cells is the electromotive force. **(b)** When the configuration is reversed (i.e., the external pathway for electrons is replaced by an aqueous pathway for protons), the potential is the protonmotive force. In mitochondria, protons are translocated across the inner membrane when electrons are transported within the membrane by the electron transport chain.

Electrons from XH_2 pass along a wire that connects the two electrodes where the oxidation and reduction half-reactions occur. Electrons flow from the electrode where XH_2 is oxidized

$$XH_2 \rightleftharpoons X + 2\,H^{\oplus} + 2\,e^{\ominus} \tag{14.2}$$

to the electrode where O_2 is reduced.

$$\tfrac{1}{2}O_2 + 2\,H^{\oplus} + 2\,e^{\ominus} \rightleftharpoons H_2O \tag{14.3}$$

In the electrochemical cell, protons pass freely from one reaction cell to the other through the solvent in a salt bridge. Electrons move through an external wire because of a potential difference between the cells. This potential, measured in volts, is the electromotive force. The direction of electron flow and the extent of reduction of the oxidizing agent depend on the difference in free energy between XH_2 and O_2 that in turn depends on their respective reduction potentials.

In mitochondria, it is protons—not electrons—that flow through the external connection, an aqueous circuit connecting the membrane-associated electron transport chain and ATP synthase. This connection is analogous to the wire of the electrochemical reaction. The electrons still pass from the reducing agent XH_2 to the oxidizing agent O_2 but in this case it is through the membrane-associated electron transport chain. The free energy of these oxidation–reduction reactions is stored as the protonmotive force of the proton concentration gradient and is recovered in the phosphorylation of ADP.

Recall from Section 9.10 that the Gibbs free energy change for transport of a charged molecule is

$$\Delta G_{transport} = 2.303\,RT \log \frac{[A_{in}]}{[A_{out}]} + zF\Delta\Psi \tag{14.4}$$

where the first term is the Gibbs free energy due to the concentration gradient and the second term $(zF\Delta\Psi)$ is due to the charge difference across the membrane. For protons the charge per molecule is 1 ($z = 1.0$) and the overall Gibbs free energy change of the proton gradient is

$$\Delta G = 2.303\,RT \log \frac{[H^{\oplus}_{in}]}{[H^{\oplus}_{out}]} + F\Delta\Psi = 2.303\,RT\,\Delta pH + F\Delta\Psi \tag{14.5}$$

This equation can be used to calculate the protonmotive force generated by the proton gradient and the charge difference across the membrane. In liver mitochondria the membrane potential ($\Delta\Psi$) is -0.17 V (inside negative, Section 9.10A) and the pH difference is -0.5 ($\Delta pH = pH_{out} - pH_{in}$). The membrane potential is favorable for movement of protons into the mitochondrial matrix so the $F\Delta\Psi$ term will be negative because $\Delta\Psi$ is negative. The pH gradient is also favorable so the first term in Equation 14.5 must be negative. Thus, the equation for protonmotive force is

$$\Delta G_{in} = F\Delta\Psi + 2.303\,RT\,\Delta pH \tag{14.6}$$

Using the above values at 37° ($T = 310$ K) the available Gibbs free energy is

$$\begin{aligned}\Delta G &= [96485 \times -0.17] + [2.303 \times 8.315 \times 310 \times -0.5] \\ &= -16402\ \text{J mol}^{-1} - 2968\ \text{J mol}^{-1} = -19.4\ \text{kJ mol}^{-1}\end{aligned} \tag{14.7}$$

This means that the transport of a single mole of protons back across the membrane is associated with a free energy change of -19.4 kJ. That's a lot of energy for moving such a small ion!

The standard Gibbs free energy change for the synthesis of one molecule of ATP from ADP is 32 kJ mol^{-1} ($\Delta G^{\circ\prime}$ = 32 kJ mol^{-1}) but the actual Gibbs free energy change is about –48 kJ mol^{-1} (Section 10.6). At least three protons must be translocated in order to drive synthesis of one ATP molecule ($3 \times -19.4 = -58.2$ kJ mol^{-1}) .

Note that 85% ($-16.4/-19.4 = 85\%$) of the Gibbs free energy change is due to the charge gradient across the membrane and only 15% ($-3.0/-19.4 = 15\%$) is due to the proton concentration gradient. Keep in mind that the energy required to *create* the proton gradient is +19.4 kJ mol^{-1}.

KEY CONCEPT

The protonmotive force is due to the combined effect of a charge difference and a proton concentration difference across the membrane.

14.4 Electron Transport

We now consider the individual reactions of the membrane-associated electron transport chain. Four oligomeric assemblies of proteins are found in the inner membrane of mitochondria or the plasma membrane of bacteria. These enzyme complexes have been isolated in their active forms by careful solubilization using detergents. Each complex catalyzes a separate portion of the energy transduction process. The numbers I through IV are assigned to these complexes. Complex V is ATP synthase.

A. Complexes I Through IV

The four enzyme complexes contain a wide variety of oxidation–reduction centers. These may be cofactors such as FAD, FMN, or ubiquinone (Q). Other centers include Fe–S clusters, heme-containing cytochromes, and copper proteins. Electron flow occurs via the sequential reduction and oxidation of these redox centers with flow proceeding from a reducing agent to an oxidizing agent. There are many reactions that involve electron transport processes in biochemistry. We have already seen several of these reactions in previous chapters—the flow of electrons in the pyruvate dehydrogenase complex is a good example (Section 13.1).

Electrons flow through the components of an electron transport chain in the direction of increasing reduction potential. The reduction potentials of each redox center fall between that of the strong reducing agent, NADH, and that of the terminal oxidizing agent, O_2. The mobile coenzymes ubiquinone (Q) and cytochrome *c* serve as links between different complexes of the electron transport chain. Q transfers electrons from complexes I or II to complex III. Cytochrome *c* transfers electrons from complex III to complex IV. Complex IV uses the electrons for the reduction of O_2 to water.

The order of the electron transport reactions is shown in Figure 14.6 against a scale of standard reduction potential on the left and a relative scale of Gibbs standard free energy change on the right. Recall from Section 10.9 that the standard reduction potential (in units of volts) is directly related to the standard Gibbs free energy change (in units of kJ mol^{-1}) by the formula

$$\Delta G^{\circ\prime} = -n\mathrm{F}\ \Delta E^{\circ\prime} \qquad (14.8)$$

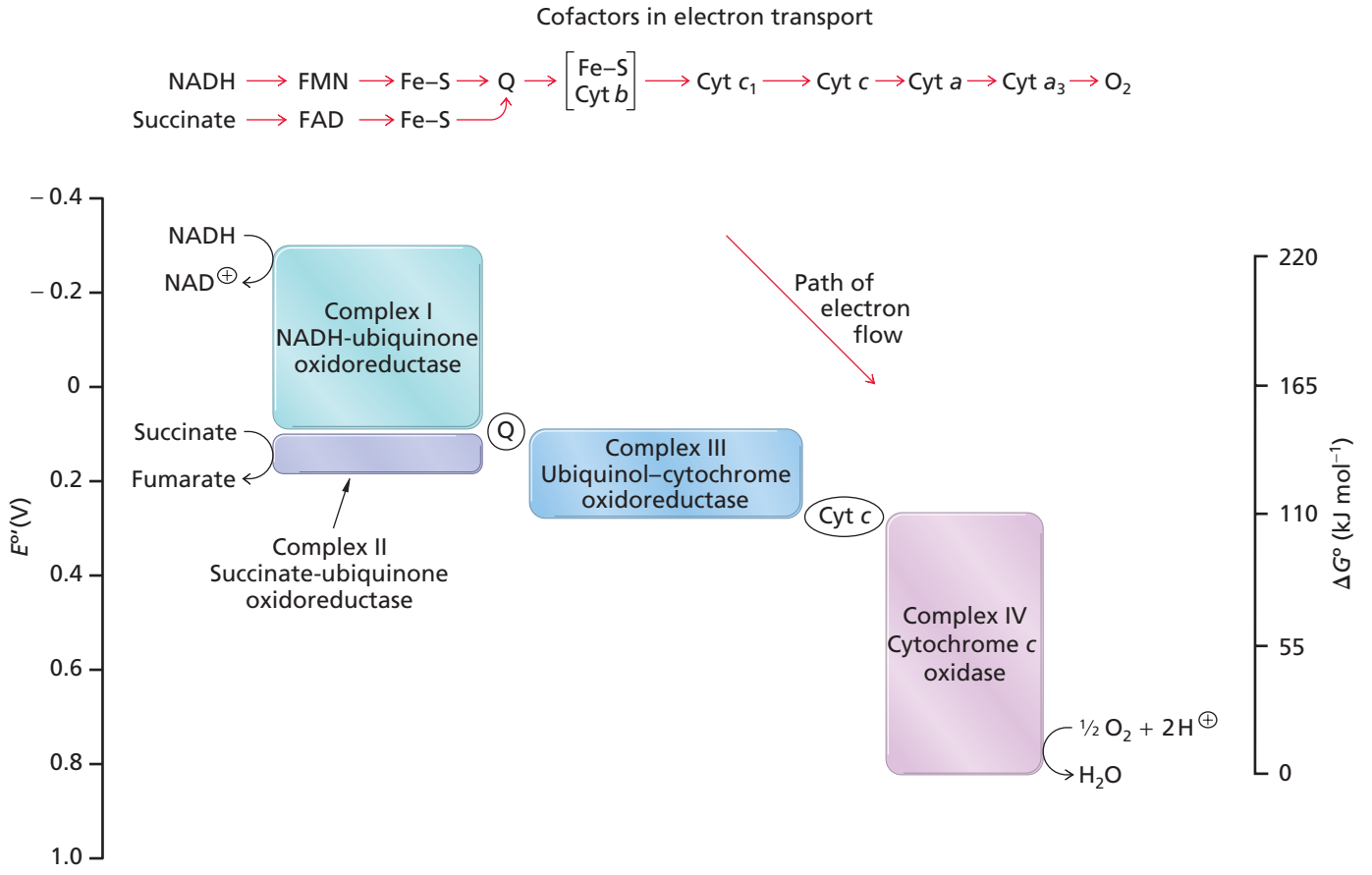
As you can see from Figure 14.6, a substantial amount of energy is released during the electron transport process. Much of this energy is stored in the protonmotive force that drives ATP synthesis. It is this coupling of electron transport to the generation of a protonmotive force that distinguishes membrane-associated electron transport from other examples of electron transport.

The values shown in Figure 14.6 are strictly true only under standard conditions where the temperature is 25°C, the pH is 7.0, and the concentrations of reactants and products are equal (1M each). The relationship between actual reduction potentials (E) and standard ones ($E^{\circ\prime}$) is similar to the relationship between actual and standard free energy (Section 1.4B),

$$E = E^{\circ\prime} - \frac{RT}{n\mathrm{F}} \ln \frac{[\mathrm{S_{red}}]}{[\mathrm{S_{ox}}]} = E^{\circ\prime} - \frac{2.303RT}{n\mathrm{F}} \log \frac{[\mathrm{S_{red}}]}{[\mathrm{S_{ox}}]} \qquad (14.9)$$

The Gibbs Free Energy of Electron Transport

$$E^{\circ\prime} = E^{\circ\prime}_{\text{acceptor}} - E^{\circ\prime}_{\text{donor}} \qquad (10.26)$$
$$= E^{\circ\prime}_{\mathrm{O_2}} - E^{\circ\prime}_{\mathrm{NADH}}$$
$$= +0.82 - (-0.32) \text{ (Table 10.4)}$$
$$= 1.14 \text{ V}$$
$$\Delta G^{\circ\prime} = -n\mathrm{F}\Delta E^{\circ\prime}$$
$$= -2(96485)(1.14)$$
$$= 220 \text{ kJ mol}^{-1}$$



▲ **Figure 14.6**
Electron transport. Each of the four complexes of the electron transport chain, composed of several protein subunits and cofactors, undergoes cyclic reduction and oxidation. The complexes are linked by the mobile carriers ubiquinone (Q) and cytochrome *c*. The height of each complex indicates the $\Delta E°'$ between its reducing agent (substrate) and its oxidizing agent (which becomes the reduced product). Standard reduction potentials are plotted with the lowest value at the top pf the graph (see Section 10.9B).

KEY CONCEPT

Aerobic organisms need oxygen because it serves as the terminal electron acceptor in membrane-associated electron transport.

where $[S_{red}]$ and $[S_{ox}]$ represent the actual concentrations of the two oxidation states of the electron carrier. Under standard conditions, the concentrations of reduced and oxidized carrier molecules are equal; thus, the ratio $[S_{red}]/[S_{ox}]$ is one, and the second term in Equation 14.9 is zero. In this case, the actual reduction potential is equal to the standard reduction potential (at 25°C and pH 7.0). In order for electron carriers to be efficiently reduced and reoxidized in a linear fashion, appreciable quantities of both the reduced and oxidized forms of the carriers must be present under steady state conditions. This is the situation found in mitochondria. We can therefore assume that for any given oxidation–reduction reaction in the electron transport complexes the concentrations of the two oxidation states of the electron carriers are fairly similar. Since physiological pH is close to 7 under most circumstances and since most electron transport processes operate at temperatures close to 25°C, we can safely assume that E is not much different from $E°'$. From now on, our discussion refers only to $E°'$ values.

The standard reduction potentials of the substrates and cofactors of the electron transport chain are listed in Table 14.1. Note that the values progress from negative to positive so that, in general, each substrate or intermediate is oxidized by a cofactor or substrate that has a more positive $E°'$. In fact, one consideration in determining the actual sequence of the electron carriers was their reduction potentials.

The Gibbs standard free energy available from the reactions catalyzed by each complex is shown in Table 14.2. The overall free energy totals −220 kJ mol^{-1} as shown in Figure 14.6. Complexes I, III, and IV translocate protons across the membrane as electrons pass through the complex. Complex II, which is also the succinate dehydrogenase complex we examined as a component of the citric acid cycle, does not directly contribute to formation of the proton concentration gradient. Complex II transfers electrons from succinate to Q and thus represents a tributary of the respiratory chain.

B. Cofactors in Electron Transport

As shown at the top of Figure 14.6, the electrons that flow through complexes I through IV are actually transferred between coupled cofactors. Electrons enter the membrane-associated electron transport chain two at a time from the reduced substrates NADH and succinate. The flavin coenzymes FMN and FAD are reduced in complexes I and II, respectively. The reduced coenzymes $FMNH_2$ and $FADH_2$ donate one electron at a time and all subsequent steps in the electron transport chain proceed by single electron transfers. Iron–sulfur (Fe–S) clusters of both the [2 Fe–2 S] and [4 Fe–4 S] type are present in complexes I, II, and III. Each iron–sulfur cluster can accept or donate one electron as an iron atom undergoes reduction and oxidation between the ferric [Fe^{3+}, Fe(III)] and ferrous [Fe^{2+}, Fe(II)] states. Copper ions and cytochromes are also single electron oxidation–reduction agents.

Several different cytochromes are present in the mammalian mitochondrial enzyme complexes. These include cytochrome b_L, cytochrome b_H, cytochrome c_1, cytochrome *a*, and cytochrome a_3. Very similar cytochromes are found in other species. Cytochromes transfer electrons from a reducing agent to an oxidizing agent by cycling between the ferric and ferrous oxidation states of the iron atoms of their heme prosthetic groups (Section 7.17). Individual cytochromes have different reduction potentials because of differences in the structures of their apoproteins and sometimes their heme groups (Table 14.1). These differences allow heme groups to function as electron carriers at several points in the electron transport chain. Similarly, the reduction potentials of iron–sulfur clusters can vary widely depending on the local protein environment.

The membrane-associated electron transport complexes are functionally linked by the mobile electron carriers ubiquinone (Q) and cytochrome *c*. Q is a lipid-soluble molecule that can accept and donate two electrons, one at a time (Section 7.15). Q diffuses within the lipid bilayer accepting electrons from complexes I and II and passing them to complex III. The other mobile electron carrier is cytochrome *c*, a peripheral membrane protein associated with the outer face of the membrane. Cytochrome *c* carries electrons from complex III to complex IV. The structures and the oxidation–reduction reactions of each of the four electron transport complexes are examined in detail in the following sections.

Table 14.1 Standard reduction potentials of mitochondrial oxidation–reduction components

Substrate of Complex	$E^{\circ\prime}$ (V)
NADH	−0.32
Complex I	
FMN	−0.30
Fe–S clusters	−0.25 to −0.05
Succinate	+0.03
Complex II	
FAD	0.0
Fe–S clusters	−0.26 to 0.00
QH_2/Q	+0.04
($\cdot Q^{\ominus}/Q$	−0.16)
($QH_2/\cdot Q^{\ominus}$	+0.28)
Complex III	
Cytochrome b_1.	−0.01
Cytochrome b_H	+0.03
Fe–S cluster	+0.28
Cytochrome c_1	+0.22
Cytochrome *c*	+0.22
Complex IV	
Cytochrome *a*	+0.21
Cu_A	+0.24
Cytochrome a_3	+0.39
Cu_B	+0.34
O_2	+0.82

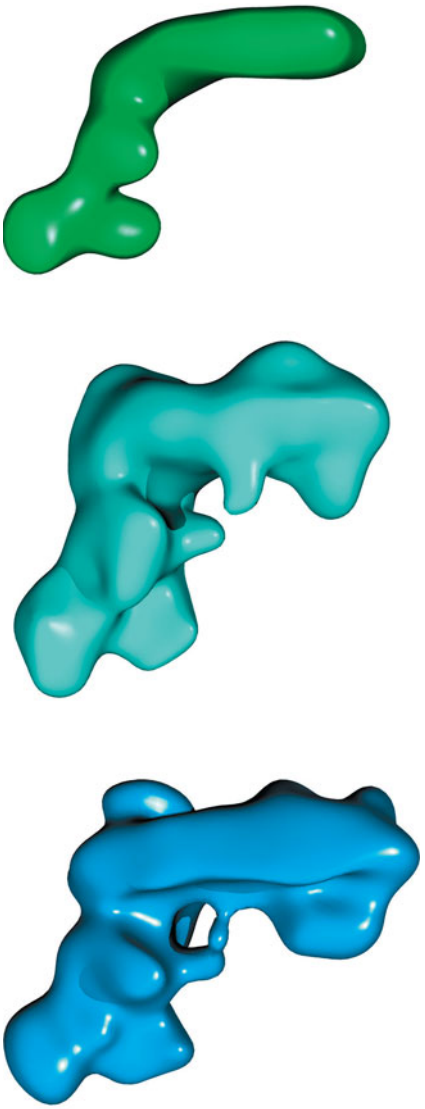
KEY CONCEPT

The transfer of electrons from NADH to O_2 releases enough energy to drive synthesis of many ATP molecules.

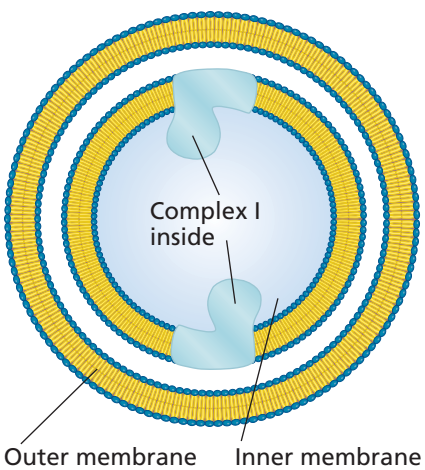
Table 14.2 Standard free energy released in the oxidation reaction catalyzed by each complex

Complex	$E^{\circ\prime}_{reductant}$ (V)	$E^{\circ\prime}_{oxidant}$ (V)	$\Delta E^{\circ\prime}$ [a] (V)	$\Delta G^{\circ\prime}$ [b] (kJ mol^{-1})
I (NADH/Q)	−0.32	−0.04	+0.36	−60
II (Succinate/Q)	+0.03	+0.04	+0.01	−2
III (QH_2/Cytochrome *c*)	+0.04	+0.22	+0.18	−35
IV (Cytochrome *c*/O_2)	+0.22	+0.82	+0.59	−116

[a] $\Delta E^{\circ\prime}$ was calculated as the difference between $E^{\circ\prime}_{reductant}$ and $E^{\circ\prime}_{oxidant}$.
[b] The **Gibbs** standard free energy was calculated using Equation **14.8** where $n = 2$ electrons.



▲ **Figure 14.7**
Structure of complex I. The structures of complex I have been determined at low resolution by analyzing electron micrographic images. **(a)** Complex I from the bacterium *Aquifex aeolicus*. **(b)** Complex I from cow, *Bos taurus*. **(c)** Complex I from the yeast, *Yarrowia lipolytica*.



14.5 Complex I

Complex I catalyzes the transfer of two electrons from NADH to Q. The systematic name of this enzyme is NADH:ubiquinone oxidoreductase. It is a very complicated enzyme whose structure has not been completely solved. The prokaryotic versions contain 14 different polypeptide chains. The eukaryotic forms have 14 homologous subunits plus 20–32 additional subunits, depending on the species. The extra eukaryotic subunits probably stabilize the complex and prevent electron leakage.

The structure of the complex is L-shaped as seen in the electron microscope (Figure 14.7). The membrane-bound component consists of multiple subunits that span the membrane. This module contains a proton transporter activity. A larger component projects into the mitochondrial matrix, or the cytoplasm in bacteria (Figure 14.8). This arm contains a terminal NADH dehydrogenase activity and FMN. The connector module is composed of multiple subunits with 8 or 9 Fe–S clusters (Figure 14.9).

NADH molecules on the inside surface of the membrane donate electrons to complex I. The electrons are passed two at a time as a hydride ion ($H^{\ominus}$, two electrons and a proton). In the first step of electron transfer the hydride ion is transferred to FMN forming $FMNH_2$. $FMNH_2$ is then oxidized in two steps via a semiquinone intermediate. The two electrons are transferred one at a time to the next oxidizing agent, an iron–sulfur cluster.

$$\text{FMN} \xrightarrow{+H^{\oplus},\,+H^{\ominus}} \text{FMNH}_2 \xrightarrow{-H^{\oplus},\,-e^{\ominus}} \text{FMNH}\cdot \xrightarrow{-H^{\oplus},\,-e^{\ominus}} \text{FMN} \qquad \textbf{(14.10)}$$

FMN is a transducer that converts two-electron transfer from NAD-linked dehydrogenases to one-electron transfer for the rest of the electron transport chain. In complex I the cofactor $FMNH_2$ transfers electrons to sequentially linked iron–sulfur clusters. There are at least eight Fe–S clusters positioned within the same arm of complex I that contains the NADH dehydrogenase activity. These Fe–S clusters provide a channel for electrons by directing them to the membrane-bound portion of the complex where ubiquinone (Q) accepts electrons one at a time passing through a semiquinone anion intermediate ($\cdot Q^{\ominus}$) before reaching its fully reduced state, ubiquinol (QH_2).

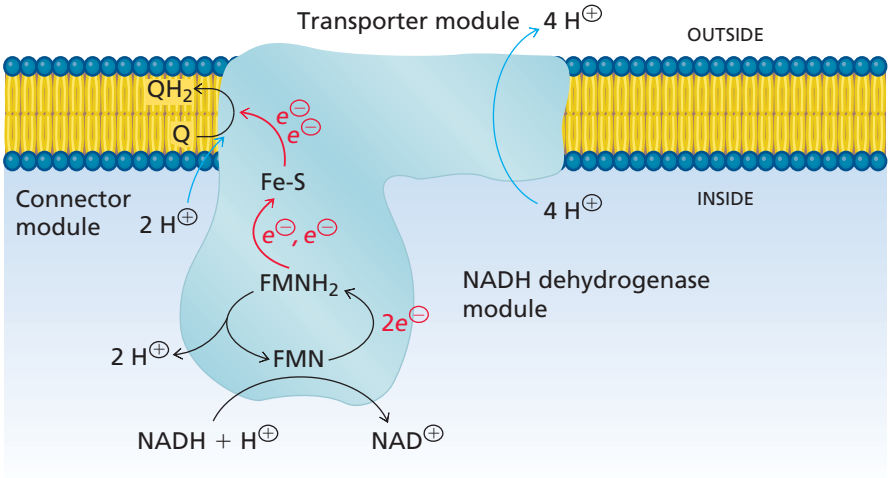
$$Q \xrightarrow{+e^{\ominus}} \cdot Q^{\ominus} \xrightarrow{+e^{\ominus},\,+2\,H^{\oplus}} QH_2 \qquad \textbf{(14.11)}$$

Q and QH_2 are lipid-soluble cofactors. They remain within the lipid bilayer and can diffuse freely in two dimensions. Note that the Q binding site of complex I is within the membrane. One of the reasons for the complicated electron transport chain within complex I is to carry electrons from an aqueous environment to a hydrophobic environment within the membrane.

As electrons move through complex I, two protons (one originating from the hydride ion of NADH and one from the interior) are transferred to FMN to form $FMNH_2$. These two protons or their equivalents are consumed in the reduction of Q to QH_2. Thus, two protons are taken up from the interior and transferred to QH_2. They are not released to the exterior in the complex I reactions. (QH_2 is subsequently reoxidized by complex III and the protons are then released to the exterior. This is part of the proton translocation activity of complex III described in Section 14.7.)

In complex I, four protons are directly translocated across the membrane for every pair of electrons that pass from NADH to QH_2. These do *not* include the protons required for ubiquinone reduction. The proton pump is probably an $H^{\oplus}/Na^{\oplus}$ antiporter

◀ **Figure 14.8**
Complex orientation. The electron transport complexes are embedded in the inner membrane. They can be drawn with the outside of the membrane at the top or at the bottom of the figure. Both views are seen in the scientific literature. We have chosen the orientation with the outside on top and the inside of the matrix on the bottom.



◀ **Figure 14.9**
Electron transfer and proton flow in Complex I. Electrons are passed from NADH to Q via FMN and a series of Fe–S clusters. The reduction of Q to QH_2 requires two protons taken up from the inside compartment. In addition, four protons are translocated across the membrane for each pair of electrons transferred.

located in the membrane-bound module. The mechanism of proton translocation is not clear—it is likely coupled to conformational changes in the structure of complex I as electrons flow from the NADH dehydrogenase site to the ubiquinone binding site.

14.6 Complex II

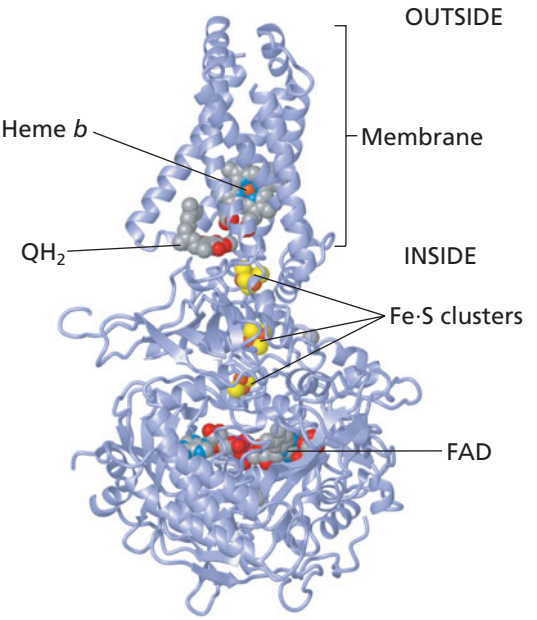
Complex II is succinate:ubiquinone oxidoreductase, also called the succinate dehydrogenase complex. This is the same enzyme that we encountered in the previous chapter (Section 13.3#6). It catalyzes one of the reactions of the citric acid cycle. Complex II accepts electrons from succinate and, like complex I, catalyzes the reduction of Q to QH_2.

Complex II contains three identical multisubunit enzymes that associate to form a trimeric structure that is firmly embedded in the membrane (Figure 14.10). The overall shape resembles a mushroom with its head projecting into the interior of the membrane compartment. Each of the three succinate dehydrogenase enzymes has two subunits forming the head and one or two subunits (depending on the species) forming the membrane-bound stalk. One of the head subunits contains the substrate binding site and a covalently bound flavin adenine dinucleotide (FAD). The other head subunit contains three Fe–S clusters.

The head subunits from all species are closely related and share significant sequence similarity with other members of the succinate dehydrogenase family (e.g., fumarate reductase, Section 14.13). The membrane subunits, on the other hand, may be very different (and unrelated) in various species. In general, the membrane component has one or two subunits that consist exclusively of membrane-spanning α helices. Most of them have a bound heme *b* molecule and this subunit is often called cytochrome *b*. All of the membrane subunits have a Q binding site positioned near the interior surface of the membrane at the point where the head subunits are in contact with the membrane subunits.

The sequence of reactions for the transfer of two electrons from succinate to Q begins with the reduction of FAD by a hydride ion. This is followed by two single electron transfers from the reduced flavin to the series of three iron–sulfur clusters (Figure 14.11). (In those species with a cytochrome *b* anchor, the heme group is not part of the electron transfer pathway.)

Very little free energy is released in the reactions catalyzed by complex II (Table 14.2). This means that the complex cannot contribute directly to the proton concentration gradient across the membrane. Instead, it supplies electrons from the oxidation of succinate midway along the electron transport sequence. Q can accept electrons from complex I or II and donate them to complex III and then to the rest of the electron transport chain. Reactions in several other pathways also donate electrons to Q. We saw one of them, the reaction catalyzed by the glycerol 3-phosphate dehydrogenase complex, in Section 12.2C.



▲ **Figure 14.10**
Structure of the *E. coli* succinate dehydrogenase complex. A single copy of the enzyme showing the positions of FAD, the three Fe–S clusters, QH_2, and heme b. Complex II contains three copies of this multisubunit enzyme. [PDB 1NEK]

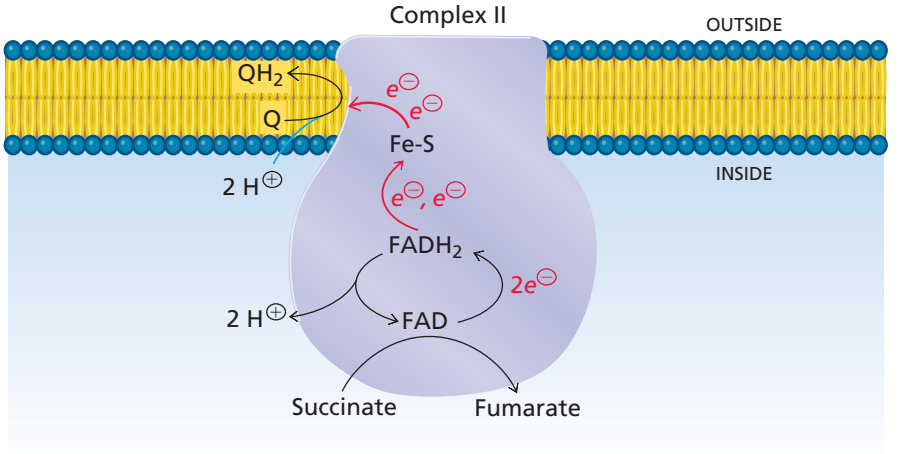


Figure 14.11 ►
Electron transfer in complex II. A pair of electrons is passed from succinate to FAD as part of the citric acid cycle. Electrons are transferred one at a time from $FADH_2$ to three Fe–S clusters and then to Q. (Only one Fe–S cluster is shown in the figure.) Complex II does not directly contribute to the proton concentration gradient but serves as a tributary that supplies electrons (as QH_2) to the rest of the electron transport chain.

14.7 Complex III

Complex III is, arguably, the most important enzyme in metabolism. Very similar complexes are present in chloroplasts where they participate in electron transport and proton translocation during photosynthesis.

Complex III is ubiquinol:cytochrome *c* oxidoreductase, also called the cytochrome bc_1 complex. This enzyme catalyzes the oxidation of ubiquinol (QH_2) molecules in the membrane and the reduction of a mobile water-soluble cytochrome *c* molecule on the exterior surface. Electron transport through complex III is coupled to the transfer of $H^{\oplus}$ across the membrane by a process known as the Q cycle.

The structures of the cytochrome bc_1 complexes from many bacterial and eukaryotic species have been solved by X-ray crystallography. Complex III contains two copies of the enzyme and is firmly anchored to the membrane by a large number of α helices that span the lipid bilayer (Figure 14.12). The functional enzyme consists of three main subunits: cytochrome c_1, cytochrome *b*, and the Rieske iron sulfur protein (ISP) (Figure 14.13). (Note that the cytochrome c_1 subunit is a different protein than the mobile cytochrome *c* product of the reaction.) Other subunits are present on the inside surface but they do not play a direct role in the ubiquinol:cytochrome *c* oxidoreductase reaction. The mobile cytochrome *c* electron acceptor binds at the top of the complex—the part that is exposed to the exterior side of the membrane.

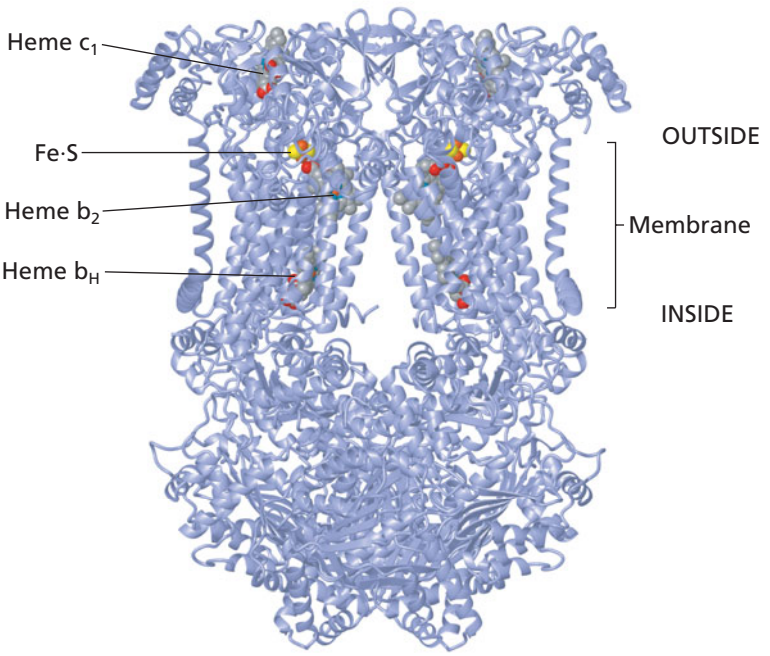


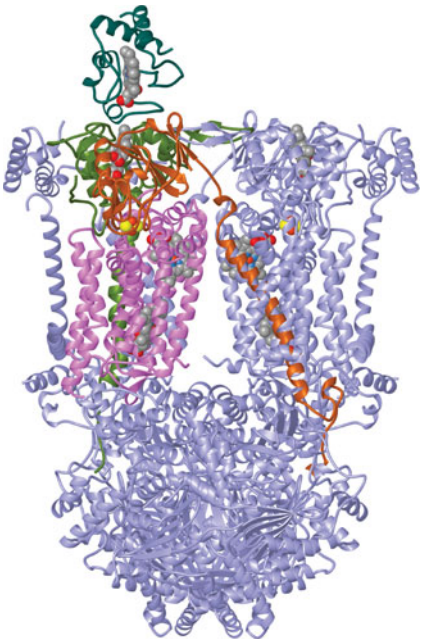
Figure 14.12 ►
Complex III from cow (*Bos taurus*) mitochondria. The cytochrome bc_1 complex contains two copies of the enzyme ubiquinone: cytochrome *c* oxidoreductase. [PDB 1PP9]

The path of electrons through the complex is shown in Figure 14.14. The reaction begins when QH_2 (from complex I or complex II) binds to the Q_0 site in the cytochrome *b* subunit. QH_2 is oxidized to the semiquinone and a single electron is passed to the adjacent Fe–S complex in the ISP subunit. From there, the electron transfers to the heme group in cytochrome c_1. This transfer is facilitated by movement of the head group of ISP. In the electron accepting position, the Fe–S cluster is adjacent to the Q_0 site and in the electron donating position the Fe–S cluster shifts to a position near the heme group in cytochrome c_1. Soluble cytochrome *c* is oxidized by transfer of an electron from the membrane-bound cytochrome c_1 subunit of complex III.

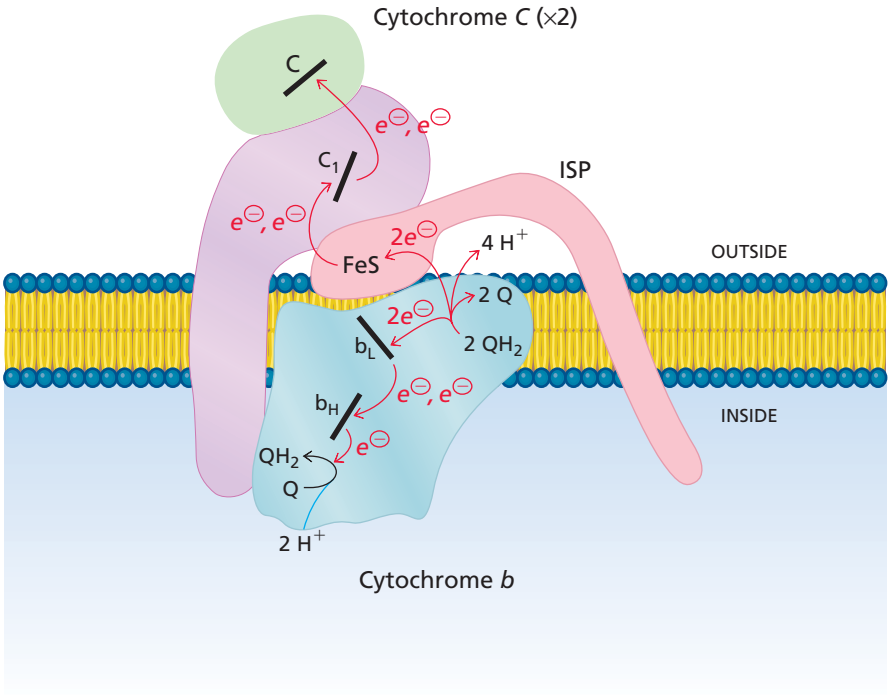
In this reaction, the terminal electron acceptor is cytochrome *c* (Section 7.17). This molecule serves as a mobile electron carrier transferring electrons to complex IV, the next component of the chain (Figure 14.15). The role of reduced cytochrome *c* is similar to that of QH_2, which carries electrons from complex I to complex III. The structures of cytochrome *c* electron carriers from all species are remarkably similar (Section 4.7B, Figure 4.21) and the amino acid sequences of the polypeptide chain are well conserved (Section 3.11, Figure 3.23).

The oxidation of QH_2 at the Q_0 site is a two-step process with a single electron transferred at each step. The path of electrons from the second step, oxidation of the semiquinone intermediate, follows a different route than the first electron. In this case, the electron is passed sequentially to two different *b*-type hemes within the membrane portion of the complex. The first heme group (b_L) has a lower reduction potential and the second heme (b_H) has a higher reduction potential (Table 14.1).

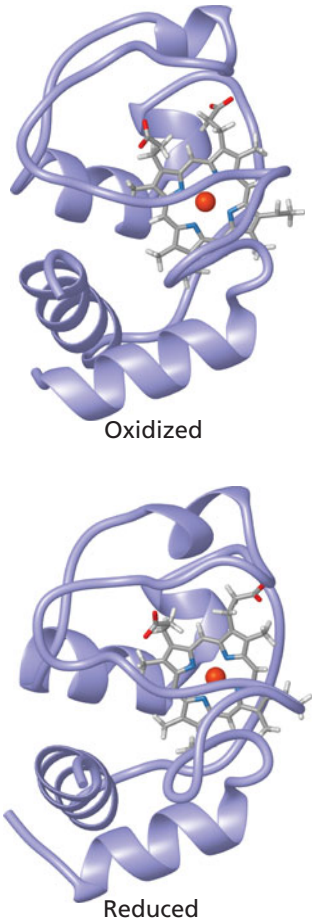
The b_H heme is part of the Q_1 site where a molecule of Q is reduced to QH_2 in a two-step reaction that involves a semiquinone intermediate. A single electron is transported from b_L (at the Q_0 site) to b_H (at the Q_1 site) to Q to produce the semiquinone.



▲ **Figure 14.13**
Subunits of complex III. The three catalytic subunits of each dimer are Cytochrome c_1 (green), cytochrome *b* (blue) and the Rieske iron sulfur protein (ISP) (red). Cytochrome *c* (dark blue) binds to the Cytochrome c_1 subunit. [PDB 1PP9]



▲ **Figure 14.14**
Electron transfer and proton flow in complex III. Two pairs of electrons are passed separately from two molecules of QH_2 at the Q_0 site. Each pair of electrons is split so that individual electrons follow separate pathways. One electron is transferred to an Fe–S cluster then to cytochrome c_1 and finally to cytochrome *c*, the terminal electron carrier. The other electron from each oxidation of QH_2 is transferred to heme b_H (Q_1 site) and then to Q. A total of four protons are translocated across the membrane: two from the inside compartment and two from QH_2. (Only the left-hand half of the dimer is shown and the bottom subunits that project into the matrix are not shown.)



▲ **Figure 14.15**
Cytochrome *c*. Oxidized (top) and reduced (bottom) forms of cytochrome *c* from horse (*Equus caballus*). The iron atom in the center of the heme group (orange) shifts from Fe^{3+} to Fe^{2+} as it gains an electron from complex III. This reduction is accompanied by small changes in the conformation of the protein. [PDB 1OCD (top) 1GIW (bottom)]

KEY CONCEPT

The net effect of the Q cycle is transfer of four protons to the exterior of the membrane for every two electrons transferred from QH_2 to cytochrome *c*.

Table 14.3

Q_0 :	$2\ QH_2 + 2\ \text{cyt}\ c(Fe^{3+}) \longrightarrow 2\ Q + 2\ \text{cyt}\ c(Fe^{2+}) + 2\ e^{\ominus} + 4\ H^{\oplus}_{out}$
Q_1 :	$Q + 2\ H^{\oplus}_{in} + 2\ e^{\ominus} \longrightarrow QH_2$
Sum:	$QH_2 + 2\ \text{cyt}\ c(Fe^{3+}) + 2\ H^{\oplus}_{in} \longrightarrow Q + 2\ \text{cyt}\ c(Fe^{2+}) + 4\ H^{\oplus}_{out}$

Then, a second electron is transferred to reduce the semiquinone to QH_2. The second electron is derived from the oxidation of a second molecule of QH_2 at the Q_0 site. This second oxidation of QH_2 also results in the reduction of a second molecule of cytochrome *c* since the two electrons from the second QH_2 follow separate paths. The net result is that the oxidation of two molecules of QH_2 at the Q_0 site produces two molecules of reduced cytochrome *c* and regenerates a molecule of QH_2 at the Q_1 site. The two cycles of QH_2 oxidation are shown in Figure 14.16. The entire pathway is known as the Q cycle and it is one of the most important reactions in all of metabolism because it is the one most responsible for creating the protonmotive force.

Four protons are produced during the oxidation of two molecules of QH_2 at the Q_0 site. These protons are released to the exterior of the membrane compartment and they contribute to the proton gradient that is formed during membrane-associated electron transport. The protons originate in the interior compartment. They may have been taken up in the reactions catalyzed by complex I or complex II or they may be derived from protons taken up on the inside of the membrane during reduction of Q at the Q_1 site in complex III as shown in Figure 14.16.

The stoichiometry of the complete Q-cycle reaction is shown in Table 14.3. For every pair of electrons that pass through complex III from QH_2 to cytochrome *c* there are four protons translocated across the membrane. Two molecules of cytochrome *c* are reduced and these mobile carriers transport one electron each to complex IV. Note that there are actually two molecules of QH_2 oxidized (giving up four electrons) but two of these electrons are recycled to regenerate a molecule of QH_2.

The complete reaction catalyzed by ubiquinone:cytochrome *c* oxidoreductase (complex III) includes the Q cycle and proton translocation across the membrane. The complex III reaction is a fine example of the relationship between structure and function. While the stoichiometry of the Q cycle had been known for many years, the actual mechanism of the reaction only became apparent once the complete structure was solved in 1998.

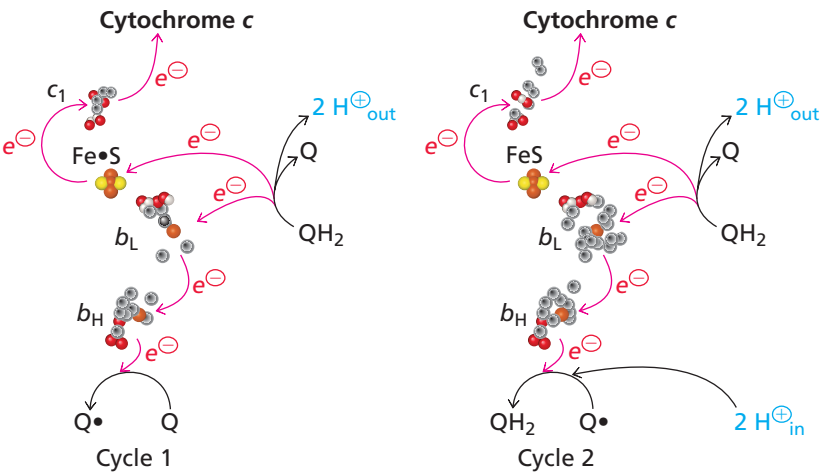


Figure 14.16 ▶
Q cycle. A molecule of QH_2 is oxidized in cycle 1 and a separate molecule is oxidized in cycle 2. Each cycle produces a molecule of reduced cytochrome *c*. The combination of cycle 1 and cycle 2 results in a two-stage reduction of Q to QH_2. Four protons are released on the exterior side of the membrane.

14.8 Complex IV

Complex IV is cytochrome *c* oxidase. This complex catalyzes the oxidation of the reduced cytochrome *c* molecules produced by complex III. The reaction includes a four-electron reduction of molecular oxygen (O_2) to water (2 H_2O) and translocation of four protons across the membrane.

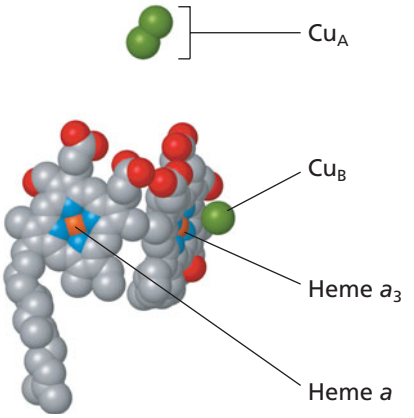
Complex IV contains two functional units of cytochrome *c* oxidase. Each cytochrome *c* oxidase contains single copies of subunits I, II, and III (Figure 14.17). The bacterial enzymes contain only one additional subunit in each functional unit but the eukaryotic (mitochondrial) enzymes have up to ten additional subunits. Additional subunits in the eukaryotic complexes play a role in assembling complex IV and in stabilizing the structure.

The core structure of cytochrome *c* oxidase is formed from the three conserved subunits—I, II, and III. These polypeptides are encoded by mitochondrial genes in all eukaryotes. Subunit I is almost entirely embedded in the membrane. The bulk of this polypeptide consists of 12 transmembrane α helices. There are three redox centers buried within subunit I—two of them are *a*-type hemes (heme *a* and heme a_3) and the third is a copper ion (Cu_B). The copper atom is in close proximity to the iron atom of heme a_3 forming a binuclear center where the reduction of molecular oxygen takes place (Figure 14.18).

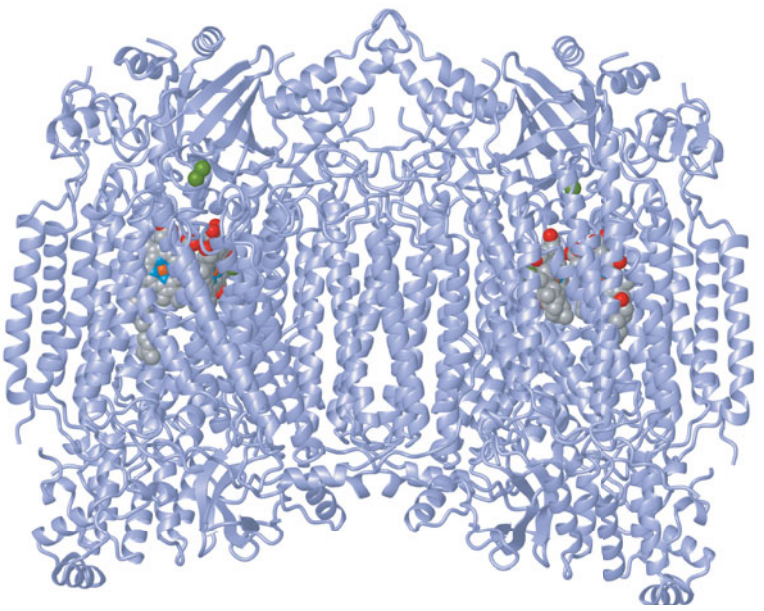
Subunit II has two transmembrane helices that anchor it to the membrane. Most of the polypeptide chain forms a β-barrel domain located on the exterior surface of the membrane. This domain contains a copper redox center (Cu_A) composed of two copper ions. These two copper atoms share electrons forming a mixed valence state. The external domain of subunit II is the site where cytochrome *c* binds to cytochrome *c* oxidase.

Subunit III has seven transmembrane helices and is completely embedded in the membrane. There are no redox centers in subunit III and it can be artificially removed without loss of catalytic activity. Its role *in vivo* is to stabilize subunits I and II and help protect the redox centers from inappropriate oxidation–reduction reactions.

Figure 14.19 shows the sequence of electron transfers in complex IV. Cytochrome *c* binds to subunit II and transfers an electron to the Cu_A site. The pair of copper ions at the Cu_A site can accept and donate one electron at a time—much like an Fe–S cluster. The complete oxidation of O_2 requires four electrons. Thus, four cytochrome *c* molecules have to bind and sequentially transfer a single electron each to the Cu_A redox center.



▲ **Figure 14.18**
Redox centers in cytochrome *c* oxidase. Organization of the heme and copper cofactors in one of the cytochrome *c* oxidase units. [PDB 1OCC]



◀ **Figure 14.17**
Structure of cow (*Bos taurus*) complex IV from mitochondria. The complex consists of two functional units of cytochrome *c* oxidase. Each unit is composed of 13 subunits with multiple membrane-spanning α helices. [PDB 1OCC]

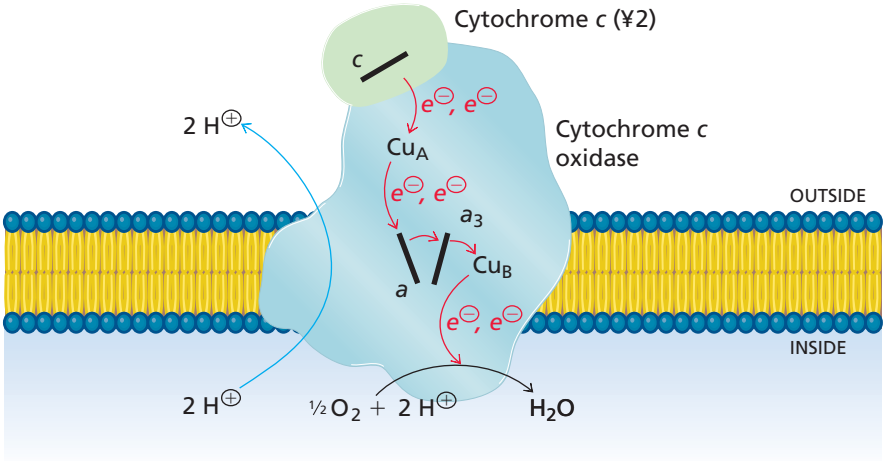


Figure 14.19 ▸
Electron transfer and proton flow in complex IV. The iron atoms of the heme groups in the *a* cytochromes and the copper atoms are both oxidized and reduced as electrons flow from cytochrome *c* to oxygen. Electron transport through complex IV is coupled to the transfer of protons across the membrane. The diagram shows the stoichiometry for transfer of a pair of electrons as in previous figures. The actual reaction involves the transfer of four electrons to a molecule of O_2 to form two molecules of water.

Electrons are transferred one at a time from the Cu_A site to the heme *a* prosthetic group in subunit I. From there they are transferred to the heme a_3–Cu_B binuclear center. The two heme groups (*a* and a_3) have identical structures but differ in their standard reduction potentials due to the local microenvironment formed by surrounding amino acid side chains in subunit I. Electrons can accumulate at the binuclear center as the heme iron alternates between Fe^{3+} and Fe^{2+} states and the copper atom shifts from Cu^{2+} to Cu^{+}. The detailed mechanism for reduction of molecular oxygen at the binuclear center is under active investigation in a number of laboratories. The first step involves the rapid splitting of molecular oxygen. One oxygen atom is bound to the iron atom of the a_3-heme group and the other is bound to the copper atom. Subsequent protonation and electron transfer results in the release of a water molecule from the copper site followed by release of a second water molecule from the iron ligand. The overall reaction requires the uptake of four protons from the inside surface of the membrane

$$O_2 + 4\,e^- + 4\,H^+_{in} \longrightarrow 2\,H_2O \qquad \textbf{(14.12)}$$

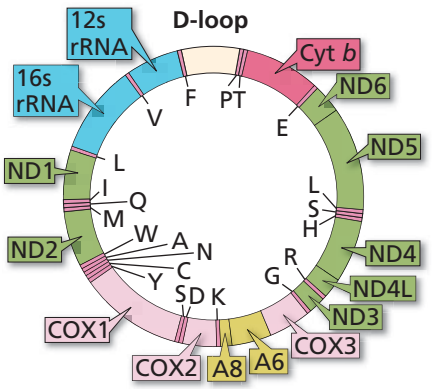
The site where oxygen is reduced is buried within the protein in the middle of the lipid bilayer of the membrane. Charged protons cannot access this site by passive diffusion—instead, the enzyme contains a channel leading from the inside of the membrane to the active site. This channel is filled with a single line of water molecules that rapidly exchange protons leading to the net movement of protons along this "water wire."

The reactions of cytochrome *c* oxidase are coupled to the transfer of protons across the membrane. One proton is translocated for each electron that passes from cytochrome *c* to the final product (H_2O). The protons move through a water-filled channel in complex IV and this movement is driven by conformational changes in the enzyme as oxygen is reduced. The stoichiometry of the complete reaction catalyzed by complex IV is

$$4\ \text{cyt}\ c^{2+} + O_2 + 8\,H^+_{in} \longrightarrow 4\ \text{cyt}\ c^{3+} + 2\,H_2O + 4\,H^+_{out} \qquad \textbf{(14.13)}$$

Complex IV contributes to the proton gradient that will drive ATP synthesis. *Two* protons are translocated for each *pair* of electrons that pass through this complex. Recall that complex I transfers four protons for each pair of electrons and complex III also translocates four protons for each electron pair. Thus, the membrane-associated electron transport system pumps ten protons across the membrane for every molecule of NADH that is oxidized.

The genes encoding the various subunits of the mitochondrial complexes may be in the nucleus or the mitochondria, depending on the species. The genes for cytochrome *c* oxidase subunits are always found in the mitochondrial genome (Figure 14.20).



▲ Figure 14.20
Mitochondrial genome. Mitochondrial genomes are small, circular, double-stranded DNA molecules. They contain genes for ribosomal RNAs (12S rRNA, 16S rRNA) and tRNAs (labeled according to the amino acid they carry). The human mitochondrial genome, shown here, is only 16,589 bp in size and it encodes only a few of the subunits of the electron transport complexes. Genes for the subunits of complex I are colored green, a complex III subunit is purple, complex IV subunits are pink, and complex V subunits are yellow. The D-loop is a highly variable region required for DNA replication. Sequences of individual D-loop regions have been used to trace the evolution of modern humans providing early evidence that we all descend from a population in Africa.

14.9 Complex V: ATP Synthase

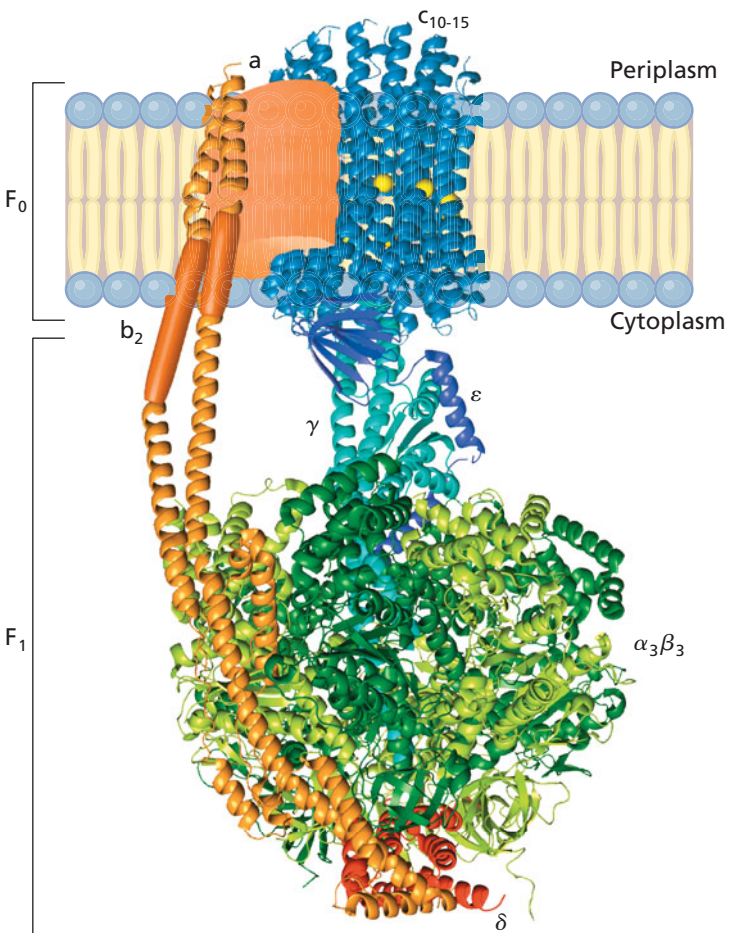
Complex V is ATP synthase. It catalyzes the synthesis of ATP from ADP + P_i in a reaction that is driven by the proton gradient generated during membrane-associated electron transport. ATP synthase is a specific F-type ATPase called F_0F_1 ATPase—named after the reverse reaction. In spite of its name, F-type ATPase is responsible for *synthesizing* ATP—not hydrolyzing it. The enzyme is embedded in the membrane and has a characteristic knob-and-stalk structure that has been observed in electron micrographs for over half a century (Figure 14.21). The F_1 (knob) component contains the catalytic subunits—when released from membrane preparations it catalyzes the hydrolysis of ATP. For this reason, it has traditionally been referred to as F_1 ATPase. This part of the enzyme projects into the mitochondrial matrix in eukaryotes and into the cytoplasm in bacteria. (ATP synthase is also found in chloroplast membranes, as we will see in the next chapter.) The F_0 (stalk) component is embedded in the membrane. It has a proton channel that spans the membrane, and the passage of protons through this channel from the outside of the membrane to the inside is coupled to the formation of ATP by the F_1 component.

Recent cryo electron micrograph structures of ATP synthase have revealed details of its overall structure. These can be correlated with the X-ray crystallographic structures of the various components (Figure 14.22).

The subunit composition of the F_1 component (knob) is $\alpha_3\beta_3\gamma\delta\varepsilon$ and that of the F_0 membrane component is $a_1b_2c_{10–14}$. The c subunits of F_0 interact to form a cylindrical base within the membrane. The core of the F_1 (knob) structure is formed from three copies each of subunits α and β arranged as a cylindrical hexamer. The nucleotide binding sites lie in the clefts between adjacent α and β subunits. Thus, the binding sites are spaced 120° apart on the surface of the $\alpha_3\beta_3$ cylinder. The catalytic site of ATP synthesis is mostly associated with amino acid residues of the β subunit.



▲ **Figure 14.21**
Knobs and stalks. The internal mitochondrial membranes are studded with structures that look like knobs projecting into the mitochondrial matrix at the end of short membrane-embedded stalks.



◀ **Figure 14.22**
ATP synthase structure. The F_1 component is on the inner face of the membrane. The F_0 component, which spans the membrane, forms a proton channel at the interface between the *a* and *c* subunits. The passage of protons through this channel causes the *c* subunit rotor (blue) to rotate relative to the stator of *a* and *b* subunits (orange). The torque of these rotations is transmitted to F_1 where it is used to drive ATP synthesis as the γ subunit (cyan) rotates within the head formed by α and β subunits (green). (The ε subunit is part of the stalk—it lies behind the γ subunit in this view.) (Modified from von Ballmoos et al., 2009.)

The $\alpha_3\beta_3$ oligomer of F_1 is connected to the transmembrane c subunits by a multisubunit stalk made up of the γ and ε subunits. The c-ε-γ unit forms a "rotor" that spins within the membrane. Rotation of the γ subunit inside the $\alpha_3\beta_3$ hexamer alters the conformation of the β subunits, opening and closing the active sites. The a, b, and δ subunits form an arm that also attaches the F_0 component to the $\alpha_3\beta_3$ oligomer. This a-b-δ-$\alpha_3\beta_3$ unit is termed the "stator." Passage of protons through the channel at the interface between the a and c subunits causes the rotor assembly to spin in one direction relative to the stator. The entire structure is often called a molecular motor.

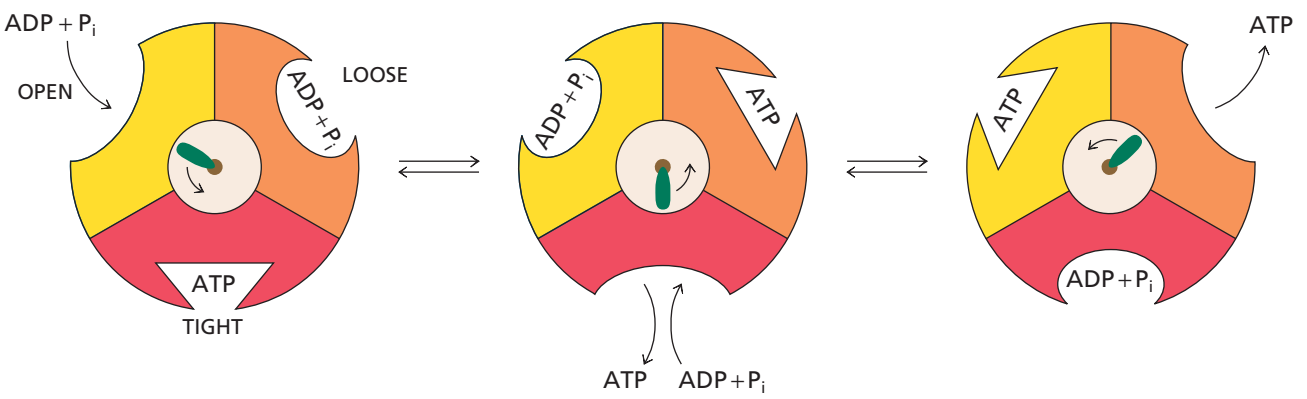
There are 10–14 c subunits in the membrane-associated c-ring at the base of the rotor. The number of subunits depends on the species—yeast and *E. coli* have a 10-subunit ring but plants and animals have up to 14 subunits. There is good evidence to indicate that the rotation of each c subunit past the stator is driven by translocation of a single proton. Rotation of the γ subunit within the F_1 component takes place in a stepwise, jerky manner where each step is 120° of rotation. As the c-ring rotates it twists the γ shaft until enough tension builds up to cause it to snap into the next position within the $\alpha_3\beta_3$ hexamer. If the c-ring has ten subunits then a complete rotation requires translocation of ten protons and results in the production of three ATP molecules but the exact stoichiometry is still being worked out. The results of many experiments indicate that, on average, three protons must be translocated for each ATP molecule synthesized and that's the value that we will use in the rest of this book. It suggests that only nine proton translocations are required for one complete rotation of the c-ring.

V-ATPases have a similar structure. They use ATP hydrolysis to drive the import of protons into acidic vesicles (vacuoles). This is the reverse of the reaction catalyzed by ATP synthase.

The mechanism of ATP synthesis from ADP and P_i has been the target of intensive research for several decades. In 1979 Paul Boyer proposed the *binding change mechanism* based on observations suggesting that the substrate and product binding properties of the active site could change as protons moved across the membrane. The $\alpha_3\beta_3$ oligomer of ATP synthase contains three catalytic sites. At any given time, each site can be in one of three different conformations: (1) open: newly synthesized ATP can be released and ADP + P_i can bind; (2) loose: bound ADP + P_i cannot be released; (3) tight: ATP is very tightly bound and condensation of ADP + P_i is favored. All three sites pass sequentially through these conformations as the γ subunit rotates within the knob. The rate of this reaction is comparable to that of many enzymes. The rotor turns at ten revolutions per second producing 30 ATP molecules per second. Typical turnover numbers (k_{cat}) are in the range of 100–1000 reactions per second.

The formation and release of ATP are believed to occur by the following steps, summarized in Figure 14.23:

1. One molecule of ADP and one molecule of P_i bind to an open site.
2. Rotation of the γ shaft causes each of the three catalytic sites to change conformation. The open conformation (containing the newly bound ADP and P_i) becomes a loose site. The loose site, already filled with ADP and P_i, becomes a tight site. The site containing ATP becomes an open site.
3. ATP is released from the open site and ADP and P_i condense to form ATP in the tight site.



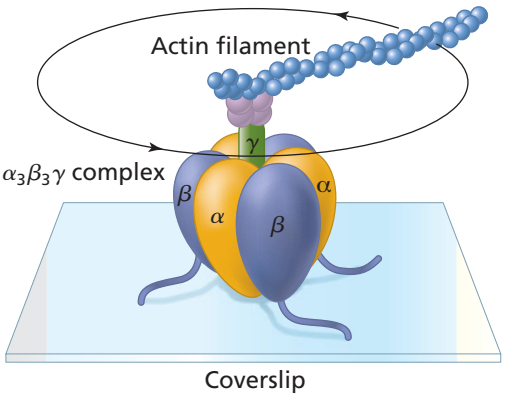
▼ **Figure 14.23**
Binding change mechanism of ATP synthase. The different conformations of the three catalytic sites are indicated by different shapes. ADP and P_i bind to the yellow site in the open conformation. As the γ shaft rotates in the counterclockwise direction (viewed from the cytoplasmic/matrix end of the F_1 component), the yellow site is converted to a loose conformation where ADP and P_i are more firmly bound. Following the next step of the rotation the yellow site is converted to a tight conformation and ATP is synthesized. Meanwhile, the site that had bound ATP tightly has become an open site and a loose site containing other molecules of ADP and P_i has become a tight site. ATP is released from the open site and ATP is synthesized in the tight site.

BOX 14.2 PROTON LEAKS AND HEAT PRODUCTION

Proton leaks appear to be a major consumer of free energy in mammals. In a resting adult mammal, about 90% of oxygen consumption takes place in the mitochondria and about 80% of this is coupled to ATP synthesis. Quantitative estimates indicate that the ATP produced by mitochondria is used for protein synthesis (almost 30% of the available ATP), for active transport of ions by $Na^{\oplus}$—$K^{\oplus}$ ATPase and Ca^{2+} ATPase (25% to 35%), for gluconeogenesis (up to 10%), and for other metabolic processes including heat generation. A significant amount of the energy from oxidation is not used for the synthesis of ATP. In resting mammals, at least 20% of the oxygen consumed by mitochondria is uncoupled by mitochondrial proton leakage. This leakage produces heat directly without apparent use.

The generation of heat in newborns and hibernating animals is a special example of deliberate uncoupling of proton translocation and ATP synthesis. This physiological uncoupling occurs in brown adipose tissue, whose brown color is due to its many mitochondria. Brown adipose tissue is found in abundance in newborn mammals and in species that hibernate. The free energy of NADH is not conserved as ATP but is lost as heat because oxidation is uncoupled from phosphorylation. The uncoupling is due to uncoupling protein 1 (UCP1, thermogenin) that forms a channel for the re-entry of protons into the mitochondrial matrix. When UCP1 is active the free energy released is dissipated as heat, raising the body temperature of the animal.

The strongest evidence that ATP synthase is a rotating motor has been obtained using the $\alpha_3\beta_3\gamma$ complex immobilized on a glass plate and modified by attachment of a fluorescent actin filament (Figure 14.24). Rotation of single molecules was observed by microscopy in the presence of ATP. In this experiment, the labeled γ subunit rotates inside the $\alpha_3\beta_3$ oligomer in response to ATP hydrolysis. This rotation is counterclockwise as depicted in Figure 14.24. Note that rotation driven by ATP hydrolysis is in the opposite direction to that observed when rotation is driven by the proton gradient and ATP is synthesized. The rotation of the γ shaft took place in 120° increments with one step for each ATP molecule hydrolyzed. Under ideal conditions, rates of more than 130 revolutions per second have been observed. This is the expected rotation rate based on the measured rate of ATP hydrolysis. It is much faster than the *in vivo* rate of rotation during ATP synthesis.



▲ **Figure 14.24**
Demonstration of the rotation of a single molecule of ATP synthase. $\alpha_3\beta_3$ complexes were bound to a glass coverslip and the γ subunit was attached to a long fluorescent protein arm. The arms on the molecules rotated when ATP was added. [Adapted from Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr. (1997). Direct observation of rotation of F_1-ATPase. *Nature* 386:299–302.]

14.10 Active Transport of ATP, ADP, and P_i Across the Mitochondrial Membrane

A large fraction of the total ATP synthesized in eukaryotic cells is made in the mitochondria. These molecules must be exported since most of them are used in the cytoplasm. An active transporter is required to allow ADP to enter and ATP to leave mitochondria because the inner mitochondrial membrane is impermeable to charged substances. This transporter is called the adenine nucleotide translocase—it exchanges mitochondrial ATP and cytosolic ADP (Figure 14.25). Normally adenine nucleotides are complexed with Mg^{2+} but this is not the case when they are transported across the membrane. Exchange of ADP^{3-} and ATP^{4-} causes the loss of a net charge of –1 in the matrix. This type of exchange draws on the electrical part of the protonmotive force ($\Delta\Psi$) and some of the free energy of the proton concentration gradient is expended to drive this transport process.

The formation of ATP from ADP and P_i in the mitochondrial matrix also requires a phosphate transporter to import P_i from the cytosol. Phosphate ($H_2PO_4^-$) is transported into mitochondria in electroneutral symport with $H^{\oplus}$ (Figure 14.25). The phosphate carrier does not draw on the electrical component of the protonmotive force but does draw on the concentration difference, ΔpH. Thus, both transporters necessary for ATP formation use up some of the protonmotive force generated by proton translocation. The combined energy cost of transporting ATP out of the matrix and ADP and P_i into it is approximately equivalent to the influx of one proton. Therefore, the synthesis of one molecule of cytoplasmic ATP by ATP synthase requires the

KEY CONCEPT

The chemical energy of the protonmotive force is converted to mechanical energy by causing the rotation of the ATP synthase rotor.

Active transport by ATPases is discussed in Section 9.11D.

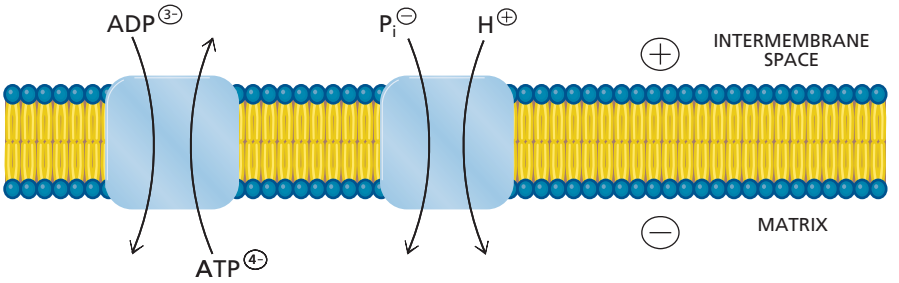


Figure 14.25 ►
Transport of ATP, ADP, and P_i across the inner mitochondrial membrane. The adenine nucleotide translocase carries out unidirectional exchange of ATP for ADP (antiport). Note that the symport of P_i and H^+ is electroneutral.

net influx of four protons from the intermembrane space—one for transport and three that pass through the F_0 component of ATP synthase. Bacteria do not need to transport ATP or ADP across a membrane so the overall expense of ATP synthesis is less than that in eukaryotic cells.

14.11 The P/O Ratio

Before the chemiosmotic theory was proposed, many researchers were searching for a "high energy" intermediate capable of forming ATP by direct phosphoryl group transfer. They assumed that complexes I, III, and IV each contributed to ATP formation with one-to-one stoichiometry. We now know that energy transduction occurs by generating and consuming a proton concentration gradient. The yield of ATP need not be equivalent for each proton translocating electron transport complex nor must the yield of ATP per molecule of substrate oxidized be an integral number.

Many different membrane-associated electron transport complexes contribute simultaneously to the proton concentration gradient. This common energy reservoir is drawn on by many ATP synthase complexes. We saw in the preceding sections that the formation of one molecule of ATP from ADP and P_i catalyzed by ATP synthase requires the inward passage of about three protons and one more proton is needed to transport P_i, ADP, and ATP across the inner membrane.

The first biochemists who studied these processes were primarily interested in the relationship between oxygen consumption (respiration) and ATP synthesis (phosphorylation). The P/O ratio is the ratio of molecules phosphorylated to atoms of oxygen reduced. It takes two electrons to reduce a single atom of oxygen (1/2 O_2) so we are interested in the number of protons translocated for each pair of electrons that pass through complexes I, III, and IV. Four protons are translocated by complex I, four by complex III, and two by complex IV. Thus, for each pair of electrons that pass through these complexes from NADH to O_2 a total of ten protons are moved across the membrane.

KEY CONCEPT

The oxidation of a molecule of NADH results in the synthesis of 2.5 molecules of ATP. In terms of metabolic currency, one NADH molecule is 2.5 ATP equivalents.

Since four protons are moved back across the membrane for each molecule of cytoplasmic ATP, the P/O ratio is $10 \div 4 = 2.5$. The P/O ratio for succinate is only $6 \div 4 = 1.5$ since electrons contributed by succinate oxidation do not pass through complex I. These calculated values are close to the P/O ratios that have been observed in experiments measuring the amount of O_2 reduced when a given amount of ADP is phosphorylated (Figure 14.3a). Recall that the overall energy available in the oxidation–reduction reactions is 220 kJ mol^{-1} (Section 14.4A) and this is more than enough for the synthesis of 2.5 molecules of ATP.

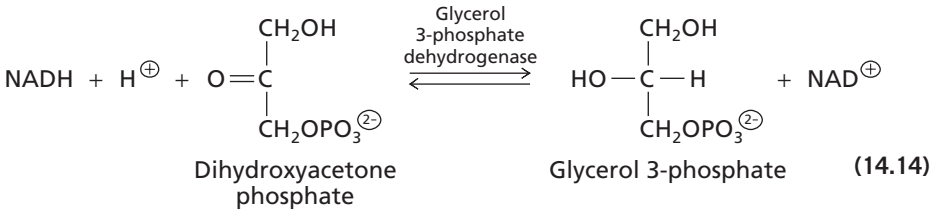
14.12 NADH Shuttle Mechanisms in Eukaryotes

NADH is produced by a variety of different reactions, especially the reactions catalyzed by glyceraldehyde-3-phosphate dehydrogenase during glycolysis and those of the citric acid cycle. NADH can be used directly in biosynthesis reactions such as amino acid synthesis and gluconeogenesis (where glceraldehyde-3-phosphate dehydrogenase operates in the reverse direction).

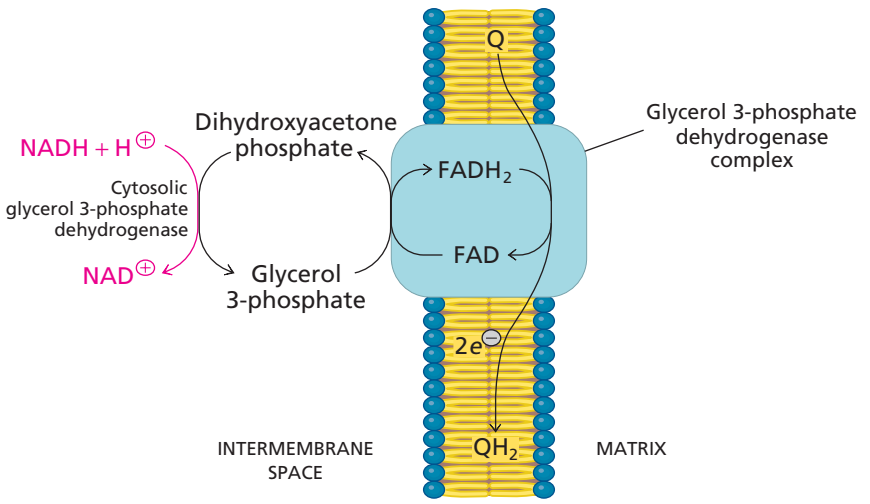
Excess NADH is used to produce ATP by the process that we have described in this chapter. In bacteria, the oxidation of NADH from all sources is readily accomplished since the membrane-associated electron transport system is embedded in the plasma membrane and the inside surface is exposed to the cytosol. In eukaryotic cells on the other hand, the only NADH molecules that have direct access to complex I are those found in the mitochondrial matrix. This is not a problem for reducing equivalents produced by the citric acid cycle since that pathway is localized to the mitochondria. However, the reducing equivalents produced by glycolysis in the cytosol must enter mitochondria in order to fuel ATP synthesis. Because neither NADH nor $NAD^{\oplus}$ can diffuse across the inner mitochondrial membrane, reducing equivalents must enter the mitochondrion by shuttle mechanisms. The glycerol phosphate shuttle and malate–aspartate shuttles are pathways by which a reduced coenzyme in the cytosol passes its reducing power to a mitochondrial molecule that then becomes a substrate for the electron transport chain.

A simplified version of the malate–aspartate shuttle is described in Section 13.4.

The glycerol phosphate shuttle (Figure 14.26) is prominent in insect flight muscles that sustain very high rates of ATP synthesis. It is also present to a lesser extent in most mammalian cells. Two glycerol 3-phosphate dehydrogenases are required—an $NAD^{\oplus}$-dependent cytosolic enzyme and a membrane-embedded dehydrogenase complex that contains an FAD prosthetic group and has a substrate binding site on the outer face of the inner mitochondrial membrane. In the cytosol, NADH reduces dihydroxyacetone phosphate in a reaction catalyzed by cytosolic glycerol 3-phosphate dehydrogenase.



Glycerol 3-phosphate is then converted back to dihydroxyacetone phosphate by the membrane dehydrogenase complex and two electrons are transferred to the FAD prosthetic group of the enzyme. $FADH_2$ transfers two electrons to the mobile electron carrier Q, that then carries the electrons to ubiquinol:cytochrome *c* oxidoreductase (complex III). The oxidation of cytosolic NADH equivalents by this pathway produces less energy (1.5 ATP per molecule of cytosolic NADH) than the oxidation of mitochondrial NADH because the reducing equivalents introduced by the shuttle bypass NADH:ubiquinone oxidoreductase (complex I).



◀ **Figure 14.26**
Glycerol phosphate shuttle. Cytosolic NADH reduces dihydroxyacetone phosphate to glycerol 3-phosphate in a reaction catalyzed by cytosolic glycerol 3-phosphate dehydrogenase. The reverse reaction is catalyzed by an integral membrane flavoprotein that transfers electrons to ubiquinone.

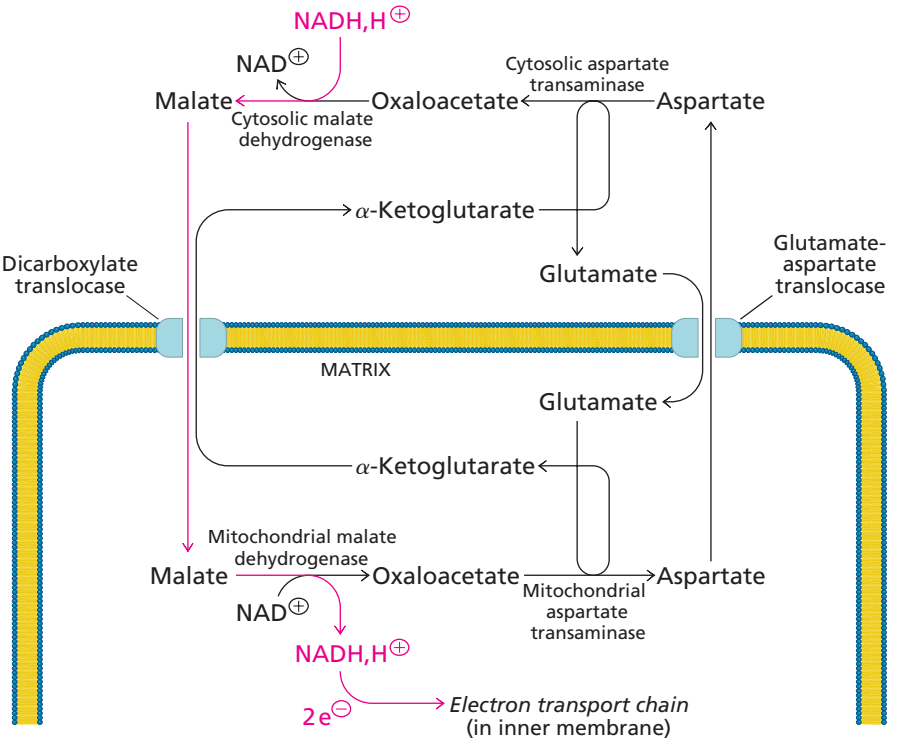
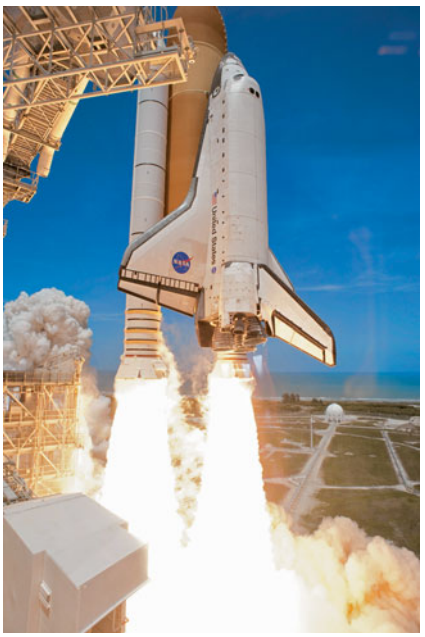


Figure 14.27 ▸
Malate–aspartate shuttle. NADH in the cytosol reduces oxaloacetate to malate that is transported into the mitochondrial matrix. The reoxidation of malate generates NADH that can pass electrons to the electron transport chain. Completion of the shuttle cycle requires the activities of mitochondrial and cytosolic aspartate transaminase.

The malate–aspartate shuttle is more common. This shuttle requires cytosolic versions of malate dehydrogenase—the same enzyme used to convert cytosolic malate to oxaloacetate for gluconeogenesis. The reverse reaction is required for the malate–aspartate shuttle. The operation of the shuttle is diagrammed in Figure 14.27. NADH in the cytosol reduces oxaloacetate to malate in a reaction catalyzed by cytosolic malate dehydrogenase. Malate enters the mitochondrial matrix via the dicarboxylate translocase in electroneutral exchange for α-ketoglutarate. Inside the mitochondria, the citric acid cycle version of malate dehydrogenase catalyzes the reoxidation of malate to oxaloacetate with the reduction of mitochondrial $NAD^{\oplus}$ to NADH. NADH is then oxidized by complex I of the membrane-associated electron transport chain.

Continued operation of the shuttle requires the return of oxaloacetate to the cytosol but oxaloacetate cannot be directly transported across the inner mitochondrial membrane. Instead, oxaloacetate reacts with glutamate in a reversible reaction catalyzed by mitochondrial aspartate transaminase (Section 17.7C). This reaction transfers an amino group to oxaloacetate producing aspartate and α-ketoglutarate. Each molecule of α-ketoglutarate exits the mitochondrion via the dicarboxylate translocase in exchange for malate. Aspartate exits through the glutamate–aspartate translocase in exchange for glutamate. Once they are in the cytosol, aspartate and α-ketoglutarate become the substrates for a cytosolic form of aspartate transaminase that catalyzes the formation of glutamate and oxaloacetate. Glutamate re-enters the mitochondrion in antiport with aspartate and oxaloacetate reacts with another molecule of cytosolic NADH, repeating the cycle.

This complex shuttle system requires several enzymes that have distinctive cytoplasmic and mitochondrial versions (e.g., malate dehydrogenase). As a general rule, these enzymes are encoded by different, but related, genes that are descended from a common ancestor by an ancient gene duplication event. The compartmentalization of metabolic pathways in eukaryotic cells provides them with some advantages over bacterial cells but it requires mechanisms for moving metabolites across internal membranes. Part of the cost of compartmentalization is the duplication of enzymes that need to be present in several compartments. This partly explains why eukaryotic genomes contain so many families of related genes while bacterial genomes usually have only a single copy. One of the striking features of the human genome sequence is the presence of many gene families of this sort. Another major discovery is the presence



▲ **Another kind of shuttle.** This one required a great deal of energy.

BOX 14.3 THE HIGH COST OF LIVING

The average active adult needs about 2400 kilocalories (10,080 kJ) per day. If all of this energy was translated to ATP equivalents, then it would correspond to the hydrolysis of 210 moles of ATP per day. (Assuming that the Gibbs free energy of hydrolysis is 48 kJ mol^{-1}.) This is approximately equal to 100 kg of ATP ($M_r = 507$).

All these ATP molecules have to be synthesized and by far the most common pathway is the synthesis of ATP driven by mitochondrial proton gradients. Actual calculated and measured values suggest that the average person makes 9×10^{20} molecules of ATP per second or 78×10^{24} molecules per day. This is 130 moles or 66 kg of ATP.

Thus, a significant percentage of our calorie intake is converted into a mitochondrial proton gradient in order to drive ATP synthesis. These calculations also tell us that ATP molecules turn over very rapidly since our bodies don't contain 66 kg of ATP.

Rich, P. (2003). The cost of living. *Nature* 421, 583.

of hundreds of genes involved in the translocation of molecules across membranes. The dicarboxylate translocase and glutamate–aspartate translocase described here (Figure 14.27) are examples of transport proteins.

14.13 Other Terminal Electron Acceptors and Donors

Up to this point we have only considered NADH and succinate as important sources of electrons in membrane-associated electron transport. These reduced compounds are mostly derived from catabolic oxidation–reduction reactions such as those in glycolysis and the citric acid cycle. You can imagine that the ultimate source of glucose is a biosynthesis pathway within a photosynthetic organism. The electrons in the chemical bonds of glucose were put there using light energy—the energy from sunlight is ultimately what powers ATP synthesis in mitochondria.

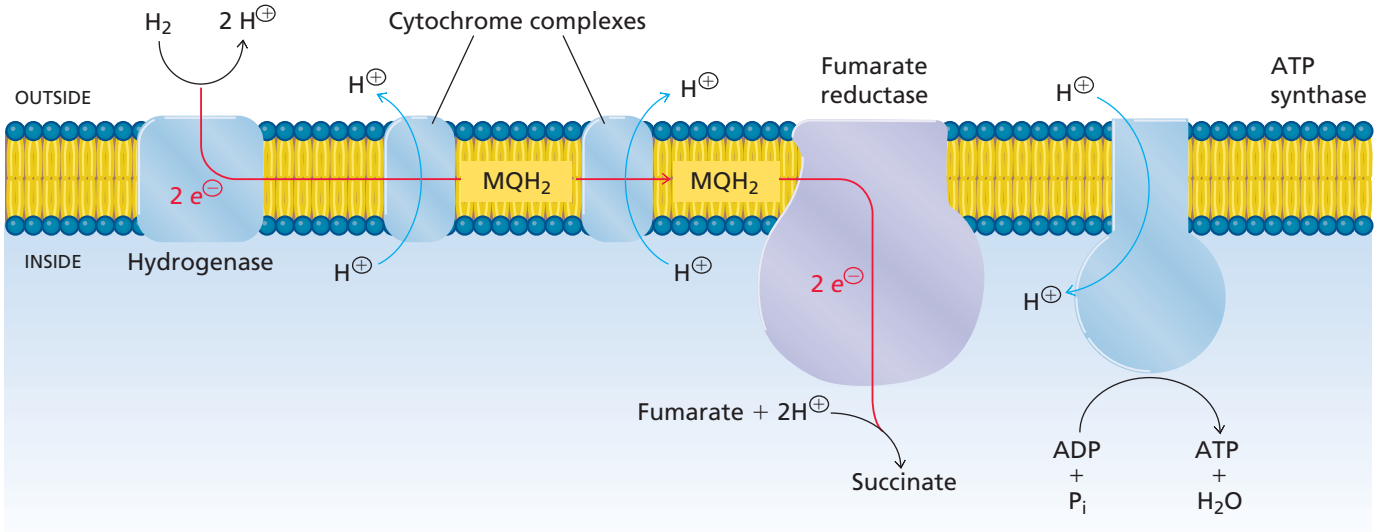
This is a reasonably accurate picture of energy flow in the modern biosphere but it doesn't explain how life survived before photosynthesis evolved. Not only did photosynthesis provide an abundant source of carbon compounds but it is also responsible for the increase in oxygen levels in the atmosphere. As we will see in the next chapter, photosynthesis also requires a membrane-associated electron transport system coupled to ATP synthesis. It's quite likely that respiratory electron transport, as described in this chapter, evolved first and the photosynthesis mechanism came later. There was probably life on this planet for several hundred million years before photosynthesis became commonplace.

What was the ultimate source of energy before sunlight? We have a pretty good idea of how metabolism worked in the beginning because there are still chemoautotrophic bacteria alive today. These species do not need organic molecules as carbon or energy sources and they do not capture energy from sunlight.

Chemoautotrophs derive their energy from oxidizing inorganic compounds such as H_2, $NH_4^{\oplus}$, $NO_2^{\ominus}$, H_2S, S, or Fe^{2+}. These inorganic molecules serve as a direct source of energetic electrons in membrane-associated electron transport. The terminal electron acceptors can be O_2, fumarate, or a wide variety of other molecules. As electrons pass through their electron transport chain a protonmotive force is generated and ATP is synthesized. An example of such a pathway is shown in Figure 14.28.

The electron donor is hydrogen in this example. A membrane-bound hydrogenase oxidizes hydrogen to protons. Such hydrogenases are common in a wide variety of bacteria species. Electrons pass through cytochrome complexes similar to those of respiratory electron transport. In most bacteria, the mobile quinone is not ubiquinone but a related molecule called menaquinone (Section 7.15). Fumarate reductase catalyzes the reduction of fumarate to succinate using reduced menaquinone (MQH_2) as the electron donor.

E. coli can use fumarate instead of oxygen as a terminal electron acceptor when it is growing under anaerobic conditions. Fumarate reductase is a multisubunit enzyme embedded in the plasma membrane. It is homologous to succinate dehydrogenase and the two enzymes catalyze a very similar reaction but in different directions. In *E. coli*,



▲ **Figure 14.28**
One possible pathway for ATP synthesis in chemoautotrophic bacteria. Hydrogen is oxidized by a membrane-bound hydrogenase and electrons are passed through various membrane cytochrome complexes. Electron transfer is coupled to the translocation of protons across the membrane and the resulting protonmotive force is used to drive ATP synthesis. The terminal electron acceptor is fumarate. Fumarate is reduced to succinate by fumarate reductase.

these two enzymes are not expressed at the same time, and *in vivo* each catalyzes its reaction in only one direction (the direction related to the enzyme name). This is one of the few cases where bacterial genomes contain a family of related genes. Each gene encodes a slightly different version of the same enzyme.

In addition to oxygen and fumarate, nitrate and sulfate and many other inorganic molecules can serve as electron acceptors. There are many different combinations of electron donors, acceptors, and electron transport complexes in chemoautotrophic bacteria. The important point is that these bacteria extract energy from inorganic compounds in the absence of light and they may survive without oxygen.

Chemoautotrophic bacteria represent possible metabolic strategies that were present in very ancient organisms but there are still modern bacteria that grow and reproduce in the absence of sunlight and oxygen such as the extreme thermophiles described in Box 2.1 and species that live deep underground.

14.14 Superoxide Anions

One of the unfortunate consequences of oxygen metabolism is the production of reactive oxygen species such as the superoxide radical ($\cdot O_2^{\ominus}$), hydroxyl radical ($OH\cdot$), and hydrogen peroxide (H_2O_2). All of these species are highly toxic to cells. They are produced by flavoproteins, quinones, and iron–sulfur proteins. Almost all of the electron transport reactions produce small amounts of these reactive species, especially $\cdot O_2^{\ominus}$. If a superoxide radical is not rapidly removed by superoxide dismutase it will cause breakdown of proteins and nucleic acids.

We have already discussed superoxide dismutase as an example of an enzyme with a diffusion controlled mechanism (Section 6.4B). The overall reaction catalyzed by this enzyme is the dismutation of two superoxide anions to hydrogen peroxide. This reaction proceeds extremely rapidly.

$$2\ \cdot O_2^{\ominus} + 2\ H^{\oplus} \longrightarrow H_2O_2 + O_2 \qquad \textbf{(14.15)}$$

The rapidity of this process is typical of electron transfer reactions. In this case, a copper ion is the only electron transfer agent bound to the enzyme. The copper ion is reduced by superoxide anion ($\cdot O_2^{\ominus}$), and it then reduces another molecule of $\cdot O_2^{\ominus}$. The hydrogen peroxide formed can be converted to H_2O and O_2 by the action of catalase.

$$2\ H_2O_2 \longrightarrow 2\ H_2O + O_2 \qquad \textbf{(14.16)}$$

Some bacteria species are obligate anaerobes. They die in the presence of oxygen because they cannot deplete reactive oxygen species that arise as a by-product of oxidation–reduction reactions. These species do not have superoxide dismutase. All aerobic species have enzymes that scavenge reactive oxygen molecules.

Summary

1. The energy in reduced coenzymes is recovered as ATP through a membrane-associated electron transport system coupled to ATP synthesis.
2. Mitochondria are surrounded by a double membrane. The electron transport complexes and ATP synthase are embedded in the inner membrane. This inner membrane is highly folded.
3. The chemiosmotic theory explains how the energy of a proton gradient can be used to synthesize ATP. The free energy associated with the protonmotive force is mostly due to the charge difference across the membrane.
4. The electron transport complexes I through IV contain multiple polypeptides and cofactors. The electron carriers are arranged roughly in order of increasing reduction potential. The mobile carriers ubiquinone (Q) and cytochrome *c* link the oxidation–reduction reactions of the complexes.
5. The transfer of a pair of electrons from NADH to Q by complex I contributes four protons to the proton concentration gradient.
6. Complex II does not directly contribute to the proton concentration gradient but rather supplies electrons from succinate oxidation to the electron transport chain.
7. The transfer of a pair of electrons from QH_2 to cytochrome *c* by complex III is coupled to the transport of four protons by the Q cycle.
8. The transfer of a pair electrons from cytochrome *c* and the reduction of 1/2 O_2 to H_2O by complex IV contributes two protons to the gradient.
9. Protons move back across the membrane through complex V (ATP synthase). Proton flow drives ATP synthesis from ADP + P_i by conformational changes produced by the operation of a molecular motor.
10. The transport of ADP and P_i into and ATP out of the mitochondrial matrix consumes the equivalent of one proton.
11. The P/O ratio, the ATP yield per pair of electrons transferred by complexes I through IV, depends on the number of protons translocated. The oxidation of mitochondrial NADH generates 2.5 ATP; the oxidation of succinate generates 1.5 ATP.
12. Cytosolic NADH can contribute to oxidative phosphorylation when the reducing power is transferred to mitochondria by the action of shuttles.
13. Superoxide dismutase converts superoxide radicals to hydrogen peroxide. Hydrogen peroxide is removed by catalase.

Problems

1. In a typical marine bacterium the membrane potential across the inner membrane is –0.15 V. The protonmotive force is –21.2 kJ mol^{-1}. If the pH in the periplasmic space is 6.35, what is the pH in the cytoplasm if the cells are at 25°C?
2. The iron atoms of six different cytochromes in the respiratory electron transport chain participate in one-electron transfer reactions and cycle between the Fe(II) and the Fe(III) states. Explain why the reduction potentials of the cytochromes are not identical but range from −0.10 V to 0.39 V.
3. Functional electron transport systems can be reconstituted from purified respiratory electron transport chain components and membrane particles. For each of the following sets of components, determine the final electron acceptor. Assume O_2 is present.
   (a) NADH, Q, complexes I, III, and IV
   (b) NADH, Q, cytochrome *c*, complexes II and III
   (c) succinate, Q, cytochrome *c*, complexes II, III, and IV
   (d) succinate, Q, cytochrome *c*, complexes II and III
4. A gene has been identified in humans that appears to play a role in the efficiency with which calories are utilized, and anti-obesity drugs have been proposed to regulate the amount of the uncoupling protein-2 (UCP-2) produced by this gene. The UCP-2 protein is present in many human tissues and has been shown to be a proton translocator in mitochondrial membranes. Explain how increasing the presence of the UCP-2 protein might lead to weight loss in humans.
5. (a) When the widely prescribed painkiller Demerol (mepiridine) is added to a suspension of respiring mitochondria, the ratios $NADH/NAD^{\oplus}$ and Q/QH_2 increase. Which electron transport complex is inhibited by Demerol?
   (b) When the antibiotic myxothiazole is added to respiring mitochondria, the ratios cytochrome $c_1(Fe^{3+})$/cytochrome $c_1(Fe^{2+})$ and cytochrome $b_{566}(Fe^{3+})$/cytochrome $b_L(Fe^{2+})$ increase. Where does myxothiazole inhibit the electron transport chain?
6. (a) The toxicity of cyanide ($CN^{\ominus}$) results from its binding to the iron atoms of the cytochrome a,a_3 complex and subsequent inhibition of mitochondrial electron transport. How does this cyanide–iron complex prevent oxygen from accepting electrons from the electron transport chain?
   (b) Patients who have been exposed to cyanide can be given nitrites that convert the Fe^{2+} iron in oxyhemoglobin to Fe^{3+} (methemoglobin). Given the affinity of cyanide for Fe^{3+}, suggest how this nitrite treat mentmight function to decrease the effects of cyanide on the electron transport chain.
7. Acyl CoA dehydrogenase catalyzes the oxidation of fatty acids. Electrons from the oxidation reactions are transferred to FAD and enter the electron transport chain via Q. The reduction potential of the fatty acid in the dehydrogenase-catalyzed reaction is about –0.05 V. Calculate the free energy changes to show why FAD—not $NAD^{\oplus}$—is the preferred oxidizing agent.
8. For each of the following two-electron donors, state the number of protons translocated from the mitochondrion, the number of ATP molecules synthesized, and the P/O ratio. Assume that electrons pass eventually to O_2, NADH is generated in the mitochondrion, and the electron *transport and oxidative phosphorylation systems are fully* functional.
   (a) NADH
   (b) succinate
   (c) ascorbate/tetramethyl-*p*-phenylenediamine (donates two electrons to cytochrome *c*)
9. (a) Why is the outward transport of ATP favored over the outward transport of ADP by the adenine nucleotide transporter?
   (b) Does this ATP translocation have an energy cost to the cell?

10. Atractyloside is a toxic glycoside from a Mediterranean thistle that specifically inhibits the ADP/ATP carrier. Why does atractyloside cause electron transport to be inhibited as well?

11. (a) Calculate the protonmotive force across the inner mitochondrial membrane at 25°C when the electrical difference is –0.18 V (inside negative), the pH outside is 6.7, and the pH inside is 7.5.

(b) What percentage of the energy is from the chemical (pH) gradient, and what percentage is from the charge gradient?

(c) What is the total free energy available for the phosphorylation of ADP?

12. (a) Why does NADH generated in the cytosol and transported into the mitochondrion by the malate–aspartate shuttle produce fewer ATP molecules than NADH generated in the mitochondrion?

(b) Calculate the number of *ATP equivalents produced* from the complete oxidation of one molecule of glucose to six molecules of CO_2 in the liver when the malate–aspartate shuttle is operating. Assume aerobic conditions and fully functional electron transport and oxidative phosphorylation systems.

Selected Readings

Mitochondria

Mentel, M., and Martin, W. (2010). Anaerobic animals from an ancient, anoxic ecological niche. *BMC Biology* 8:32–38.

Taylor, R. W., and Turnbull, D. M. (2005). Mitochondrial DNA mutations in human disease. *Nature Reviews: Genetics* 6:390–402.

Chemiosmotic Theory

Lane, N. (2006) Batteries not included. *Nature* 441:274–277.

Mitchell, P. (1979). Keilin's respiratory chain concept and its chemiosmotic consequences. *Science* 206:1148–1159.

Mitchell, P., and Moyle J. (1965). Stoichiometry of proton translocation through the respiratory chain and adenosine triphosphatase systems of rat liver mitochondria. *Nature* 208:147–151.

Schultz, B., and Chan, S. I. (2001). Structures and proton-pumping strategies of mitochondrial respiratory enzymes. *Annu. Rev. Biophys. Biomol. Struct.* 30:23–65.

Electron Transport Complexes

Berry, E. A., Guergova-Kuras, M., Huang, L., and Crofts, A. R. (2000). Structure and function of cytochrome *bc* complexes. *Annu. Rev. Biochem.* 69:1005–1075.

Brandt, U. (2006). Energy converting NADH: quinone oxidoreductase (complex I). *Annu. Rev. Biochem.* 75:69–92.

Cecchini, G. (2003). Function and structure of Complex II of the respiratory chain. *Annu. Rev. Biochem.* 72:77–100.

Clason, T., Ruiz, T., Schägger, H., Peng, G., Zickerman, V., Brandt, U., Michel, H., and Radermacher, M. (2010). The structure of eukaryotic and prokaryotic complex I. *J. Struct. Biol.* 169:81–88.

Clason, T., Ruiz, T., Schägger, H., Peng, G., Zickerman, V., Brandt, U., Michel, H., and Radermacher, M. (2010). The structure of eukaryotic and prokaryotic complex I. *J. Struct. Biol.* 169:81–88.

Crofts, A. R. (2004). The cytochrome bc_1 complex: function in the context of structure. *Annu. Rev. Physiol.* 66:689–733.

Hosler, J. P., Ferguson-Miller, S., and Mills, D. A. (2006). Energy transduction: proton transfer through the respiratory complexes. *Annu. Rev. Biochem.* 75:165–187.

Hunte, C., Palsdottir, H., and Trumpower, B. L. (2003). Protonmotive pathways and mechanisms in the cytochrome bc_1 complex. *FEBS Letters* 545:39–46.

Hunte, C., Zickerman, V., and Brandt, U. (2010). Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329:448–457.

Richter, O.-M., and Ludwig, B. (2003). Cytochrome *c* oxidase—structure, function, and physiology of a redox-driven molecular machine. *Rev. Physiol. Biochem. Pharmacol.* 147:47–74.

ATP Synthase

Capaldi, R. A., and Aggler, R. (2002). Mechanism of the F_1F_0-type ATP synthase, a biological rotary motor. *Trends in Biochem. Sci.* 27:154–160.

Lau, W. C. Y., and Rubinstein, J. (2010). Structure of intact *Thermus thermophilusV-ATPase* by cryo-EM reveals organization of the membrane-bound V_o motor. *Proc. Natl. Acad. Sci. (USA)* 107:1367–1372.

Nishio, K., Iwamoto-Kihara, A., Yamamoto, A., Wada, Y., and Futai, M. (2002). Subunit rotation of ATP synthase: α or β subunit rotation relative to the *c* subunit ring. *Proc. Natl. Acad. Sci. (USA)* 99:13448–13452.

Oster, G., and Wang, H. (2003). Rotary protein motors. *Trends in Cell Biology* 13:114–121.

Other Electron Donors and Acceptors

Hederstedt, L. (1999). Respiration without O_2. *Science* 284:1941–1942.

Iverson, T. M., Luna-Chavez, C., Cecchini, G., and Rees, D. C. (1999). Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science* 284:1961–1966.

Peters, J. W., Lanzilotta, W. N., Lemon, B. J., and Seefeldt, L. C. (1998). X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 Ångstrom resolution. *Science* 282:1853–1858.

Tielens, A. G. M., Rotte, C., van Hellemond, J. J., and Martin, W. (2002). Mitochondria as we don't know them. *Trends in Biochem. Sci.* 27:564-572.

von Ballmoos, C., Cook, G. M., and Dimroth, P. (2008). Unique rotary ATP synthase and its biological diversity. *Annu. Rev. Biophys.* 37:43–64.

von Ballmoos, C., Wiedenmann, A., and Dimroth, P. (2009). Essentials for ATP synthesis by F_1F_0 ATP synthases. *Annu. Rev. Biochem.* 78:649–672.

Yankovskaya, V., Horsefield, R., Törnroth, S., Luna-Chavez, C., Miyoshi, H., Léger, C., Byrne, B., and Iwata, S. (2003). Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* 299:700–704.





Photosynthesis

Why does this particular group of radiations, rather than some other, make the leaves grow and the flowers burst forth, cause the mating of fireflies and the spawning of palolo worms, and, when reflecting off the surface of the moon, excite the imagination of poets and lovers?

Helena Curtis and Sue Barnes (1989). *Biology*, 5th ed.

The most important part of photosynthesis is the conversion of light energy into chemical energy in the form of ATP. The basic principle behind this fundamental reaction is similar to that of membrane-associated electron transport covered in the previous chapter. In photosynthesis, light shines on a pigment molecule (e.g., chlorophyll) and an electron is excited to a higher energy level. As the electron falls back to its initial state it gives up energy and this energy is used to translocate protons across a membrane. This creates a proton gradient that is used to drive phosphorylation of ADP in a reaction catalyzed by ATP synthase. In some cases, reducing equivalents in the form of NADPH are synthesized directly when the excited electron is used to reduce $NADP^{\oplus}$. These reactions are called the **light reactions** since they are absolutely dependent on sunlight.

Photosynthetic species use their abundant supply of cheap ATP and NADPH to carry out all of the metabolic reactions that require energy. This includes synthesis of proteins, nucleic acids, carbohydrates, and lipids. This is why photosynthetic bacteria and algae are such successful organisms.

Most photosynthetic organisms have a special CO_2 fixing pathway called the *Calvin cycle*. Strictly speaking, the fixation of CO_2 does not require light and is not directly coupled to the light reactions. For this reason, these reactions are often called the **dark reactions** but this does not mean they take place in the dark. This pathway is closely related to the pentose–phosphate pathway described in Section 12.4.

The details of photosynthesis reactions are extremely important in understanding the biochemistry of all life on the planet. The ability to harvest light energy to synthesize macromolecules led to a rapid expansion of photosynthetic organisms. This, in turn, created opportunities for species that could secondarily exploit photosynthetic organisms as food sources. Animals, such as us, ultimately derive much of their energy by degrading molecules that were originally synthesized using the energy from sunlight. In addition, oxygen is a by-product of photosynthesis in plants and some bacteria. The buildup of oxygen in Earth's atmosphere led to its role as an electron acceptor in membrane-associated electron transport. With few exceptions, modern eukaryotes now absolutely depend on the supply of oxygen produced by photosynthesis in order to synthesize ATP in their mitochondria.

Top: Sunlight on trillium in the woods. Solar energy captured by photosynthetic organisms ultimately sustains the activities of nearly all organisms on Earth.



▲ **Photosynthetic organisms.** Left: cyanobacteria. Middle: leaves of a flowering plant. Right: purple bacteria.

The major components of the photosynthesis reactions are large complexes of proteins, pigments, and cofactors embedded in a membrane. A complex containing the light-sensitive pigments is called a **photosystem**. Different species employ a variety of different strategies to utilize light energy in order to synthesize ATP and/or NADPH. We will first describe the structure and function of photosystems in bacteria and then move on to the more complex photosynthesis pathway in eukaryotes such as algae and plants. The eukaryotic photosynthesis complexes clearly evolved from the simple bacterial ones.

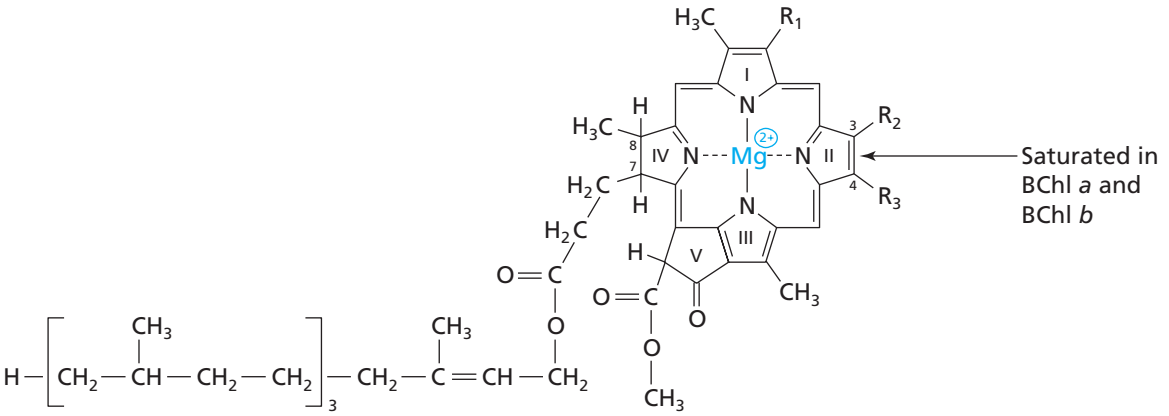
15.1 Light-Gathering Pigments

There are several kinds of light-gathering pigments. They have different structures, different properties, and different functions.

A. The Structures of Chlorophylls

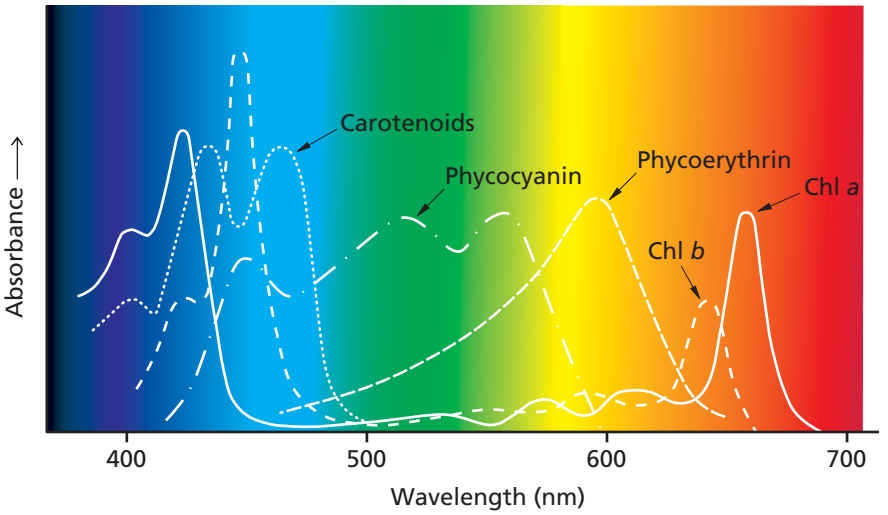
Chlorophylls are the most important pigments in photosynthesis. The structures of the most common chlorophyll molecules are shown in Figure 15.1. Note that the tetrapyrrole ring of chlorophylls is similar to that of heme (Figure 7.38) except that the chlorophyll ring is reduced—it has one less double bond in the conjugated ring system between position 7 and 8 in ring IV. Chlorophylls contain a central chelated Mg^{2+} ion instead of the Fe^{2+} found in heme. Another distinguishing feature of chlorophylls is that they possess a long phytol side chain that contributes to their hydrophobicity.

There are many different types of chlorophylls. They differ mostly in the side chains labeled R_1, R_2, and R_3 in Figure 15.1. Chlorophyll *a* (Chl *a*) and chlorophyll *b*



Chl species	R_1	R_2	R_3
Chl *a*	$-CH=CH_2$	$-CH_3$	$-CH_2-CH_3$
Chl *b*	$-CH=CH_2$	$-C(=O)-H$	$-CH_2-CH_3$
BChl *a*	$-C(=O)-CH_3$	$-CH_3$	$-CH_2-CH_3$
BChl *b*	$-C(=O)-CH_3$	$-CH_3$	$-CH=CH-CH_3$

▲ **Figure 15.1**
Structures of chlorophyll and bacteriochlorophyll pigments. Differences in substituent groups indicated as R_1, R_2 and R_3 are shown in the table. In the bacteriochlorophylls, the double bond indicated in ring II is saturated. In some molecules of bacteriochlorophyll *a*, the phytol side chain has three additional double bonds. The hydrophobic phytol side chain and hydrophilic porphyrin ring give chlorophyll amphipathic characteristics. Chlorophyll (bound to proteins) is found in photosystems and in associated light-harvesting complexes.



◂ **Figure 15.2**
Absorption spectra of major photosynthetic pigments. Collectively, the pigments absorb radiant energy across the spectrum of visible light.

(Chl *b*) are found in a large number of species. Bacteriochlorophyll *a* (BChl *a*) and bacteriochlorophyll *b* (BChl *b*) are only found in photosynthetic bacteria. They differ from the other chlorophylls because they have one less double bond in ring II. Pheophytin (Ph) and bacteriopheophytin (BPh) are similar pigments where the Mg^{2+} in the central cavity is replaced by two covalently bound hydrogens.

Chlorophyll molecules are specifically oriented in the membrane by noncovalent binding to integral membrane proteins. The hydrophobic phytol side chain helps anchor chlorophyll in the membrane. The light-absorbing ability of chlorophyll is due to the tetrapyrrole ring with its network of conjugated double bonds. Chlorophylls absorb light in the violet-to-blue region (absorption maximum 400 to 500 nm) and the orange-to-red region (absorption maximum 650 to 700 nm) of the electromagnetic spectrum (Figure 15.2). This is why chlorophylls are green—that's the part of the spectrum that is reflected, not absorbed. The exact absorption maxima of chlorophylls depend on their structures; for example, Chl *a* differs from Chl *b*. The absorption maxima of particular chlorophyll molecules is also affected by their microenvironment within the pigment–protein complex.

KEY CONCEPT

Chlorophyll molecules are oxidized (loss of an electron) when they absorb a photon of light.

B. Light Energy

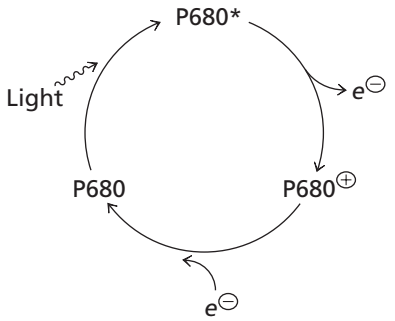
A single quantum of light energy is called a **photon**. When a chlorophyll molecule absorbs a photon, an electron from a low energy orbital in the pigment is promoted to a higher energy molecular orbital. The energy of the absorbed photon must match the difference in energy between the ground state and higher energy orbitals—this is why chlorophyll absorbs only certain wavelengths of light. The excited "high energy" electron can be transferred to nearby oxidation–reduction centers in the same way that "high energy" electrons can be transferred from NADH to FMN in complex I during respiratory electron transport (Section 14.5). The main difference between photosynthesis and respiratory electron transport is the source of excited electrons. In respiratory electron transport the electrons are derived from chemical oxidation–reduction reactions that produce NADH and QH_2. In photosynthesis the electrons are directly promoted to a "high energy" state by absorption of a photon of light.

Chlorophyll molecules can exist in three different states. In the ground state (Chl or Chl^0), all electrons are at their normal stable level. In the excited state (Chl^*) a photon of light has been absorbed. Following electron transfer, the chlorophyll molecule is in the oxidized state ($Chl^{\oplus}$) and must be regenerated by receiving an electron from an electron donor.

The energy of a photon of light can be calculated from the following equation

$$E = \frac{hc}{\lambda} \qquad \textbf{(15.1)}$$

where h is Planck's constant (6.63×10^{-34} J s), c is the velocity of light (3.00×10^{8} m s^{-1}), and λ is the wavelength of light. It's often convenient to calculate the total energy of a



▲ **The states of chlorophyll.** Reduction, excitation, and oxidation of chlorophyll P680. P680* is the excited state following absorption of a photon of light. Loss of an electron produces the oxidized state, P680⊕. Gain of an electron from an outside source (such as the oxidation of water) yields the reduced P680 state.

The Gibbs free energy change associated with the protonmotive force is calculated in Section 14.3B

"mole" of photons by multiplying E by 6.022×10^{23} (Avogadro's number). Thus, for light at a wavelength of 680 nm, the energy is 176 kJ mol^{-1}. This is similar to a standard Gibbs free energy change. It means that when a mole of chlorophyll molecules absorbs a mole of photons the excited electrons acquire an amount of energy equal to 176 kJ mol^{-1}. As they fall back to their ground state they give up this energy and some of it is captured and used to pump protons across the membrane or to synthesize NADPH.

C. The Special Pair and Antenna Chlorophylls

A typical photosystem contains dozens of chlorophyll molecules but only two special chlorophyll molecules actually give up electrons to begin the electron transfer chain. These two chlorophyll molecules are called the **special pair**. In most cases the special pair is identified simply as pigments (P) that absorb light at a specific wavelength. Thus, P680 is the special pair of chlorophyll molecules that absorbs light at 680 nm (red). Its three states are P680, P680*, and P680⊕. P680 is the ground state. P680* is the state following absorption of a photon of light when the chlorophyll macromolecules have an excited electron. P680⊕ is the electron-deficient (oxidized) state following transfer of an electron to another molecule. P680⊕ is reduced to P680 by transfer of an electron from an electron donor.

In addition to the special pair there are other specialized chlorophyll molecules that function as part of the electron transfer chain. They accept electrons from the special pair and transfer them to the next molecule on the pathway. Not all chlorophylls are directly involved in electron transfer. The remaining chlorophylls act as antenna molecules by capturing light energy and transferring it to the special pair. These antenna chlorophylls are much more numerous than the molecules in the electron transfer chain. The mode of excitation energy transfer between antenna chlorophylls is called **resonance energy transfer**. It does not involve the movement of electrons. You can think of excitation energy transfer as a transfer of vibrational energy between adjacent chlorophyll molecules in the densely packed antenna complex.

Figure 15.3 illustrates the transfer of excitation energy from antenna chlorophylls to the special pair in one of the photosystems. The figure shows only a few of the many antenna molecules surrounding the special pair. All chlorophyll molecules are held in

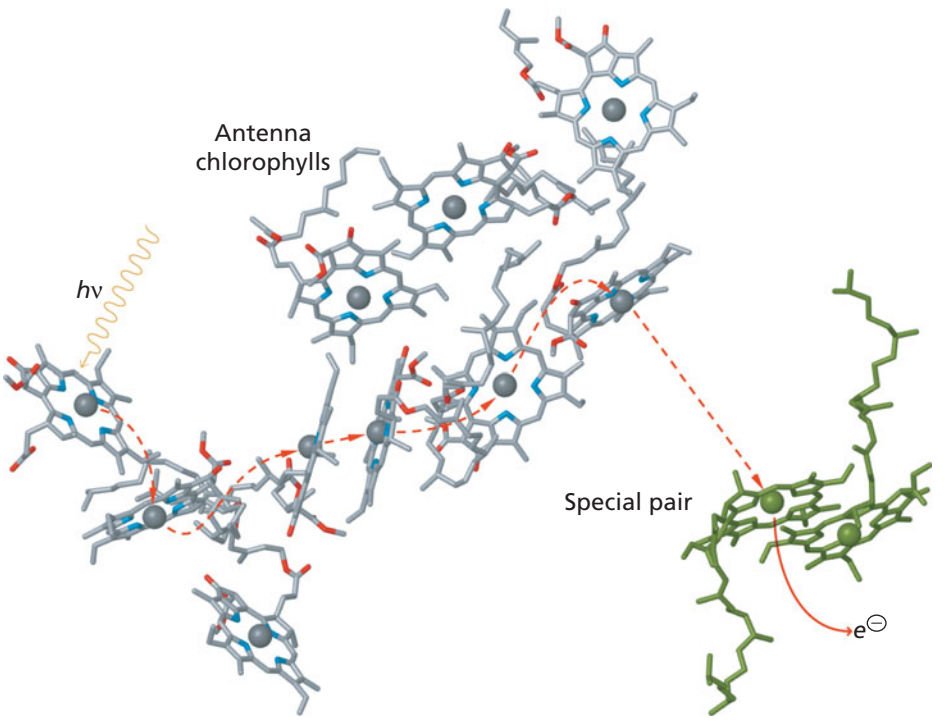


Figure 15.3 ►
Transfer of light energy from antenna chlorophyll pigments to the special pair of chlorophyll molecules. Light can be captured by the antenna pigments (gray) and excitation energy is transferred between antenna chlorophylls until it reaches the special pair of chlorophyll molecules in the electron transfer pathway (green). The path of excitation energy transfer is shown in red. The special pair gives up an electron to the electron transfer pathway. The chlorophyll molecules are held in fixed positions because they are tightly bound to membrane proteins (not shown).

BOX 15.1 MENDEL'S SEED COLOR MUTANT

One of Gregor Mendel's original mutants affected the color of the peas in a pod. The normal color of mature seeds is yellow (I) and the recessive mutant confers a green color to the seeds (i). The mutation affects the "stay-green" (*sgr*) gene that encodes a chloroplast protein responsible for the degradation of chlorophyll as the seeds mature. When the protein is defective, chlorophyll is not broken down in the chloroplasts and the seeds stay green.

In normal wild-type plants (II) the seed are yellow and in the heterozygotes (Ii) the deficiency in the amount of chlorophyll degradation protein is not sufficient to affect chlorophyll breakdown. The seeds of the heterozygotes are also yellow. In homozygous mutant plants (ii) chlorophyll is not degraded and the seeds are green. Mendel determined that the wild-type trait (I) was dominant and the mutant trait (i) was recessive. Crosses between heterozygotes (Ii x Ii) gave the famous 3:1 ratio of yellow seeds to green seeds.

Some strains of food plants are homozygous for mutations in the genes that break down chlorophyll. These "cosmetic stay-greens," such as the one used by Mendel, produce seeds and fruit that are more attractive to consumers.

All the peas that we buy in supermarkets and farmer's markets have been genetically modified (by breeding) to be homozygous for the deficient *sgr* allele. That's why we never see the "normal" yellow peas.

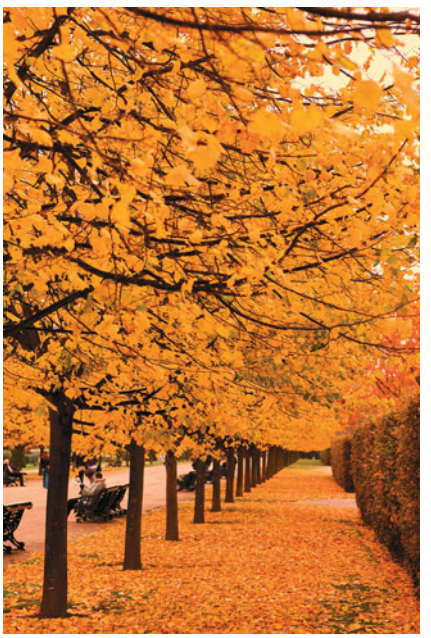


► Normal mature peas turn yellow in color as they mature (bottom) but a mutation causes the seeds to retain their green color (top). The seed coat has been removed from the lower pair of each group in order to make the color difference more obvious.

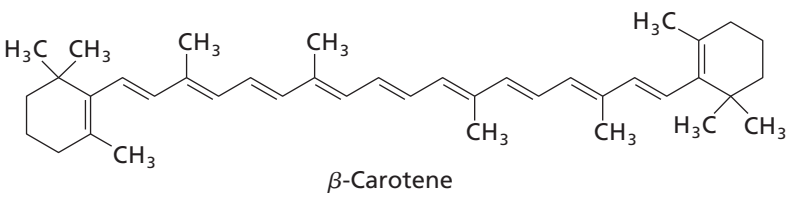
fixed positions through interactions with the side chains of amino acids in the polypeptides of the photosystem. Excitation energy is efficiently transferred from any molecule that absorbs a photon because these molecules are so close to each other.

D. Accessory Pigments

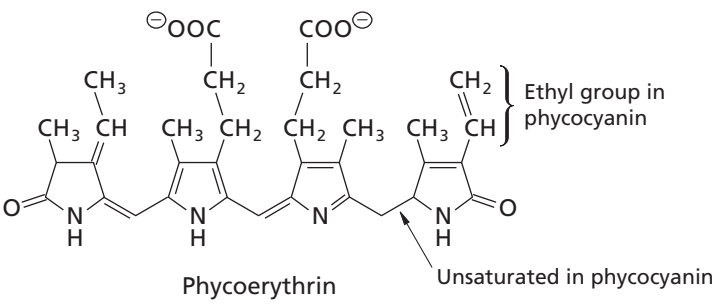
Photosynthetic membranes contain several **accessory pigments** in addition to chlorophyll. The carotenoids include β-carotene (Figure 15.4) and related pigments such as xanthophylls. Xanthophylls have extra hydroxyl groups on the two rings. Note that the carotenoids, like chlorophyll, contain a series of conjugated double bonds allowing them to absorb light. Their absorption maxima lie in the blue region of the spectrum, which is why carotenoids appear red, yellow, or brown (Figure 15.2). The autumn colors of deciduous trees are due, in part, to carotenoids, as is the brown color of sea kelp (brown algae).



▲ The autumn colors of the leaves are due, in part, to the presence of accessory carotenoid pigments that become visible when chlorophyll molecules are degraded as the leaves die.



β-Carotene



Phycoerythrin

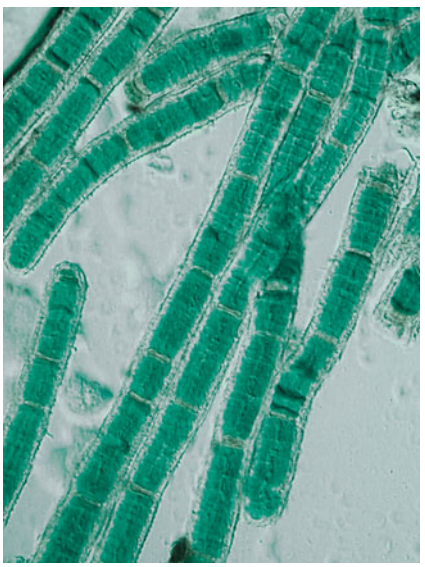
◄ **Figure 15.4**
Structures of some accessory pigments. β-Carotene is a carotenoid, and phycoerythrin and phycocyanin are phycobilins. Phycobilins are covalently attached to proteins whereas carotenoids are bound noncovalently.



▲ **Red tide.** This red tide off the coast of Fujian, China, is due to the presence of red algae.

Carotenoids are closely associated with chlorophyll molecules in antenna complexes. They absorb light and transfer excitation energy to adjacent chlorophylls. In addition to serving as light-gathering pigments carotenoids also play a protective role in photosynthesis. They take up any electrons that are accidently released from antenna chlorophylls and return them to the oxidized chlorophyll molecule. This quenching process prevents the formation of reactive oxygen species such as the superoxide radical ($\cdot O_2^{\ominus}$). If allowed to form, these reactive oxygen species can be highly toxic to cells as described in Section 14.14.

Phycobilins, such as red phycoerythrin and blue phycocyanin (Figure 15.4), are found in some algae and cyanobacteria. They resemble a linear version of chlorophyll without the central magnesium ion. Like chlorophylls and carotenoids, these molecules contain a series of conjugated double bonds that allow them to absorb light. Like carotenoids, the absorption maxima of phycobilins complement those of chlorophylls and thus broaden the range of light energy that can be absorbed. In most cases, the phycobilins are found in special antenna complexes called phycobilisomes. Unlike other pigment molecules, the phycobilins are covalently attached to their supporting polypeptides. The bluish color of blue-green cyanobacteria and the red color of red algae are due to the presence of numerous phycobilisomes associated with their photosystems.



▲ **Scytonema—a blue-green cyanobacterium.**

15.2 Bacterial Photosystems

We begin our discussion by describing simple bacterial systems. These simple systems evolved into more complicated structures in the cyanobacteria. The cyanobacterial version of photosynthesis was then adopted by algae and plants when a primitive cyanobacterium gave rise to chloroplasts.

Photosynthetic bacteria contain typical light-gathering photosystems. There are two basic types of photosystems that appear to have diverged from a common ancestor more than two billion years ago. Both types of photosystem contain a large number of antenna pigments surrounding a small reaction center located in the middle of the structure. The reaction center consists of a few chlorophyll molecules that include the special pair and others forming a short electron transfer chain.

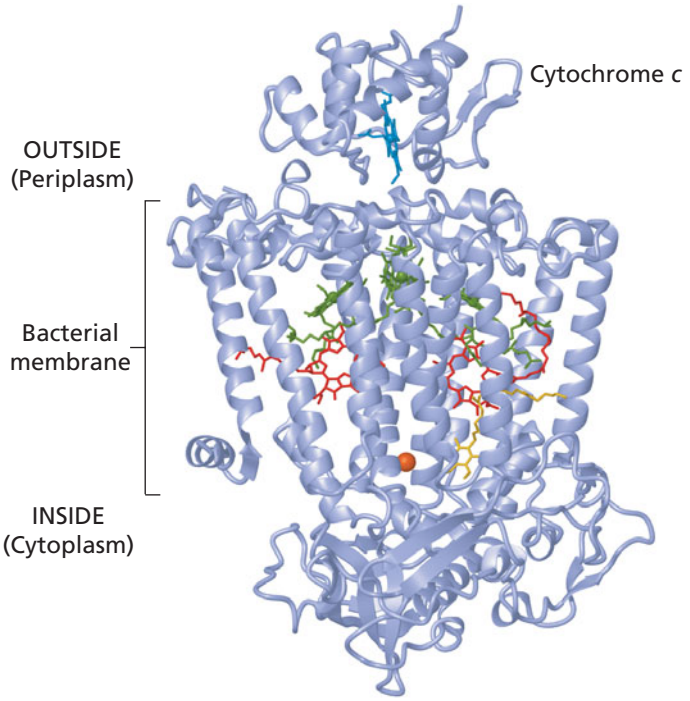
Photosystem I (PSI) contains a **type I reaction center**. **Photosystem II (PSII)** contains a **type II reaction center**. Heliobacteria and green sulfur bacteria rely on photosystems with a type I reaction center whereas purple bacteria and green filamentous bacteria use photosystems with a type II reaction center. Cyanobacteria, the most abundant class of photosynthetic bacteria, utilize both photosystem I and photosystem II coupled in series. This coupled system resembles the one found in algae and plants.

A. Photosystem II

We begin by describing photosynthesis in purple bacteria and green filamentous bacteria. Most of these species of bacteria are strict anaerobes—they cannot survive in the presence of oxygen. Thus, they do not produce oxygen as a by-product of photosynthesis or consume it in respiratory electron transport. Purple bacteria and green filamentous bacteria have photosystems with a type II reaction center. These membrane complexes are often referred to as the bacterial reaction center (BRC) but this is misleading since bacteria also contain the other type of reaction center. We will refer to it here as photosystem II since it is evolutionarily related to photosystem II in cyanobacteria and eukaryotes.

The structure of the purple bacteria photosystem is shown in Figure 15.5. The pigment molecules of the internal type II reaction center form an electron transfer chain with two branches. The special pair of bacteriochlorophylls (P870) are positioned near the periplasmic (outside) surface of the membrane. Each branch contains a molecule of bacteriochlorophyll *a* and a bacteriopheophytin molecule (Figure 15.6). The right-hand branch terminates in a tightly bound quinone molecule while the equivalent position in the left-hand branch is occupied by a loosely bound quinone that can dissociate and diffuse within the lipid bilayer. Note in Figure 15.5 that the bound quinone is buried within the α helix barrel spanning the membrane while the equivalent site on the other side of the complex is open to the lipid bilayer.

The structure of the photosystem of the purple bacterium, *Rhodopseudomas viridis*, is shown in Figure 4.25f.



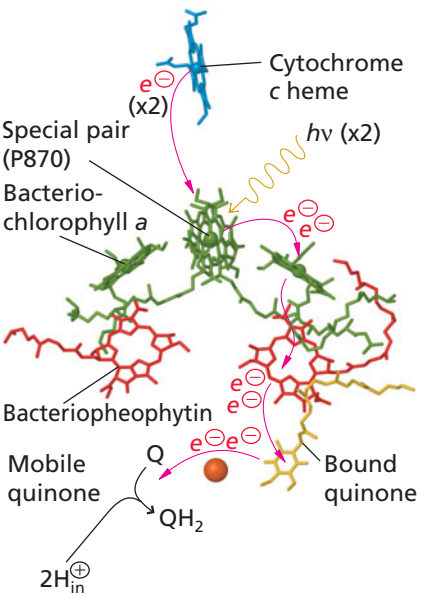
◂ **Figure 15.5**
Photosystem II in the purple bacterium *Rhodobacter spaeroides*. The core of the structure consists of two homologous membrane-spanning polypeptide subunits (L and M). Each subunit has five transmembrane α helices. The electron transfer molecules of the reaction center are sandwiched between the core polypeptides. Cytochrome *c* binds to PSII on the periplasmic side of the membrane (top). An additional subunit covers the core subunits on the cytoplasmic surface (bottom). [PDB 1L9B]

Electron transfer begins with the release of an excited electron from P870 following absorption of a photon of light or the transfer of excitation energy from antenna pigments. (Antenna pigment molecules are not shown in Figure 15.6.) Electrons are then transferred exclusively down the right-hand branch of the reaction center complex resulting in the reduction of the bound quinone molecule. From there, electrons are passed to the mobile quinone on the opposite side of the complex. This transfer is mediated by a single bound iron atom on the central axis near the cytoplasmic side of the membrane. The mobile quinone (Q) is reduced to QH_2 in a two-step process via the sequential transfer of two electrons and the uptake of two $H^{\oplus}$ from the cytoplasm. Two photons of light are absorbed for each molecule of QH_2 produced. Modern type II reaction centers probably evolved from a more primitive system in which electrons were transferred down both branches to produce QH_2 at both of the Q sites.

QH_2 diffuses within the lipid bilayer to the cytochrome bc_1 complex (complex III) of the bacterial respiratory electron transport system. This is the same complex that we described in the previous chapter (Section 14.7). The cytochrome bc_1 complex catalyzes the oxidation of QH_2 and the reduction of cytochrome *c*—the enzyme is ubiquinol: cytochrome *c* oxidoreductase. This reaction is coupled to the transfer of $H^{\oplus}$ from the cytoplasm to the periplasmic space via the Q cycle. The resulting proton gradient drives the synthesis of ATP by ATP synthase (Figure 15.7).

The $P870^{\oplus}$ special pair of chlorophyll molecules is reduced by the cytochrome *c* (Fe^{2+}) molecules produced by the cytochrome bc_1 complex. Cytochrome *c* diffuses within the periplasmic space enclosed by the two membranes surrounding the bacterial cell. The net effect is that electrons are shuffled from PSII to the cytochrome bc_1 complex and back again. Note that the structure shown in Figure 15.5 includes a bound cytochrome *c* molecule with its heme group positioned near the P870 special pair in order to facilitate electron transfer.

The movement of electrons between complexes is mediated by the mobile cofactors QH_2 and cytochrome *c* just as we saw in respiratory electron transport. The main difference between photosynthesis in purple bacteria and respiratory electron transport is that photosynthesis is a cyclic process. There is no net gain or loss of electrons to other reactions and consequently no outside source of electrons is needed. Cyclic electron flow is a characteristic of many, but not all, photosynthesis reactions. The result of coupling PSII and the cytochrome bc_1 complex is that absorption of light creates a proton



▲ **Figure 15.6**
The type II reaction center contains the electron transfer chain. The special pair (P870) is located near the periplasmic surface close to the heme group of cytochrome *c*. When light is absorbed, electrons are transferred one at a time from P870 to BChl *a* to BPh to a bound quinone and from there to a quinone located at a loosely bound site next to a central iron atom (orange). Electrons are restored to P870 from cytochrome *c*.

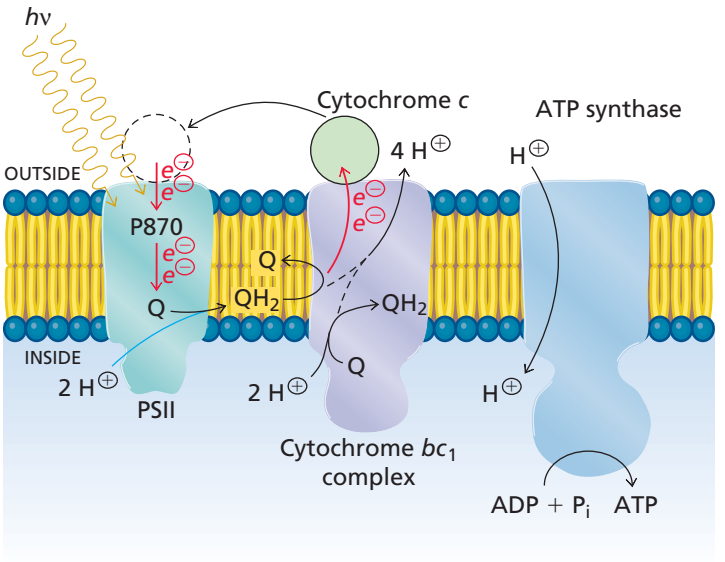


Figure 15.7 ▸
Photosynthesis in purple bacteria. Light is absorbed by the pigments of the PSII complex resulting in the transfer of electrons from P870 to QH_2 via the reaction center electron transfer chain. QH_2 diffuses to the cytochrome bc_1 complex where the electrons are transferred to cytochrome *c*. This reaction is coupled to the transfer of protons across the membrane. The proton gradient drives the synthesis of ATP. Reduced cytochrome *c* diffuses within the periplasmic space to PSII where it reduces $P870^+$. The Q-cycle reactions are shown in more detail in Figure 14.11.

KEY CONCEPT

Bacteria with photosystem II use sunlight to produce a proton gradient that drives ATP synthesis.

KEY CONCEPT

Photosynthesis in purple bacteria is a cyclic process. It does not require an external source of electrons such as H_2O or H_eS.

gradient for ATP synthesis. The reactions are listed in Table 15.1. (The cytochrome bc_1 reactions are the same ones shown in Table 14.3.) Four protons are transferred across the membrane for every two photons of light that are absorbed. The ATP molecules produced as a result of this cycle are used by bacteria to synthesize proteins, nucleic acids, carbohydrates, and lipids. Thus, captured light energy is ultimately used in biosynthesis reactions.

We can calculate the energy of two "moles" of light at 870 nm using Equation 15.1. It works out to 274 kJ mol^{-1}. This light energy is used to pump four protons across the membrane. Pumping requires approximately 4×19.4 kJ $mol^{-1} = 77.6$ kJ mol^{-1} using our estimate from the previous chapter (Section 14.3). The result suggests that the production of chemical energy from light energy is not very efficient in purple bacteria (77.6/274 = 28%).

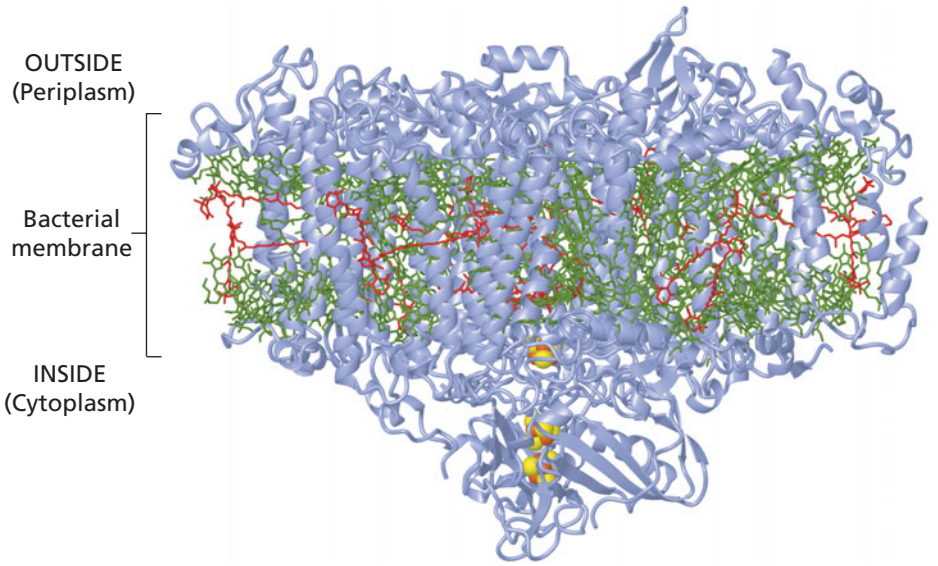
The basic principle of photosynthesis is the conversion of light energy (photons) to chemical energy (e.g. ATP). The pathway clearly evolved, in part, from the electron transport system we described in the previous chapter. Photosynthesis evolved several hundred million years after the main energy-producing pathway that uses complex III and ATP synthase. It's important to note that the ATP produced in bacterial photosynthesis is not restricted to the synthesis of carbohydrate and oxygen is not produced as part of the process.

Table 15.1 Photosystem II reactions

PSII:	$2\ P870 + 2\ photons \longrightarrow 2\ P870^{\oplus} + 2\ e^{\ominus}$
	$Q + 2\ e^{\ominus} + 2\ H^{\oplus}_{in} \longrightarrow QH_2$
Cyt bc_1:	$2\ QH_2 + 2\ cyt\ c\ (Fe^{3+}) \longrightarrow 2\ Q + 2\ cyt\ c\ (Fe^{2+}) + 4\ H^{\oplus}_{out} + 2\ e^{\ominus}$
	$Q + 2\ e^{\ominus} + 2\ H^{\oplus}_{in} \longrightarrow QH_2$
PSII:	$2\ cyt\ c\ (Fe^{2+}) + 2\ P870^{\oplus} \longrightarrow 2\ cyt\ c\ (Fe^{3+}) + 2\ P870$
Sum:	$2\ photons + 4\ H^{\oplus}_{in} \longrightarrow 4\ H^{\oplus}_{out}$

B. Photosystem I

The structure of a typical photosystem I (PSI) complex is shown in Figure 15.8. The central part of the complex is formed by two homologous polypeptides with multiple membrane-spanning α helices. Each subunit of this dimer has two domains—an interior domain that binds the electron transfer chain pigments of the type I reaction center and a peripheral domain that binds antenna pigments. The reaction center protein domains in PSI subunits are related by structure and amino acid sequence to the core polypeptides in PSII. This is strong evidence for a common ancestor of type I and type II reaction centers.



◂ **Figure 15.8**
Structure of photosystem I (PSI). This version of PSI is from the cyanobacterium *Thermosynechococcus elongatus* (*Synechococcus elongatus*). The complex contains 96 chlorophylls (green), 22 carotenoids (red), and three iron–sulfur clusters (orange). There are 14 polypeptide subunits, most of which have membrane-spanning α helices. [PBD 1JBO]

The most obvious difference between PSI and PSII is the presence of a more complex antenna structure in PSI than in PSII. The PSI antenna complex is packed with chlorophyll and carotenoid pigment molecules. The example shown in Figure 15.8 is from cyanobacteria whose PSI complexes contain 96 chlorophylls and 22 carotenoids. Many of the light-gathering pigment molecules are tightly bound to additional membrane-spanning polypeptide subunits that surround the core subunits. The contrast between the structures shown in Figure 15.5 and Figure 15.8 is a bit misleading since there are simpler forms of PSI in some bacteria and more complex versions of PSII in other species (see below). Nevertheless, as a general rule, PSI is larger and more complicated than PSII.

The organization of the electron transfer chain molecules in PSI reveals striking parallels to that of PSII (Figure 15.9). In both cases, the reaction center contains two short branches of pigment molecules that terminate at bound quinones. The PSI pigment molecules are both chlorophylls and not one chlorophyll and one pheophytin as in PSII. The bound quinones in PSI are usually phylloquinones whereas in PSII they are related to ubiquinone (or menaquinone in bacteria). The phylloquinones in type I reaction centers are tightly bound to the complex and form part of the electron transfer chain. (Recall that one of the quinones in type II reaction centers is a mobile terminal electron acceptor.)

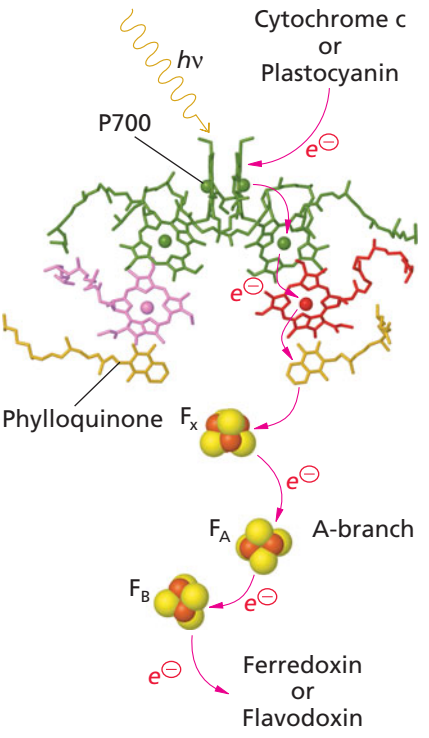
Electron transfer begins with a special pair of chlorophyll molecules located near the periplasmic surface of the membrane. This special pair is known as P700 since it absorbs light at a wavelength of 700 nm. The two chlorophyll molecules are not identical—the molecule closest to the A-branch is an epimer of chlorophyll *a* (bacteriochlorophyll *a* in bacteria). P700 is excited by absorbing a photon of light or by excitation energy transfer from antenna molecules. The excited electron is then transferred down one of the branches of the electron transfer chain to one of the bound phylloquinones. Electron transfer from P700 to phylloquinone takes about 20 picoseconds (10^{-12} s). This is extremely rapid compared to other electron transfer systems. In type II reaction centers, for example, the transfer from P680 to the bound quinone takes two or three times longer.

Electrons are subsequently transferred from bound phylloquinone to the three Fe–S clusters, F_X, F_A, and F_B. The terminal electron acceptor in PSI is ferredoxin (or flavodoxin) (Figure 7.36). Ferredoxin contains two [4Fe-4S] iron–sulfur clusters and reduction involves a $Fe^{3+} \rightarrow Fe^{2+}$ reduction with a standard reduction potential of −0.43 V (Table 10.5).

Reduced ferredoxin (Fd_{red}) becomes the substrate for an oxidation–reduction reaction catalyzed by an enzyme called ferredoxin:$NADP^{\oplus}$ oxidoreductase, more commonly known as ferredoxin:$NADP^{\oplus}$ reductase or FNR. The enzyme is a flavoprotein (containing FAD) and the reaction proceeds in three steps involving a typical semiquinone intermediate (Section 7.5). The product of the reaction is reducing equivalents in the form of NADPH. The coupled reactions involving PSI are shown in Table 15.2.

Note that the standard reduction potential of ferredoxin is considerably lower than that of $NADP^{\oplus}$, allowing for transfer of electrons from ferredoxin to $NDAP^{\oplus}$. The terminal electron acceptor is Q in photosystem II and its standard reduction potential is

Phylloquinone is also known as vitamin K (Section 7.14D, Figure 7.29).



▲ **Figure 15.9**
PSI electron transfer chain (type I reaction center). Electron transfer begins with the special pair of chlorophyll molecules (P700) and proceeds down one of the branches to phylloquinone. From there, electrons are transferred to the Fe–S clusters and eventually to ferredoxin. $P700^{\oplus}$ is reduced by cytochrome *c* or plastocyanin.

Table 15.2 The photosystem I reactions

PSI:	$2\ P700 + 2\ photons \longrightarrow 2\ P700^{\oplus} + 2\ e^{\ominus}$
	$2\ Fd_{ox} + 2\ e^{\ominus} + \longrightarrow 2\ Fd_{red}$
FNR:	$Fd_{red} + H^{\oplus} + FAD \rightleftharpoons Fd_{ox} + FADH\cdot$
	$Fd_{red} + H^{\oplus} + FADH\cdot \rightleftharpoons Fd_{ox} + FADH_2$
	$FADH_2 + NADP^{\oplus} \rightleftharpoons FAD + NADPH + H^{\oplus}$
Sum:	$2\ P700 + 2\ photons + NADP^{\oplus} + H^{\oplus} \longrightarrow 2\ P700^{\oplus} + NADPH$

KEY CONCEPT

Bacteria with photosystem I use sunlight to produce NADPH.

Ferredoxin ($Fe^{3\oplus}$) + $e^{\ominus}$ → $Fe^{2\oplus}$
$\Delta E = -0.43$ V
$NADP^{\oplus} + H^{\oplus} + 2\ e^{\ominus} \longrightarrow NADPH$
$\Delta E = -0.32$ V
Ubiquinone (Q) + $2\ H^{\oplus}$ + $2\ e^{\ominus} \longrightarrow QH_2$
$\Delta E = +0.04$ V

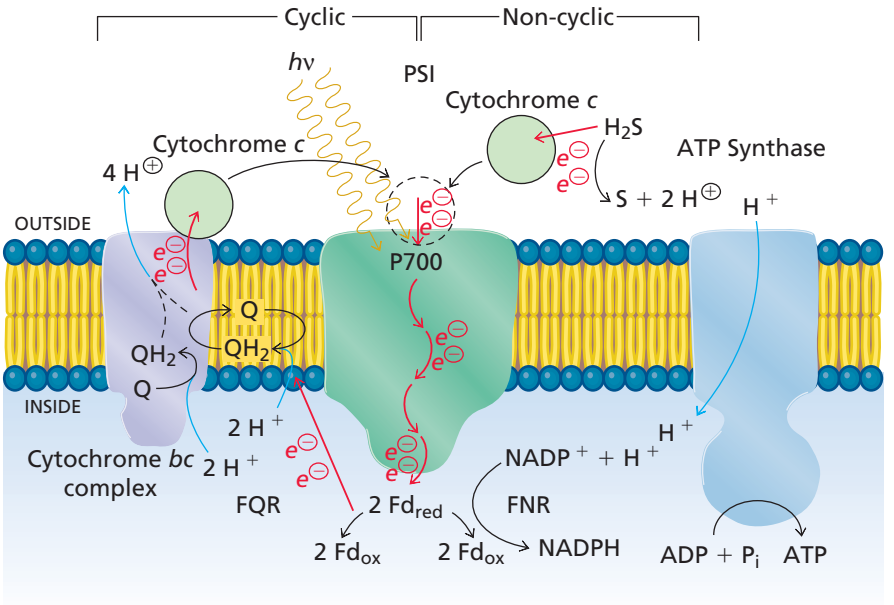
too high to allow transfer of electrons to $NADP^{\oplus}$. This means that energy capture from sunlight is more efficient in PSII than in PSI.

The reactions in PSI do not create a cyclic pathway. The oxidized special pair in type I reaction centers ($P700^{\oplus}$) must be reduced by electrons from an outside source since the excited chlorophyll electrons were eventually transferred to NADPH. Some bacteria contain versions of PSI that bind cytochrome *c* on the outside surface of the membrane next to the special pair. In these bacteria $P700^{\oplus}$ is reduced by reduced cytochrome *c* in a manner similar to the reduction of the special pair in PSII. The source of electrons for reduced cytochrome *c* depends on the species. In green sulfur bacteria it is various reduced sulfur compounds such as H_2S and $S_2O_3^{2-}$. The oxidation of these sulfur compounds is coupled to the transfer of electrons to cytochrome *c* by special enzymes that are found in these species (Figure 15.10). Green sulfur bacteria are photoautotrophs (Section 10.3) that grow in the absence of oxygen.

Noncyclic electron transfer is a characteristic feature of PSI but there can also be a cyclic process of electron transfer. Some electrons from PSI are occasionally passed from ferredoxin to a quinone—probably by ferredoxin:quinone oxidoreductase (ferredoxin: quinone reductase, FQR). Quinol (QH_2) interacts with the cytochrome bc_1 complex



▲ **Green sulfur bacteria.** Agar plate with streaks of *Chlorobium tepidum*.



▲ **Figure 15.10**
Photosynthesis in green sulfur bacteria. Photoactivation of P700 leads to production of reduced ferredoxin on the cytoplasmic side of the membrane. Ferredoxin becomes the electron donor in a reaction catalyzed by ferredoxin:$NADP^{\oplus}$ reductase (FNR) resulting in production of NADPH in the cytoplasm. Ferredoxin can also reduce Q to QH_2 in a reaction catalyzed by ferredoxin:quinone reductase (FQR). QH_2 is oxidized by the cytochrome bc_1 complex, resulting in the transfer of electrons to reduced cytochrome *c* and the transfer of protons across the membrane. $P700^{\oplus}$ is normally reduced by cytochrome *c* on the periplasmic side of the membrane. In the noncyclic process, reduced cytochrome *c* is made in reactions that are coupled to the oxidation of sulfur compounds such as H_2S. The transfer of electrons is shown by red arrows.

transferring electrons via cytochrome bc_1 to cytochrome *c* and cytochrome *c* reduces $P700^{\oplus}$ (Figure 15.10). This cyclic process is very similar to the coupled reactions involving PSII. It allows for light-mediated synthesis of ATP because the passage of electrons through cytochrome bc_1 is associated with the translocation of protons across the membrane via the Q cycle. In most cases, the noncyclic process predominates and NADPH is produced; however, if NADPH cannot be efficiently used in biosynthesis reactions, electrons will be transferred through cytochrome bc_1 to produce ATP.

Reduced ferredoxin can be used directly in other pathways, notably in nitrogen fixation (Section 17.1)

C. Coupled Photosystems and Cytochrome *bf*

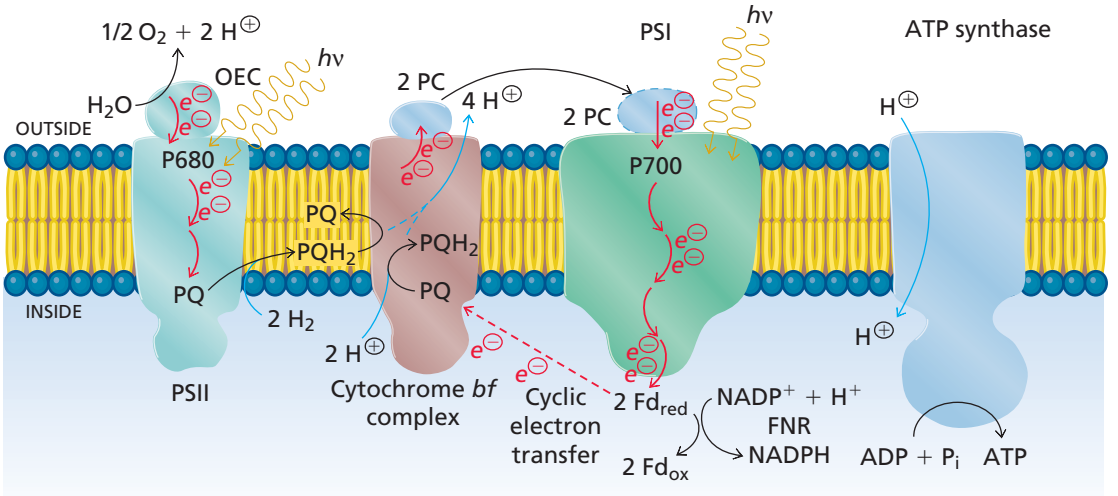
Cyanobacterial membranes contain both PSI and PSII. The two photosystems are coupled in series to produce both NADPH and ATP in response to light. The photosynthetic reactions in cyanobacteria are illustrated in Figure 15.11. Light is absorbed by PSII leading to excitation of P680 and transfer of an electron to a mobile quinone called plastoquinone (PQ, Figure 7.33). Electrons are then transferred to a cytochrome *bf* complex similar to the cytochrome bc_1 complex in respiratory electron transport. Electron transport within the cytochrome *bf* complex is coupled to the movement of $H^{\oplus}$ across the membrane by a photosynthetic Q cycle. The coupling of PSII and a cytochrome *bf* complex is similar in principle to photosynthesis reactions in purple bacteria with one major difference—in purple bacteria electrons are returned to PSII by the terminal electron acceptor of the cytochrome bc_1 complex (cytochrome *c*) whereas in cyanobacteria electrons are passed on to PSI. The terminal electron acceptor of the unique cytochrome *bf* complex is either cytochrome *c* or a blue copper-containing protein called plastocyanin (PC). Reduced cytochrome *c* and reduced plastocyanin are mobile carriers that bind to the outside (periplasmic) surface of PSI and reduce $P700^{\oplus}$. (Most cyanobacteria and algae use cytochrome *c* while some cyanobacteria and all plants use plastocyanin, or a different cytochrome called cytochrome c_6, as the terminal electron acceptor of the cytochrome *bf* complex.)

The structure of the photosynthetic cytochrome *bf* complex has been solved by X-ray crystallography (Figure 15.12). It contains a cytochrome *b* with two cytochrome reaction centers whose role in the Q cycle is similar to that of cytochrome *b* in the cytochrome bc_1 complex (complex III) of respiratory electron transport. A Rieske iron–sulfur protein (ISP) transports electrons from one of the cytochrome *b* sites to cytochrome *f* and reduced cytochrome *f* passes electrons to plastocyanin. Cytochrome *f* (*f* stands for *feuille*, the French word for leaf) is a distinct protein unrelated to cytochrome c_1 of the respiratory cytochrome bc_1 complex but cytochrome *b* and ISP are homologues of the proteins found in complex III.

The cytochrome *bf* complex evolved from the original cytochrome bc_1 complex that was present in ancient cyanobacteria. The most important adaptation was the replacement of cytochrome c_1 of the bacterial *bc* complex with cytochrome *f* in the cyanobacterial complex. This change allowed for the transfer of electrons to the copper-containing plastocyanin via cytochrome *f*. (Recall that mobile cytochrome *c*, not plastocyanin, is the normal electron acceptor of the cytochrome bc_1 complex.)

KEY CONCEPT

Organisms with coupled photosystem I and photosystem II use sunlight to produce both NADPH *and* a proton gradient that drives ATP synthesis.



▼ **Figure 15.11**
Photosynthesis in cyanobacteria. Light (wavy arrows) is captured and used to drive the transport of electrons (obtained from water) from PSII through the cytochrome *bf* complex to PSI and ferredoxin. This process can generate NADPH and a proton concentration gradient that is used to drive phosphorylation of ADP. For each water molecule oxidized to 1/2 O_2 by the oxygen evolving complex (OEC), one molecule of $NADP^{\oplus}$ is reduced to NADPH. For simplicity, PSI, PSII, and cytochrome *bf* are shown close together in the plasma membrane but in most species they are located within internal membrane structures. Plastoquinone (PQ) is the mobile carrier between PSII and the cytochrome *bf* complex. In this example, plastocyanin (PC) is the mobile carrier between the cytochrome *bf* complex and PSI.

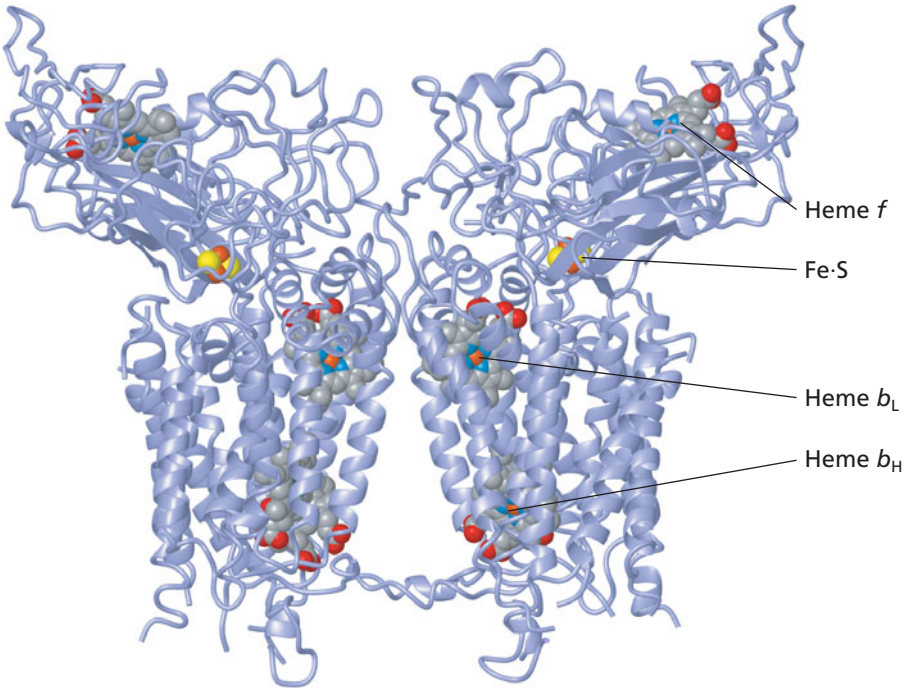


Figure 15.12 ▸
Cytochrome *bf* complex from the cyanobacterium *Mastigocladus laminosus*. The complex contains two functional enzymes as in complex III (compare Figure 14.10). The primary electron transfer components are: heme b_L and heme b_H (the sites of Q-cycle oxidation reactions), the iron–sulfur cluster (Fe–S) in ISP, and heme *f*. Each unit also contains a chlorophyll *a*, a β-carotene, and an unusual heme *x* whose function is unknown (not shown). [PDB 1UM3]

Plastocyanin binds specifically to PSI in cyanobacteria and transfers electrons to $P700^{\oplus}$. This allows for a unidirectional flow of electrons from PSII → PQH_2 → cytochrome *bf* → PC → PSI → NADPH.

Cyanobacteria do not contain cytochrome bc_1. Thus, cytochrome *bf* also plays a role in respiratory electron transport because it replaces the normal complex III. Reduced plastocyanin is the electron donor to the terminal oxidase (complex IV) possibly via an intermediate cytochrome *c* - like carrier. Plastoquinone is the mobile quinone electron carrier in both photosynthesis and respiratory electron transport.

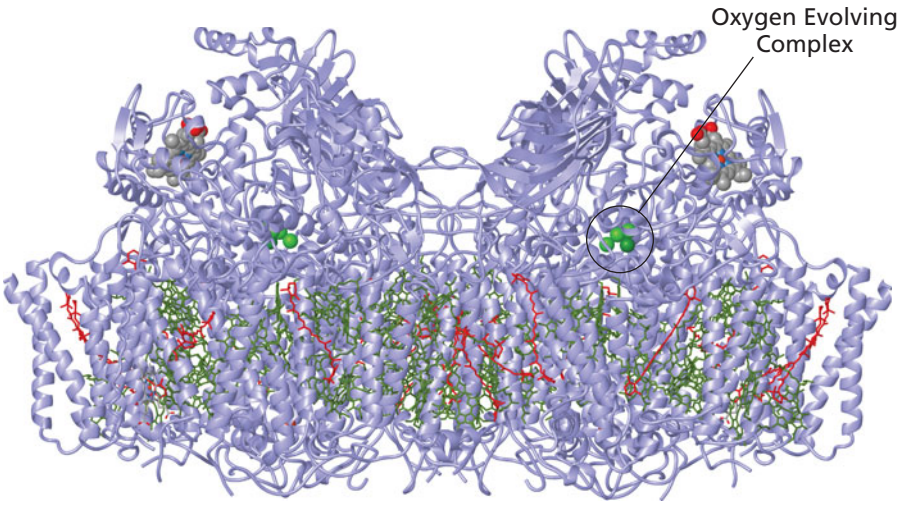
Photoactivation of PSI results in synthesis of NADPH in a manner similar to that in green sulfur bacteria. As in green sulfur bacteria, some electrons are recycled but in this case it is through the cytochrome *bf* complex. Note that PSII, cytochrome *bf*, and PSI are coupled in series and the transfer of electrons to NADPH results in a deficiency of electrons at $P680^{\oplus}$ in PSII. The reduction of $P680^{\oplus}$ in cyanobacteria is accomplished by extracting electrons from water with the production of oxygen as a by-product. The enzyme that splits water is called the oxygen evolving complex (OEC) and it is tightly bound to PSII on the outer surface of the membrane. The evolution of an oxygen evolving complex in primitive cyanobacteria was one of the most important biochemical events in the history of life.

KEY CONCEPT

The splitting of water to form molecular oxygen arose in order to supply electrons to photosystem II.

The oxygen evolving complex (OEC) contains a cluster of Mn^{2+} ions, a Ca^{2+} ion, and a Cl^{-} ion. It catalyzes a complex reaction in which four electrons are extracted, one at a time, from two molecules of water. The reaction takes place on the outside of the PSII complex near the special pair of chlorophyll molecules (P680). The electrons from the splitting of water are transferred to $P680^{\oplus}$ (Figure 15.13). The exact mechanism of the water splitting reaction is being investigated in a number of laboratories. It is similar, in principle, to the reverse reaction catalyzed by complex IV of the respiratory electron transport chain (Section 14.8). Note that the oxygen evolving complex is located on the exterior surface of the membrane and the release of protons from water contributes to the formation of the proton gradient across the membrane.

As mentioned earlier, the similarities between PSI and PSII indicate that they evolved from a common ancestor. Over time, these two photosystems diverged in those species of photosynthetic bacteria that contain only one of the two types (e.g., purple bacteria, green sulfur bacteria). At some point, about 2.5 billion years ago, a primitive ancestor of cyanobacteria acquired both types of photosystem—probably by taking up a large part of the genome from an unrelated bacterial species. At first the two types of



◂ **Figure 15.13**
PSII and the oxygen-evolving center. The PSII complex in the cyanobacterium *Thermosynechococcus elongatus* is much larger than the PSII complex in purple bacteria (Figure 15.5) but the core structures are very similar. The cyanobacteria complex contains many antenna chorolophylls and carotenoids and it is a dimer. The oxygen evolving complex (OEC) contains a Mn_3CaO_4 cluster (circled) where the splitting of water occurs. This metal ion cluster is positioned over the type II reaction center. [PDB 3BZ1]

photosystem must have worked in parallel but they began to function in series with the evolution of a photosynthetic cytochrome *bf* complex (from cytochrome bc_1) and an oxygen evolving complex. Later on, a species of cyanobacteria entered into a symbiotic relationship with a primitive eukaryotic cell and this led to the modern chloroplasts found in algae and plants.

The coupled photosystems are able to capture light energy and use it to produce both ATP (from the proton gradient) and reducing equivalents in the form of NADPH. Neither photosystem by itself can accomplish these two goals with the same efficiency.

The net result of this simplified linear pathway is the production of one molecule of NADPH and the transfer of four protons across the membrane for each *pair* of electrons excited by the absorption of light energy in each photosystem. The two separate excitation steps in PSI and PSII require a total of four photons of light energy. The splitting of water by the OEC contributes to the proton gradient and produces molecular oxygen. The individual reactions are summarized in Table 15.3.

D. Reduction Potentials and Gibbs Free Energy in Photosynthesis

The path of electron flow during photosynthesis can be depicted in a zigzag figure called the **Z-scheme** (Figure 15.14). The Z-scheme plots the reduction potentials of the photosynthetic electron transfer components in PSI, PSII, and cytochrome *bf*. It shows that the absorption of light energy converts P680 and P700—pigment molecules that are poor reducing agents—to excited molecules (P680* and P700*) that are good

Table 15.3 The photosynthesis reactions in species with both photosystems

PSII:	$2\ P680 + 2\ photons \longrightarrow 2\ P680^{\oplus} + 2\ e^{\ominus}$
	$PQ + 2\ e^{\ominus} + 2\ H^{\oplus}_{in} \longrightarrow PQH_2$
OEC:	$H_2O \longrightarrow \frac{1}{2}O_2 + 2\ H^{\oplus}_{out} + 2\ e^{\ominus}$
	$2\ P680^{\oplus} + 2\ e^{\ominus} \longrightarrow 2\ P680$
Cyt *bf*:	$2\ PQH_2 + 2\ \text{plastocyanin}\ (Cu^{\oplus}) \longrightarrow 2\ PQ + 2\ \text{plastocyanin}\ (Cu^{\oplus}) + 4\ H^{\oplus}_{out} + 2\ e^{\ominus}$
	$PQ + 2\ H^{\oplus}_{in} + 2\ e^{\ominus} \longrightarrow PQH_2$
PSI:	$2\ P700 + 2\ photons \longrightarrow 2\ P700^{\oplus} + 2\ e^{\ominus}$
	$2\ Fd_{ox} + 2\ e^{\ominus} \longrightarrow 2\ Fd_{red}$
	$2\ \text{plastocyanin}\ (Cu^{\oplus}) + 2\ P700^{\oplus} \longrightarrow 2\ \text{plastocyanin}\ (Cu^{2+}) + 2\ P700$
FNR:	$2\ Fd_{red} + H^{\oplus} + NADP^{\oplus} \rightleftharpoons 2\ Fd_{ox} + NADPH$
Sum:	$H_2O + 4\ photons + 4\ H^{\oplus}_{in} + NADP^{\oplus} + H^{\oplus} \longrightarrow \frac{1}{2}O_2 + 6\ H^{\oplus}_{out} + NADPH$

reducing agents. (Recall that a reducing agent is one that gives up electrons to reduce another molecule. The reducing agent is oxidized in such reactions.) The oxidized forms of the pigment molecules are $P680^{\oplus}$ and $P700^{\oplus}$. Energy is recovered when P680* and P700* are oxidized and electrons are passed to cytochrome *bf* and NADPH.

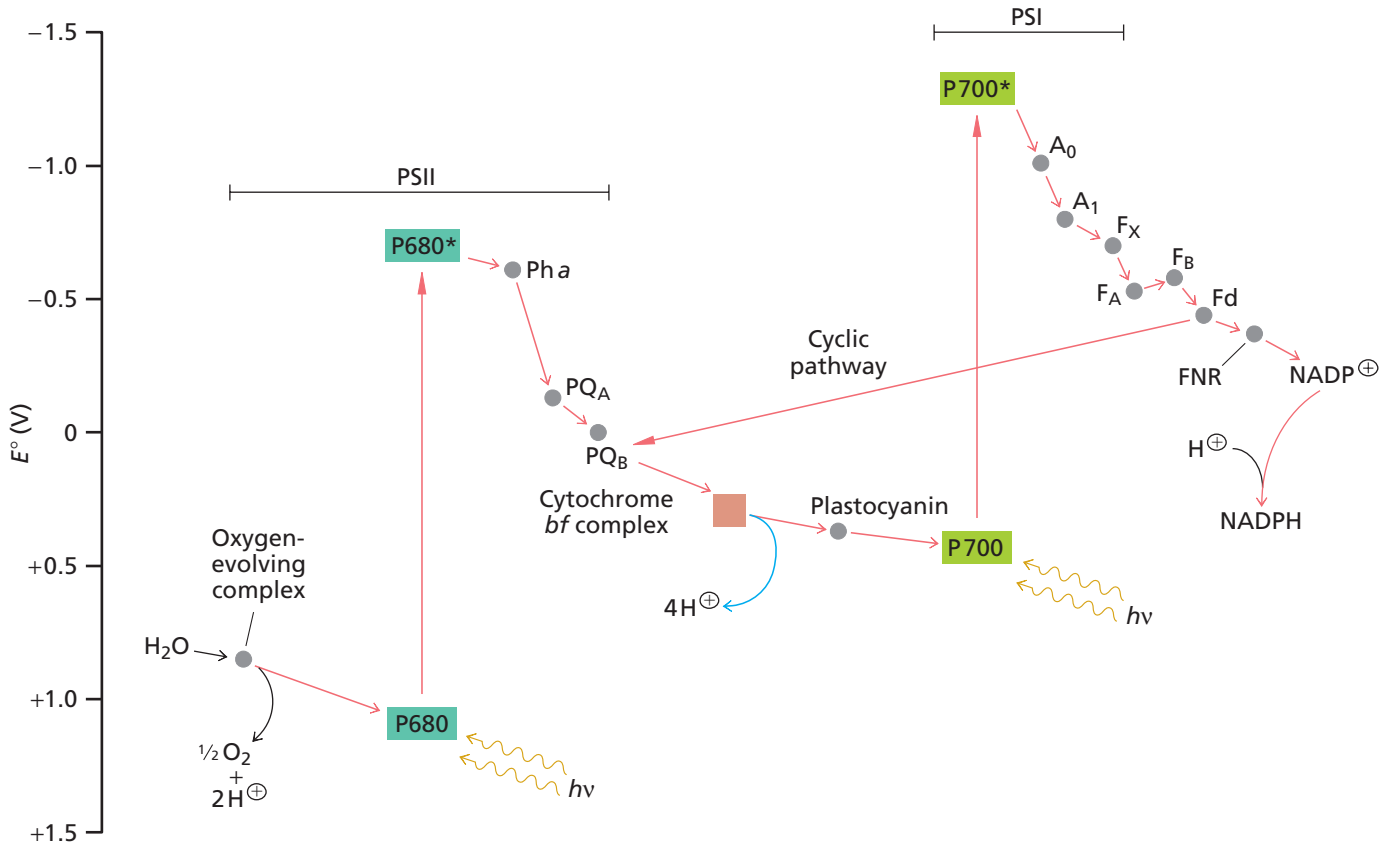
The standard reduction potentials of many of these components are listed in Table 10.5. The difference between any two reduction potentials can be converted to a standard Gibbs free energy change as we saw in Chapter 10. Looking at Figure 15.14 we can see that the absorption of a photon by either P680 or P700 lowers the standard reduction potential by about 1.85 V. In these examples, a difference of 1.85 V corresponds to a standard Gibbs free energy change of about 180 kJ mol^{-1} ($\Delta G^{\circ\prime}$ = 180 kJ mol^{-1}). This value is almost identical to the calculated energy of a "mole" of photons at a wavelength of 680 nm (176 kJ mol^{-1}, Section 15.1). What this means is that the energy of sunlight is very efficiently converted to a change in reduction potential.

There are many similarities between electron transfer in photosynthesis and the membrane-associated electron transport chain that we saw in the last chapter. In both cases electrons pass through a cytochrome complex that transports $H^{\oplus}$ across a membrane. The resulting proton gradient is expended when ATP is synthesized by ATP synthase.

The structure and orientation of cytochrome bc_1 (complex III) and cytochrome *bf* are similar. Both complexes release protons into the space between the inner and outer membranes. The orientation of ATP synthase is also identical—the "head" of the structure is located in the cytoplasm of bacterial cells or the inside compartment of mitochondria. In the next section we'll see that the orientation of ATP synthase in chloroplasts is topologically similar.

KEY CONCEPT

The energy from a photon of light is used to excite an electron in the special pair of chlorophyll molecules. The excited state has a much lower reduction potential making it easy to give up an electron to an oxidation reaction.



▲ **Figure 15.14**

Z-scheme, showing reduction potentials and electron flow during photosynthesis in cyanobacteria. Light energy is absorbed by the special pair pigments, P680 and P700. This converts these molecules into strong reducing agents as shown by the huge drop in standard reduction potential. The values shown are approximate because the reduction potentials of the carriers vary with experimental conditions. The pathway shows the stoichiometry when a *pair* of electrons is transferred from H_2O to NADPH. Abbreviations: Ph *a*, pheophytin *a*, electron acceptor of P680; PQ_A, bound plastoquinone; PQ_B, mobile plastoquinone; A_0, chlorophyll *a*, the primary electron acceptor of P700; A_1, phylloquinone; F_x, F_B, and F_A, iron–sulfur clusters; Fd, ferredoxin; FNR, ferredoxin:$NADP^+$ reductase.

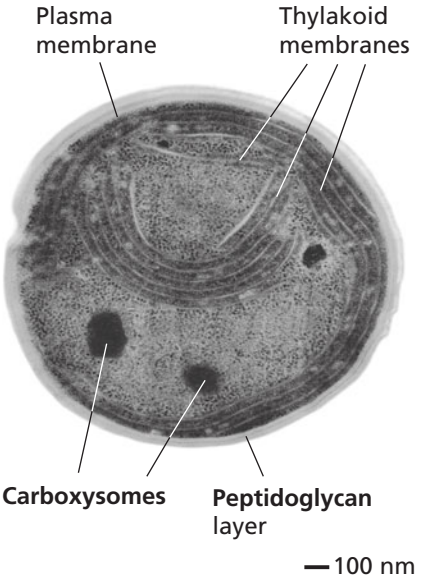
The main difference between photosynthesis and respiratory electron transport is the source of electrons and the terminal electron acceptors. In mitochondria, for example, "high energy" electrons are supplied by reducing equivalents such as NADH ($E°' = -0.32$ V) and accepted by O_2 ($E°' = +0.82$ V) to produce water. In the coupled photosynthesis pathway the flow of electrons is reversed—water ($E°' = +0.82$ V) is the electron donor and $NADP^{\oplus}$ ($E°' = -0.32$ V) is the electron acceptor. This "reversal" of electron flow is thermodynamically unfavorable unless it is coupled to other reactions with a larger Gibbs free energy change. Those other reactions are, of course, the excitation of PSI and PSII by sunlight.

In order to extract electrons from water the cell needs to generate a powerful oxidizing agent with a reduction potential greater than that of the $H_2O \rightarrow 1/2\ O_2 + 2\ H^{\oplus} + 2\ e^{\ominus}$ reaction. This strong oxidizing agent is the P680 special pair after it has given up an electron. The half reaction is $P680^{\oplus} + e^{\ominus} \rightarrow P680°$ ($E°' = +1.1$ V). Note that this standard reduction potential is higher than that of water so that electrons can flow "down" from water to $P680^{\oplus}$ as shown in Figure 15.14. $P680^{\oplus}$ is the most powerful oxidizing agent in biochemical reactions. It is much more potent than $P870^{\oplus}$ in purple bacteria even though purple bacteria have a similar type II reaction center.

Similarly, P700* is a strong reducing agent with a lower reduction potential than $NADP^{\oplus}$. In this case, the absorption of a photon of light by PSI creates an energetic electron that can be passed "down" to $NADP^{\oplus}$ to create reducing equivalents in the form of NADPH. Thus, the "reversal" of electron flow in photosynthesis, compared to respiratory electron transport, is achieved by the special light-absorbing properties of chlorophyll molecules in the two photosystems.

E. Photosynthesis Takes Place Within Internal Membranes

All four of the photosynthesis complexes (PSI, PSII, cytochrome *bf,* and ATP synthase) are embedded in membranes. Most cyanobacteria contain a complex internal network of membranes where these complexes are concentrated (Figure 15.15). The internal membranes are called **thylakoid membranes**. They form by invagination of the inner plasma membrane creating structures that are similar to the mitochondrial cristae. As the membrane folds inward it encloses a space called the **lumen** where protons accumulate during photosynthesis. The thylakoid lumen may remain connected to the periplasmic space or it may form an internal compartment if a membrane loop (or bubble) pinches off from the plasma membrane.



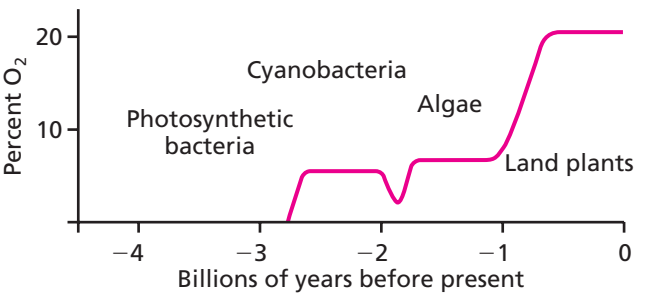
▲ **Figure 15.15**
Internal structure of the cyanobacterium *Synechocystis* PCC 6803. (Carboxysomes are described in Section 15.6A.)

BOX 15.2 OXYGEN "POLLUTION" OF EARTH'S ATMOSPHERE

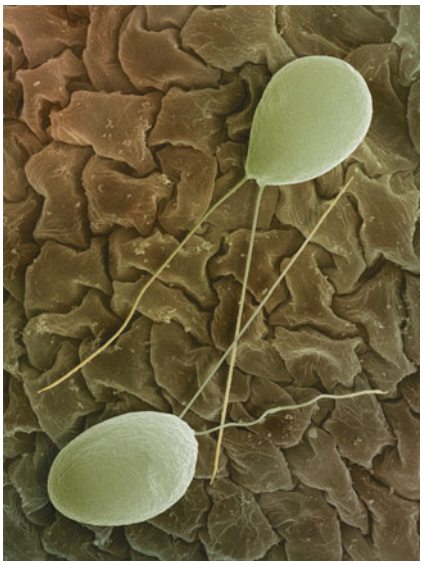
Photosynthetic bacteria probably evolved three billion years ago but the earliest fossil evidence of oxygen producing cyanobacteria dates only from 2.1 billion years ago—claims of much earlier fossils have recently been discredited. The geological record strongly indicates that bacteria began "polluting" the atmosphere with oxygen about 2.4–2.7 billion years ago. This likely corresponds to the evolution of the oxygen evolving complex in PSII and it predates the earliest cyanobacteria fossils.

At that time, oxygen levels rose to about 25% of the present level and they remained at that level for more than a billion years except for a brief drop around 1.9 billion years ago. The cause of this decline isn't known. Primitive plants—probably lichens and mosses—invaded land about 700 million years ago and this led to a steep rise in oxygen levels that eventually reached the present-day concentration of 21%.

Oxygen was highly toxic to most of the species that were around 2 billion years ago but gradually new species arose that could not only tolerate the "pollutant" but used it in respiratory electron transport.



▲ **Oxygen levels in Earth's atmosphere.**



▲ ***Chlamydomonas sp.*** Chlamydomonas species are green algae that are closely related to plants. They contain a single large chloroplast. "Chlamy" is a model organism that is easily grown in the laboratory.

The internal membrane network presents a much greater surface area for membrane proteins. As a result, cyanobacteria contain a much higher concentration of photosynthesis complexes compared to other species of photosynthetic bacteria. This means that cyanobacteria are very efficient at capturing light energy and converting it to chemical energy. This, in turn, has led to their evolutionary success and the formation of an oxygen enriched atmosphere.

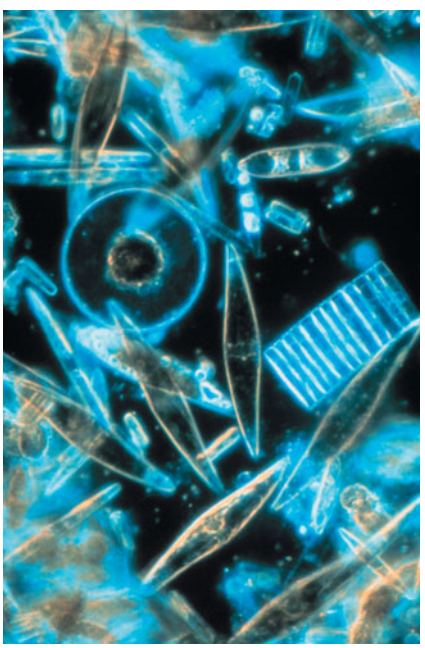
15.3 Plant Photosynthesis

Up to this point we have been describing bacterial photosynthesis but many eukaryotic species are capable of photosynthesis. The photosynthesizing eukaryotes we are most familiar with are flowering plants and other terrestrial species such as mosses and ferns. In addition to these obvious examples, there are many simpler species such as algae and diatoms.

In all photosynthesizing eukaryotes the light-gathering photosystems are localized to a specific cellular organelle called the chloroplast. Thus, unlike bacterial metabolism, photosynthesis and respiratory electron transport are not integrated since they take place in different compartments (chloroplasts and mitochondria). Chloroplasts evolved from a species of cyanobacteria that entered into a symbiotic relationship with a primitive eukaryotic cell over 1 billion years ago. Modern chloroplasts still retain a reduced form of the original bacterial genome. This DNA contains many of the genes for the proteins of the photosystems and genes for some of the enzymes involved in CO_2 fixation. The transcription of these genes and the translation of their mRNAs resemble the prokaryotic mechanisms described in Chapters 21 and 22. This prokaryotic flavor of gene expression reflects the evolutionary origin of chloroplasts.

In the modern world, a large percentage (~70%) of total atmospheric oxygen is produced by photosynthesis in land plants, especially in tropical rain forests. The remaining oxygen is produced by small marine organisms, mostly bacteria, diatoms, and algae. Almost all of the food for animals comes directly or indirectly from plants and the synthesis of these food molecules relies on the energy of sunlight.

A. Chloroplasts

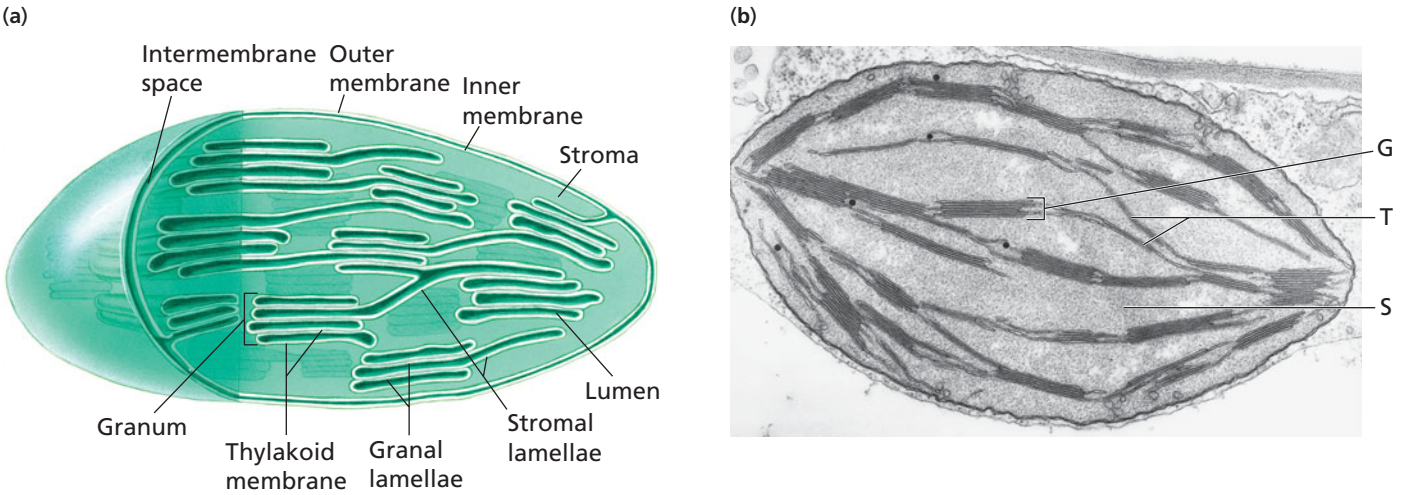


▲ **Diatoms.** About 30% of the oxygen in our atmosphere comes from marine photosynthetic organisms.

The chloroplast is enclosed by a double membrane (Figure 15.16). As in mitochondria, the outer membrane is exposed to the cytoplasm and the inner membrane forms highly folded internal structures. During photosynthesis protons are translocated from the interior of the chloroplast, called the **stroma**, to the compartments between the membranes.

The interior membrane is called the thylakoid membrane. Recall that cyanobacteria possess a similar thylakoid membrane (Figure 15.15). In the chloroplast this membrane forms an extensive network of sheets within the organelle. As the chloroplast develops, projections grow out from these sheets to form flattened disk-like structures. These disk-like structures stack on top of one another like a pile of coins to form **grana** (singular, granum). A typical chloroplast contains dozens of grana, or stacked disks of thylakoid membranes. The grana in mature chloroplasts are connected to each other by thin sheets of thylakoid membrane called stroma thylakoids. These stroma thylakoid membranes are exposed to the stroma on both surfaces whereas grana thylakoid membranes within a stack are in close contact with the membranes immediately above and below them.

The three-dimensional organization of the thylakoid membrane is shown in Figure 15.17. Each disk in the stack is connected to the stroma thylakoids by short bridges. The interior of each disk is called the lumen and it is the same compartment as the region between the two membranes of the stroma thylakoid. All thylakoid membranes are likely derived from the inner chloroplast membrane. This means that the lumen is topologically equivalent to the space between the inner and outer membranes of the chloroplast although in some cases the direct connection may be lost. The thylakoid membranes contain PSI, PSII, cytochrome *bf*, and ATP synthase complexes as in cyanobacteria. In mitochondria, protons accumulate in the compartment between the inner and outer membranes (Section 14.3); similarly, in chloroplasts, protons are translocated into the thylakoid lumen and the space between the two membranes of



▲ **Figure 15.16**
Structure of the chloroplast. (a) Illustration. **(b)** Electron micrograph: cross-section of a chloroplast from a spinach leaf. Shown are grana (G), the thylakoid membrane (T), and the stroma (S).

the stroma thylakoids. It's important to keep in mind that the chloroplast stroma is equivalent to the cytoplasm in bacteria and the matrix in mitochondria.

B. Plant Photosystems

The photosynthesis complexes in eukaryotic chloroplasts evolved from the complexes present in primitive cyanobacteria. Chloroplast PSI is structurally and functionally similar to its bacterial ancestor—the only significant structural difference is that eukaryotic PSI contains chlorophyll molecules instead of bacteriochlorophyll in the electron transfer chain of the reaction center. The eukaryotic version oxidizes plastocyanin (or cytochrome *c*) and reduces ferredoxin (or flavodoxin). Eukaryotic PSI associates with a light-harvesting complex called LHCI that resembles the complex found in some bacteria.

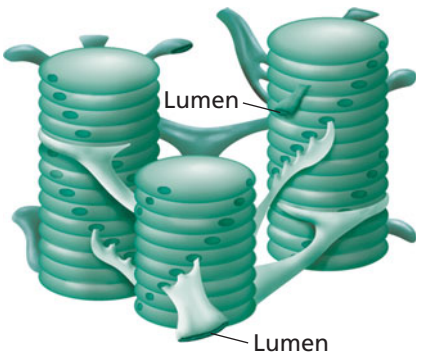
Chloroplast PSII is also similar to the one in cyanobacteria. Plant chloroplasts contain a light-harvesting complex called LHCII that associates with PSII in the chloroplast membrane. LHCII is a large structure containing 140 chlorophylls and 40 carotenoids and it completely surrounds PSII. As a result, photon capture in plants is more efficient than in bacteria. Cyanobacteria and chloroplasts contain similar cytochrome *bf* complexes.

The ATP synthase in chloroplasts is related to the cyanobacterial ATP synthase, as expected. The protein components differ from the mitochondrial version described in the previous chapter. This is not surprising since the mitochondrial ATP synthase evolved from the proteobacterial ancestor of bacteria and proteobacteria are distantly related to cyanobacteria. Species such as algae, diatoms, and plants that contain both mitochondria and chloroplasts have distinctive versions of ATP synthase in each organelle.

The chloroplast ATP synthase is a CF_0F_1 ATPase where the "C" stands for chloroplast. The overall molecular structure is very similar to that of mitochondria even though the various subunits of the two enzymes are encoded by different genes. As in mitochondria, the membrane component of the chloroplast ATP synthase consists of a multimeric ring and a rod that projects into a hexameric head structure. The ring rotates as protons move across the membrane and ATP is synthesized from ADP + P_i by a binding change mechanism as described in Section 14.9. The "knob" projects into the chloroplast stroma (Figure 15.18).

C. Organization of Chloroplast Photosystems

Figure 15.19 illustrates the locations of the membrane-spanning photosynthetic components within the chloroplast thylakoid membrane. PSI is located predominantly in the stroma thylakoid and is therefore exposed to the chloroplast stroma. PSII is located

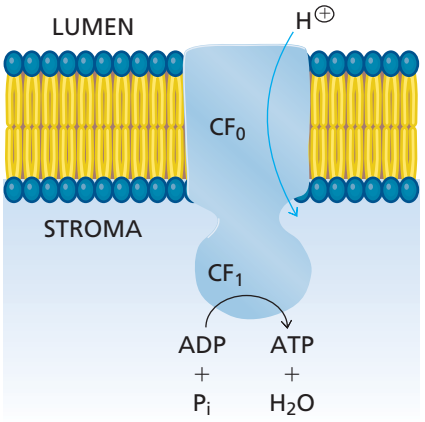


▲ **Figure 15.17**
Organization of stacked disks in a granum and their connection to the stroma thylakoids.
Adapted from Staehlin, L. A. (2003) Chloroplast structure: from chlorophyll granules to supramolecular architecture of thylakoid membranes. *Photosynthesis Research* 76:185–196.

The locations of various photosynthetic components in the stroma and grana thylakoid membranes are shown in Figure 15.19.

KEY CONCEPT

Photosynthetic bacteria and chloroplasts make use of internal thylakoid membranes to increase the number of photosystem complexes.



▲ **Figure 15.18**
Chloroplast ATP synthase.

predominantly in the grana thylakoid membrane, away from the stroma. The oxygen-evolving complex is associated with PSII on the luminal side of the thylakoid membrane. The cytochrome *bf* complex spans the thylakoid membrane and is found in both the stroma and grana thylakoid membranes. ATP synthase is found exclusively in the stroma thylakoids with the CF_1 component, the site of ATP synthesis, projecting into the stroma.

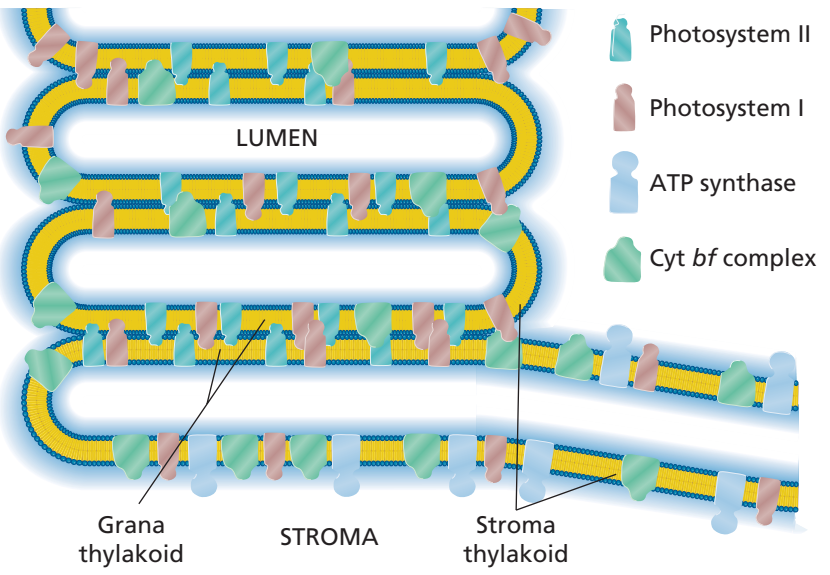
The membranes of the top and bottom surfaces of each disk in a granum are in contact with each other forming a double-membrane structure. This region is densely packed with the light-absorbing PSII complexes and their associated LHCII complexes. Light passes through the plasma membrane of the plant cell, through the cytoplasm, and across the outer membrane of the chloroplast. When light reaches the grana, the photons are efficiently absorbed by the pigment molecules in the membrane.

Excited electrons are transferred within PSII to PQ forming PQH_2. The protons for this reaction are taken up from the stroma. The PSII reaction center is replenished with electrons from the oxidation of water taking place in the lumen. PQH_2 diffuses within the membrane to the cytochrome *bf* complex where it is oxidized to PQ. The protons released in the Q cycle enter the lumen. Electrons are passed to plastocyanin that diffuses freely in the lumen to reach PSI. PSI absorbs light leading to the transfer of electrons from reduced plastocyanin to ferredoxin. Ferredoxin is formed in the stroma. It can participate in the reduction of $NADP^{\oplus}$ to NADPH in the stroma or serve as an electron donor to cytochrome *bf* complexes in the stroma thylakoid membrane (cyclic electron transport, Section 15.2B).

Note that PSII is not directly exposed to the stroma but is exposed to the thylakoid lumen. The lumen is topologically equivalent to the outside of the bacterial membrane as shown in Figure 15.11. PSI projects into the stroma compartment since it produces ferredoxin that accumulates within chloroplasts. The stroma is topologically equivalent to the bacterial cytoplasm (inside the cell). The distribution of cytochrome *bf* complexes is explained by the fact they can receive electrons from both PSII and PSI. Supercomplexes of PSII and cytochrome *bf* in the grana participate in linear electron transfer from water to plastocyanin. In the stroma thylakoids there are complexes of PSI, cytochrome *bf*, and ferredoxin:quinone oxidoreductase (FQR) that are involved in cyclic electron flow.

The proton gradient is used to generate ATP. As protons are translocated from the lumen compartment to the stroma, ATP is synthesized from ADP and P_i in the stroma. Both ATP and NADPH accumulate in the stroma where they can be used in biosynthesis reactions. In plants, but not other photosynthetic species, a high percentage of ATP and NADPH molecules are used in the fixation of CO_2 and the synthesis of carbohydrates.

Figure 15.19 ▶
Distribution of membrane-spanning photosynthetic components between stroma and granal thylakoids. PSI is found predominantly in stroma thylakoids. PSII is found predominantly in grana thylakoids. The cytochrome *bf* complex in found in both stroma and grana thylakoid membranes. ATP synthase is localized exclusively to stroma thylakoids.

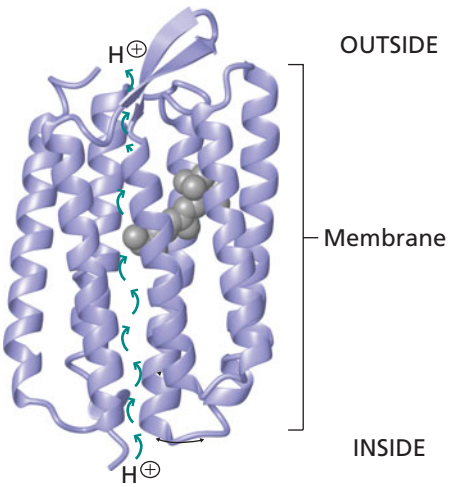


BOX 15.3 BACTERIORHODOPSIN

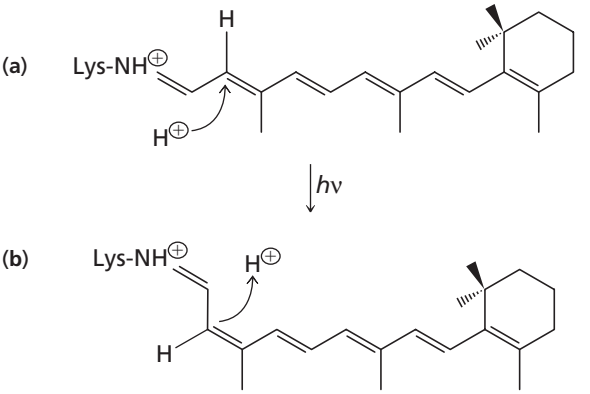
Bacteriorhodopsin is a membrane protein found in a few specialized species of archaebacteria such as *Halobacterium salinarium*. The protein has seven membrane-spanning α helices that form a channel in the membrane. (See ribbon structure below.) A single retinal molecule is covalently bound to a lysine side chain in the middle of the channel. The normal configuration of the retinal is all-*trans* but when it absorbs a photon of light it converts to the 13-*cis* configuration. (See structure below.) The light-induced change in configuration is coupled to deprotonation and reprotonation of the retinol molecule.

When light is absorbed, the shift in configuration to 13-*cis* retinal releases a proton that then passes up the channel to be released on the outside of the membrane. This proton is replaced by a proton that is taken up from the cytosol and the retinol configuration shifts back to the all-*trans* form. For every photon of light that is absorbed by bacteriorhodopsin a single proton is translocated across the membrane.

Bacteriorhodopsin creates a light-induced proton gradient and this proton gradient drives ATP synthesis by ATP synthase.

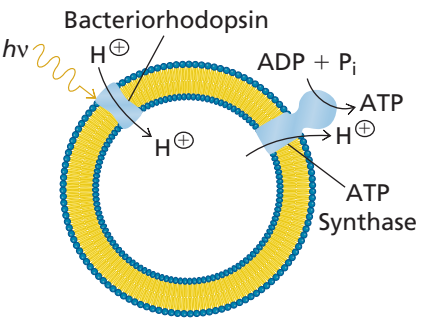


▲ **Bacteriorhodopsin.**



▲ **Two configurations of retinal-lysine in bacteriorhodopsin. (a)** All-*trans* retinal. **(b)** 13-*cis* retinal. The configuration shifts from the all-*trans* form to the 13-*cis* form when a photon of light is absorbed.

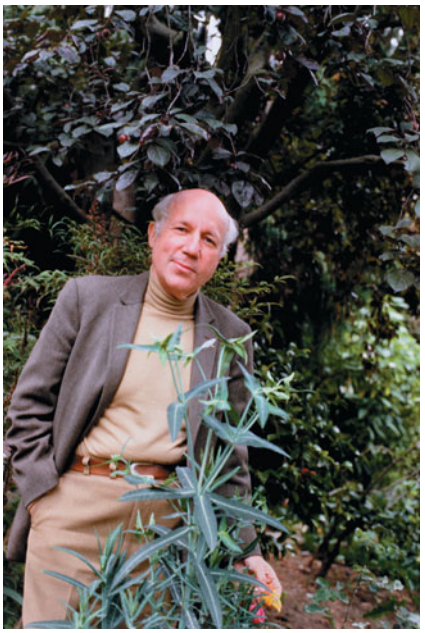
The coupling of bacteriorhodopsin and ATP synthase can be directly demonstrated by artificially synthesizing lipid vesicles containing both complexes. In the orientation shown below, the vesicles will synthesize ATP from ADP + P_i when they are illuminated. This experiment, first carried out by Efraim Racker and his colleagues in 1974, was one of the first confirmations of the chemiosmotic theory (Section 14.3).



▲ **Bacteriorhodopsin creates a proton gradient that drives ATP synthesis.** Artificial lipid vesicles containing bacteriorhodopsin and ATP synthase were created with the orientation shown. When these vesicles were illuminated, bacteriorhodopsin pumped protons into the vesicle and the resulting proton gradient activated ATP synthase.

15.4 Fixation of CO_2: The Calvin Cycle

In photosynthetic species there is a special pathway for the reductive conversion of atmospheric CO_2 to carbohydrates. The reactions are powered by the ATP and NADPH formed during the light reactions of photosynthesis. The fixation of CO_2 and the synthesis of carbohydrates occurs in the cytoplasm of bacteria and in the chloroplast stroma. This biosynthesis pathway is a cycle of enzyme-catalyzed reactions with three major stages: (1) the carboxylation of a five-carbon sugar molecule, (2) the reductive synthesis of carbohydrate for use in other pathways, and (3) the regeneration of the molecule that accepts CO_2. This pathway of carbon assimilation has several names, such



▲ **Melvin Calvin (1911–1997).** Calvin won the Nobel Prize in Chemistry in 1961 for his work on carbon dioxide assimilation in plants.

lbl.gov/Science-Articles/Research-Review/Magazine/1997/story12.html]

KEY CONCEPT

The Calvin cycle utilizes the products of photosynthesis, ATP and NADPH, to fix CO_2 into carbohydrates.

as *the reductive pentose phosphate cycle*, *the C_3 pathway* (the first intermediate is a three-carbon molecule), and *the Calvin cycle*. (Workers in Melvin Calvin's laboratory discovered the carbon-fixing pathway using $^{14}CO_2$ tracer experiments in algae.) We refer to the pathway as the **Calvin cycle**.

The fixation of CO_2 and the synthesis of carbohydrates are often described as "photosynthesis." In this textbook we refer to photosynthesis and the Calvin cycle as two separate pathways.

A. The Calvin Cycle

The Calvin cycle is outlined in Figure 15.20. The first stage is the carboxylation of ribulose 1,5-*bis*phosphate, a reaction catalyzed by the enzyme ribulose 1,5-*bis*phosphate carboxylase–oxygenase, better known as Rubisco. The second stage is a reduction stage where 3-phosphoglycerate is converted to glyceraldehyde 3-phosphate. Most of the glyceraldehyde 3-phosphate is converted to ribulose 1,5-*bis*phosphate in the third (regeneration) stage. Some of the glyceraldehyde 3-phosphate produced in the Calvin cycle is used in carbohydrate synthesis pathways. Glyceraldehye 3-phosphate is the main product of the Calvin cycle.

Figure 15.21 on page 464 shows all reactions of the Calvin cycle. The pathway begins with steps for assimilating *three* molecules of carbon dioxide because the smallest carbon intermediate in the Calvin cycle is a C_3 molecule. Therefore, three CO_2 molecules must be fixed before one C_3 unit (glyceraldehyde 3-phosphate) can be removed from the cycle without diminishing the metabolic pools.

B. Rubisco: Ribulose 1,5-*bis*phosphate Carboxylase–oxygenase

Rubisco (ribulose 1,5-*bis*phosphate carboxylase–oxygenase) is the key enzyme of the Calvin cycle. It catalyzes the fixation of atmospheric CO_2 into carbon compounds. This reaction involves the carboxylation of the five-carbon sugar, ribulose 1,5-*bis*phosphate, by CO_2. This leads to the eventual release of two three-carbon molecules of 3-phosphoglycerate. The reaction mechanism of Rubisco is shown in Figure 15.22.

Rubisco makes up about 50% of the soluble protein in plant leaves, making it one of the most abundant enzymes on Earth. Interestingly, its status as an abundant enzyme is due partly to the fact that it is not very efficient—the low turnover number of ~3 s^{-1} means that large amounts of the enzyme are required to support CO_2 fixation!

The Rubisco of plants, algae, and cyanobacteria is composed of eight large (L) subunits and eight small (S) subunits (Figure 15.23). There are eight active sites located in the eight large subunits. Four additional small subunits are located at each end of the core formed by the large subunits. The Rubisco molecules in other photosynthetic bacteria have only the large subunits containing the active sites. For example, in the purple bacterium *Rhodospirillum rubrum*, Rubisco consists of a simple dimer of large subunits.

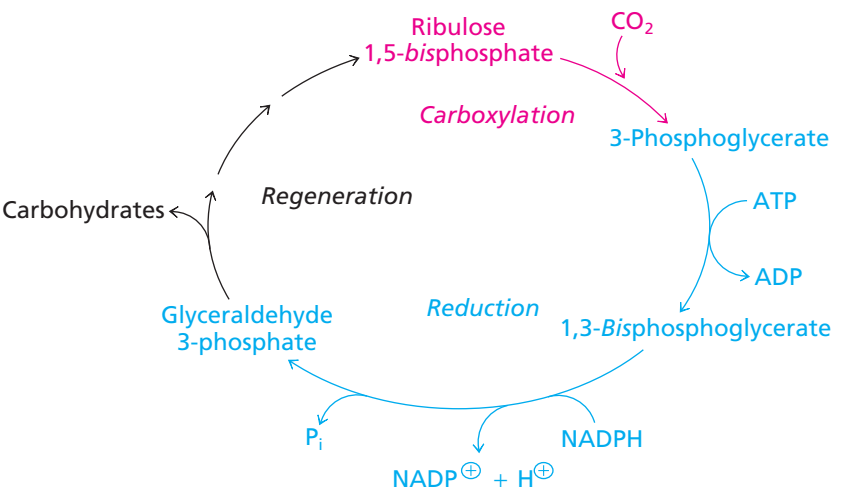
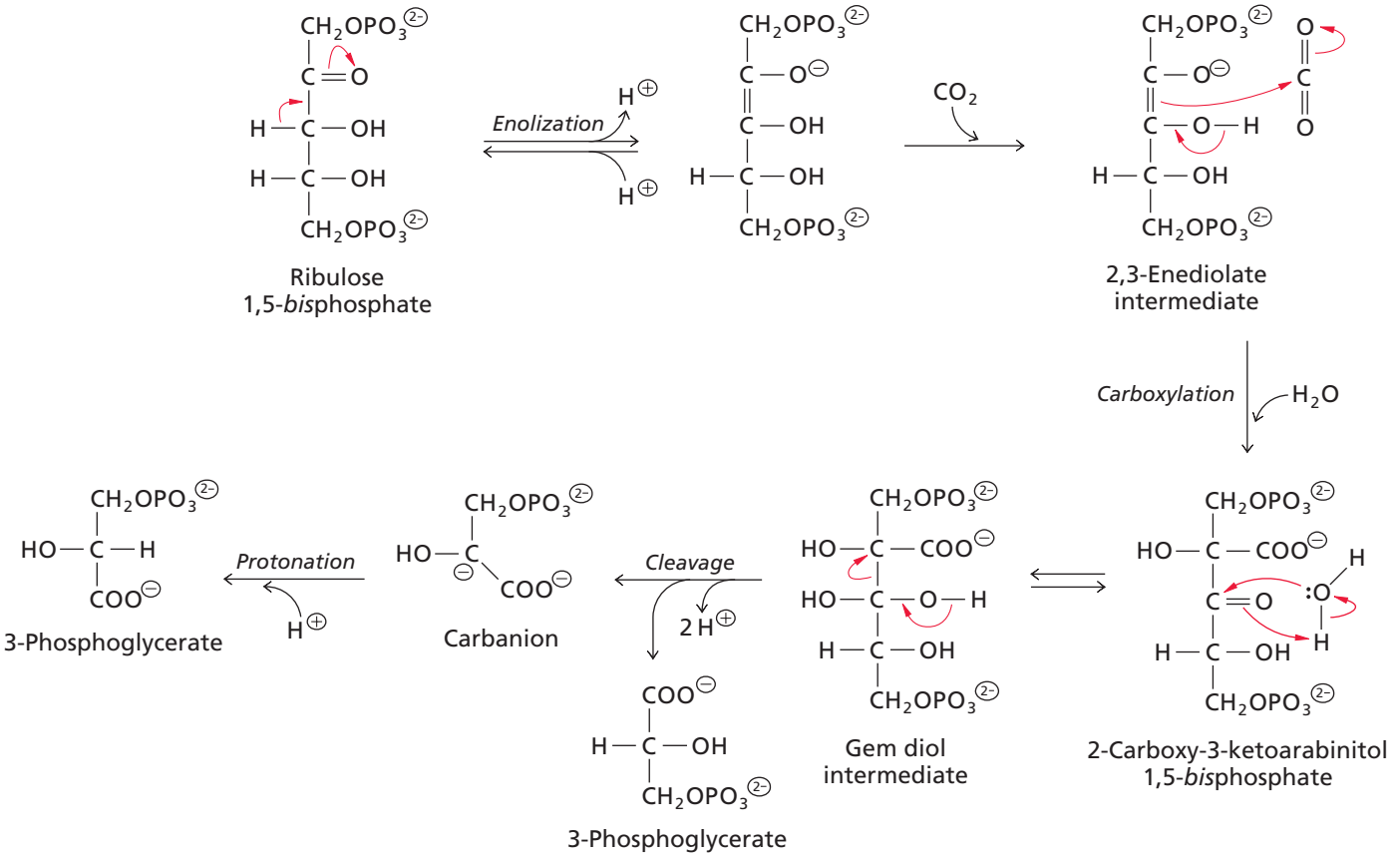
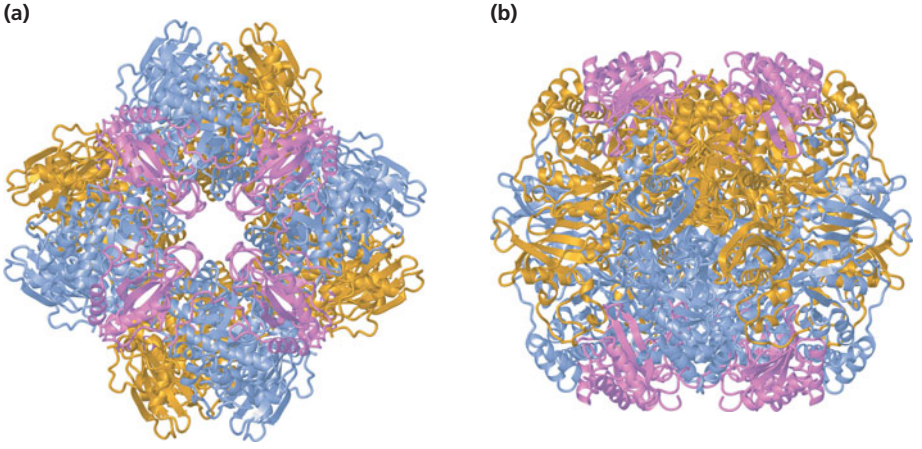


Figure 15.20 ►
Summary of the Calvin cycle. The cycle has three stages: carboxylation of ribulose 1,5-*bis*phosphate, reduction of 3-phosphoglycerate to glyceraldehyde 3-phosphate, and regeneration of ribulose 1,5-*bis*phosphate.



▲ **Figure 15.22**
Mechanism of Rubisco-catalyzed carboxylation of ribulose 1,5-*bis*phosphate to form two molecules of 3-phosphoglycerate. A proton is abstracted from C-3 of ribulose 1,5-*bis*phosphate to create a 2,3-enediolate intermediate. The nucleophilic enediolate attacks CO_2, producing 2-carboxy-3-ketoarabinitol 1,5-*bis*phosphate, which is hydrated to an unstable gem diol intermediate. The C-2–C-3 bond of the intermediate is immediately cleaved, generating a carbanion and one molecule of 3-phosphoglycerate. Stereospecific protonation of the carbanion yields a second molecule of 3-phosphoglycerate. This step completes the carbon fixation stage of the RPP cycle—two molecules of 3-phosphoglycerate are formed from CO_2 and the five-carbon sugar ribulose 1,5-*bis*phosphate.

The purple bacterium version of Rubisco has a much lower affinity for CO_2 than the more complex multisubunit enzymes in other species but it catalyzes the same reaction. In a spectacular demonstration of this functional similarity, tobacco plants were genetically engineered by replacing the normal plant gene with the one from the purple bacterium *Rhodospirillum rubrum*. The modified plants contained only the dimeric bacterial form of the enzyme but they grew normally and reproduced as long as they were kept in an atmosphere of high CO_2 concentration.



◀ **Figure 15.23**
The quaternary structure (L_8S_8) of ribulose 1,5-*bis*phosphate carboxylase–oxygenase (Rubisco). (a) Top and **(b)** side views of the enzyme from spinach (*Spinacia oleracea*). Large subunits are shown alternately yellow and blue; small subunits are purple. [PDB 1RCX].

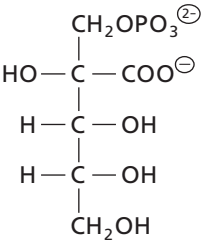
▲ **Figure 15.21**

Calvin cycle. The concentrations of Calvin cycle intermediates are maintained when one molecule of glyceraldehyde 3-phosphate (G3P) exits the cycle after three molecules of CO_2 are fixed.

Rubisco cycles between an active form (in the light) and an inactive form (in the dark). It must be activated to catalyze the fixation of CO_2. In the light, Rubisco activity increases in response to the higher, more basic pH that develops in the stroma (or bacterial cytoplasm) during proton translocation. Under alkaline conditions an activating molecule of CO_2, which is not the substrate CO_2 molecule, reacts reversibly with the side chain of a lysine residue of Rubisco to form a carbamate adduct. Mg^{2+} binds to and stabilizes this CO_2–lysine adduct. The enzyme must be carbamylated in order to carry out CO_2 fixation; however, the carbamate adduct readily dissociates, making the enzyme inactive. Carbamylation is normally inhibited because Rubico is usually in an inactive conformation. During the day, a light-activated ATP-dependent enzyme called Rubisco activase binds to Rubisco and facilitates carbamylation by inducing a conformational change. Under these conditions Rubisco is active.

When the sun goes down Rubisco activase is no longer effective in activating Rubisco and CO_2 fixation stops. This regulation makes sense since photosynthesis is not active at night and ATP + NADPH are not produced in chloroplasts during the night. These cofactors are required for the Calvin cycle so the Calvin cycle is not active at night as a result of the regulation of Rubisco activity. Inhibition of Rubisco in the dark prevents the inefficient accumulation of 3-phosphoglycerate and the wasteful oxygenation reaction described in the next section.

In plants, an additional level of inhibition is mediated by 2-carboxyarabinitol 1-phosphate (Figure 15.24). This compound is an analog of the unstable gem diol intermediate of the carboxylation reaction. It is synthesized only at night and it binds to, and inhibits, any residual carbamylated Rubisco, thus ensuring that the Calvin cycle is shut down. Some plants synthesize sufficient amounts of the inhibitor to keep Rubisco completely inactive in the dark.



▲ **Figure 15.24**
2-Carboxyarabinitol 1-phosphate.

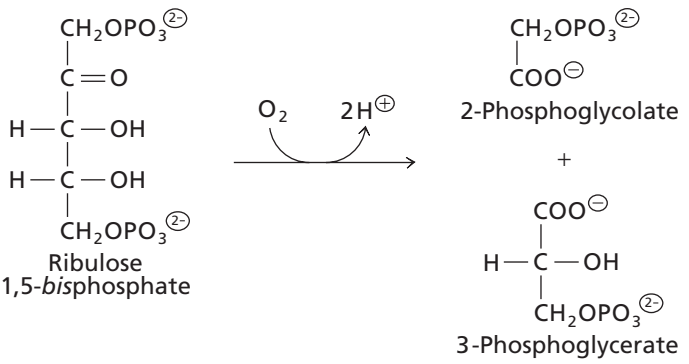
C. Oxygenation of Ribulose 1,5-*bis*phosphate

As its complete name indicates, ribulose 1,5-*bis*phosphate carboxylase–oxygenase catalyzes not only carboxylation but also the oxygenation of ribulose 1,5-*bis*phosphate. The two reactions are competitive since CO_2 and O_2 compete for the same active sites on Rubisco. The oxygenation reaction produces one molecule of 3-phosphoglycerate and one molecule of 2-phosphoglycolate (Figure 15.25). Oxygenation consumes significant amounts of ribulose 1,5-*bis*phosphate *in vivo*. Under normal growth conditions, the rate of carboxylation is only about three to four times the rate of oxygenation.

The 3-phosphoglycerate formed from the oxygenation of ribulose 1,5-*bis*phosphate enters the Calvin cycle. The other product of the oxygenation reaction follows a different pathway. Two molecules of 2-phosphoglycolate (C_2) are metabolized in peroxisomes and the mitochondria by an oxidative pathway (via glyoxylate and the amino acids glycine and serine) to one molecule of CO_2 and one molecule of 3-phosphoglycerate (C_3), which also enters the Calvin cycle. This oxidative pathway consumes NADH and ATP. The light-dependent uptake of O_2 catalyzed by Rubisco and followed by the release of CO_2 during the metabolism of 2-phosphoglycolate is called *photorespiration.* Like carboxylation, photorespiration is normally inhibited in darkness when Rubisco is inactive. The appreciable release of fixed CO_2 and the consumption of

KEY CONCEPT

Some enzymes cannot distinguish between very similar substrates.



◀ **Figure 15.25**
Oxygenation of ribulose 1,5-*bis*phosphate catalyzed by Rubisco.

BOX 15.4 BUILDING A BETTER RUBISCO

Many labs are attempting to genetically modify plants in order to enhance the carboxylation reaction and suppress the oxygenation reaction. If successful, these attempts to make a better Rubisco could greatly increase food production.

The "perfect" enzyme would have very low oxygenase activity and very efficient carboxylase activity. The kinetic parameters of the oxygenase activity of Rubisco enzymes from several species are listed in the accompanying table. The low catalytic efficiency of the enzyme is indicated by the k_{cat}/K_m values.

Kinetic parameters of Rubisco carboxylase activity in various species

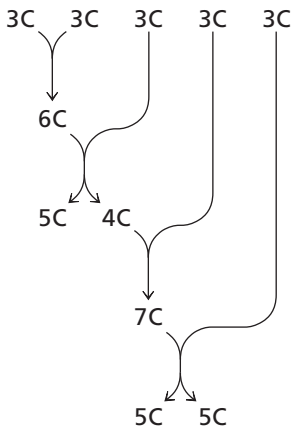
Species	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
Tobacco	3.4	10.7	3.2×10^5
Red algae	2.6	9.3	2.8×10^5
Purple bacteria	7.3	89	8.2×10^4
"Perfect" enzyme	1070	10.7	10^8

Data from Andrews, J. T., and Whitney, S. M. (2003). Manipulating ribulose *bis*phosphate carboxylase/oxygenase in the chloroplasts of higher plants. *Arch. Biochem. Biophys.* 414: 159–169.

These values should be compared to those in Table 5.2. It seems likely that the carboxylase efficiency can be improved 1000-fold by modifying the amino acid side chains in the active site.

The difficult part of the genetic modification is choosing the appropriate amino acid changes. The choice is informed by a detailed knowledge of the structures of several Rubisco enzymes from different species and by examination of the contacts between amino acid side chains and substrate molecules. Models of the presumed transition states are also important. Additional key residues can be identified by comparing the conservation of amino acid sequences in enzymes from a wide variety of species

The underlying strategy assumes that evolution has not yet selected for the most well-designed enzyme. This assumption seems reasonable since there are many examples of ongoing evolution in biochemistry. However, several billion years of evolution have not resulted in a better Rubisco and neither have several decades of human effort. It may not be possible to build a better Rubisco.



▲ **Figure 15.26**
Outline of the regeneration stage of the Calvin cycle.

energy as a result of oxygenation—with no apparent benefit to the organism—arise from the lack of absolute substrate specificity of Rubisco. This is a serious problem in agriculture because photorespiration limits crop yields.

D. Calvin Cycle: Reduction and Regeneration Stages

The reduction stage of the Calvin cycle begins with the ATP-dependent conversion of 3-phosphoglycerate to 1,3-*bis*phosphoglycerate in a reaction catalyzed by phosphoglycerate kinase. Next, 1,3-*bis*phosphoglycerate is reduced by NADPH (not NADH, as in gluconeogenesis, Section 11.2#6) in a reaction catalyzed by a glyceraldehyde 3-phosphate dehydrogenase isozyme. As in gluconeogenesis, some of the glyceraldehyde 3-phosphate is rearranged to its isomer, dihydroxyacetone phosphate, by triose phosphate isomerase. For every six glyceraldehyde 3-phosphate molecules produced by this pathway, one is removed from the cycle to be used in carbohydrate synthesis and the five others are used in the regeneration stage.

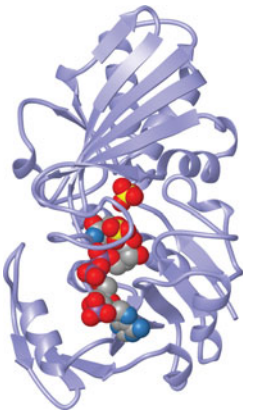
In the regeneration stage, glyceraldehyde 3-phosphate is diverted into three different branches of the pathway and is interconverted between three-carbon (3C), four-carbon (4C), five-carbon (5C), six-carbon (6C), and seven-carbon (7C) phosphorylated sugars (Figure 15.21). The pathway is schematically outlined in Figure 15.26. Two of the reactions, those catalyzed by aldolase and fructose 1.6-*bis*phosphatase, are familiar because they are part of the gluconeogenesis pathway (Section 12.1). Many of the other reactions are part of the normal pentose phosphate pathway (Section 12.4) including two tranketolase reactions. The net result of the Calvin cycle reactions is

$$3\ CO_2 + 9\ ATP + 6\ NADPH + 5\ H_2O \longrightarrow$$
$$\text{glyceraldehyde 3-phosphate} + 9\ ADP + 8\ P_i + 6\ NADP^{\oplus} + 2H^+ \qquad \textbf{(15.2)}$$

Both ATP and NADPH are required for CO_2 fixation by the Calvin cycle. These are the major products of the light reactions of photosynthesis. The fact that the requirement for ATP exceeds that of NADPH is one reason why cyclic electron flow from PSI to cytochrome *bf* is important in photosynthesis. Cyclic electron flow results in increased production of ATP relative to NADPH.

It's interesting to compare the cost of synthesizing carbohydrates from CO_2 and the energy yield from degrading it via glycolysis and the citric acid cycle. We can use Reaction 15.2 to estimate the cost of synthesizing acetyl CoA—the substrate for the citric acid cycle. Recall that the pathway from glyceraldehyde 3-phosphate to acetyl CoA is coupled to the synthesis of two molecules of NADH and two molecules of ATP (Section 11.2). If we subtract these from the cost of making glyceraldehyde 3-phosphate then the total cost of synthesizing acetyl CoA from CO_2 is 7 ATP + 4 NAD(P)H. This can be expressed as 17 ATP equivalents since each NADH is equivalent to 2.5 ATP (Section 14.11). The net gain from complete oxidation of acetyl CoA by the citric acid cycle is 10 ATP equivalents (Section13.4). The biosynthesis pathway is more expensive than the energy gained from catabolism. In this case, the "efficiency" of acetyl CoA oxidation is only about 60% (10/17 = 59%) but this value is misleading since it's actually the biosynthesis pathway (costing 17 ATP equivalents) that is complex and inefficient.

We can estimate the cost of synthesizing glucose because it is simply the cost of making two molecules of glyceraldehyde 3-phosphate. It's equivalent to 18 molecules of ATP and 12 molecules of NADPH or 48 ATP equivalents. Recall that the net gain of energy from the complete oxidation of glucose via glycolysis and the citric acid cycle is 32 ATP equivalents (Section 13.4). In this case, catabolism recovers two-thirds of amount of the ATP equivalents used in the biosynthesis pathway.



▲ **Glyceraldehyde 3-phosphate dehydrogenase.** This NADPH-dependent enzyme from spinach (*Spinacia oleracea*) crystallizes as a tetramer. Only a single subunit is shown here. NADPH is bound in the active site of the enzyme. [PDB 2PKQ]

15.5 Sucrose and Starch Metabolism in Plants

Glyceraldehyde 3-phosphate (G3P) is the main product of carbon fixation in most photosynthetic species. G3P is subsequently converted to glucose by the gluconeogenesis pathway. Newly synthesized hexoses can be used immediately as substrates in a number of biosynthesis pathways or they can be stored as polysaccharides for use later on. In bacteria, most algae, and some plants, the storage polysaccharide is glycogen, just as in animals. The storage polysaccharide in vascular plants is usually starch.

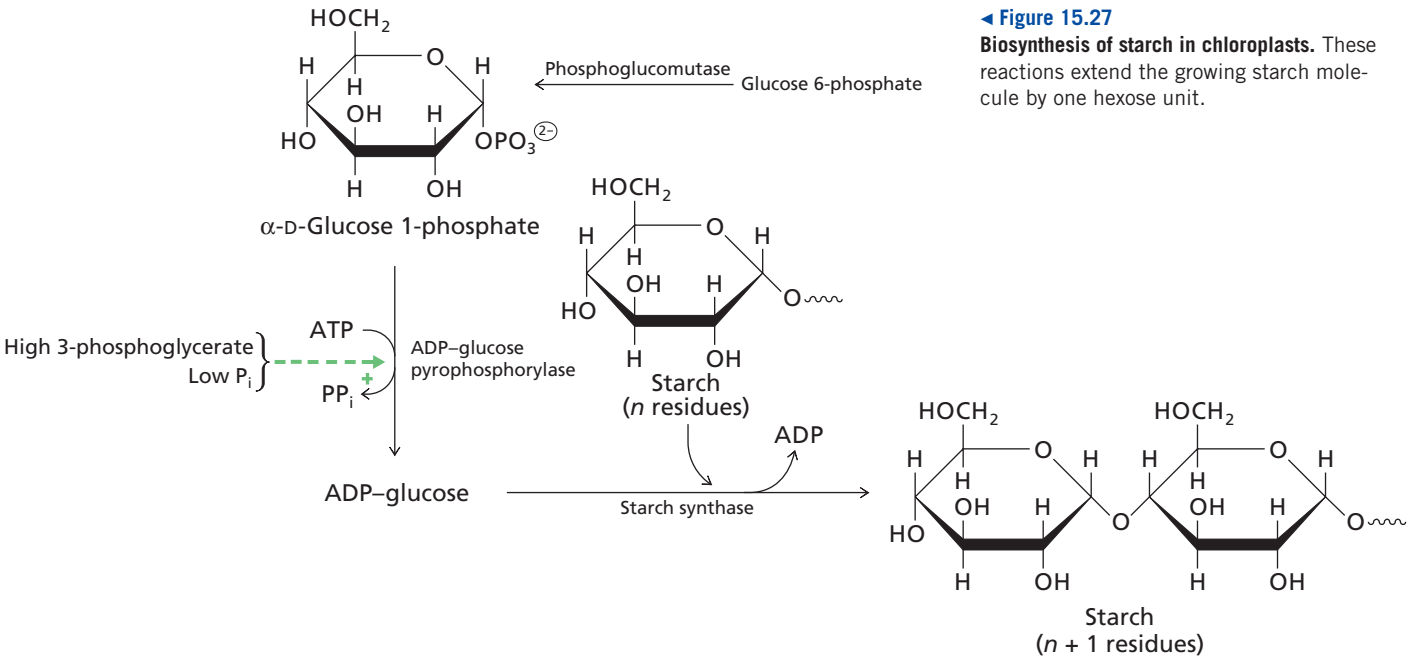
Starch is synthesized in chloroplasts from glucose 6-phosphate, the primary product of gluconeogenesis (Section 12.1D). In the first step, glucose 6-phosphate is converted to glucose 1-phosphate in a reaction catalyzed by phosphoglucomutase (Figure 15.27). This is the same enzyme we encountered in the glycogen synthesis pathway (Section 12.5A). The second step is the activation of glucose by synthesis of ADP–glucose. This reaction is catalyzed by ADP–glucose pyrophosphorylase. The metabolic strategy is similar to that of glycogen biosynthesis except that the key intermediate in glycogen

KEY CONCEPT

The energy recovered in catabolic pathways is usually about two-thirds of the energy used in biosynthesis.

The structures of starch and glycogen are described in Section 8.6A.

The nucleotide sugar ADP–glucose is also required for synthesis of glycogen by some bacteria (Section 12.5A).



◀ **Figure 15.27**
Biosynthesis of starch in chloroplasts. These reactions extend the growing starch molecule by one hexose unit.



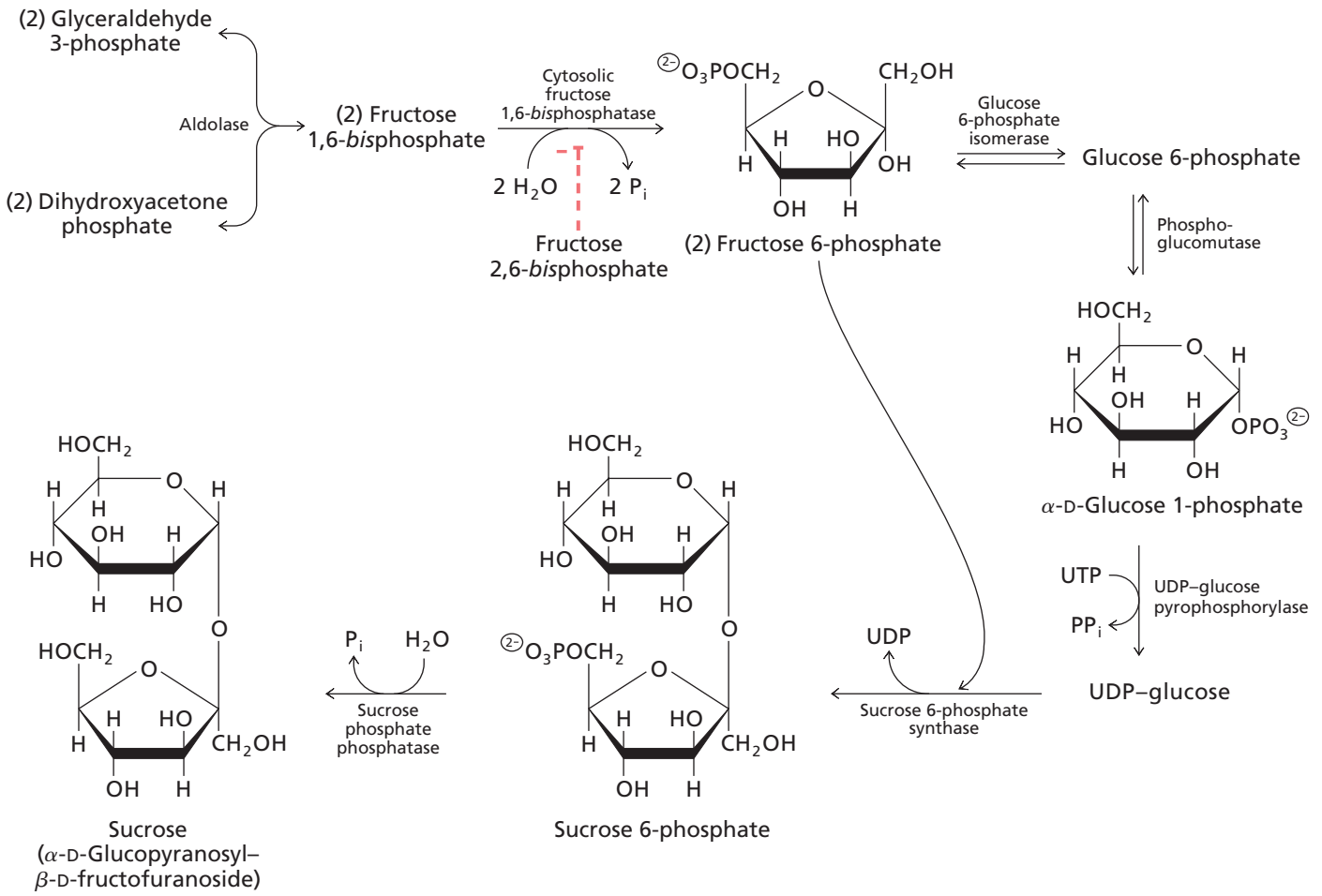
▲ **Maple syrup.** The sucrose-rich sap of maple trees is collected and concentrated to produce maple syrup.

synthesis is UDP–glucose. The polymerization reaction in starch biosynthesis is carried out by starch synthase. This pathway consumes one molecule of ATP and releases one molecule of pyrophosphate for each residue that is added to the growing polysaccharide chain. ATP is supplied by the reactions of photosynthesis.

Starch is synthesized in daylight when photosynthesis is active and ATP molecules accumulate within the chloroplast. During the night starch becomes a source of carbon and energy for the plant. The starch molecule is cleaved by the action of starch phosphorylase to generate glucose 1-phosphate that is converted to triose phosphates by glycolysis. The triose phosphates are exported from the chloroplast to the cytoplasm. Alternatively, starch can be hydrolyzed by the action of amylases to dextrins and eventually to maltose and then glucose. Glucose formed via this route is phosphorylated by the action of hexokinase and enters the glycolytic pathway.

Sucrose is a mobile form of carbohydrate in plants. It is synthesized in the cytoplasm of cells that contain chloroplasts (e.g., leaf cells) and exported to the plant vascular system where it is taken up by non-photosynthetic cells (e.g., root cells). Thus, sucrose is functionally equivalent to glucose, the mobile form of carbohydrate in those animals that possess a circulatory system (Section 12.5).

The pathway for sucrose synthesis is shown in Figure 15.28. Four molecules of triose phosphate produce one molecule of sucrose. The triose phosphates follow the gluconeogenesis pathway, condensing to form fructose 1,6-*bis*phosphate that is hydrolyzed to yield fructose 6-phosphate. Fructose 6-phosphate isomerizes to glucose 6-phosphate that is diverted from the gluconeogenesis pathway and converted to α-D-glucose 1-phosphate. Glucose 1-phosphate reacts with UTP to form UDP–glucose and this activated glucose molecule donates its glucosyl group to a molecule of fructose 6-phosphate,



▲ **Figure 15.28**
Biosynthesis of sucrose from glyceraldehyde 3-phosphate and dihydroxyacetone phosphate in the cytosol. Four molecules of triose phosphate (4 C_3) are converted to one molecule of sucrose (C_{12}).

BOX 15.5 GREGOR MENDEL'S WRINKLED PEAS

One of the genetic traits that Gregor Mendel studied was round (R) vs. wrinkled (r) peas. The wrinkled pea phenotype is caused by a defect in the gene for starch branching enzyme. Starch synthesis is partially blocked in the absence of this enzyme and the developing peas have a higher concentration of sucrose. This causes them to absorb more water than the normal peas and they swell to a larger size. When the seeds begin to dry out the mutant peas lose more water and their outer surface takes on a wrinkled appearance.

The mutation is caused by insertion of a transposon into the gene. It is a *recessive* loss-of-function mutation because a single copy of the normal wild-type allele in heterozygotes can produce enough starch branching enzyme to produce starch granules.



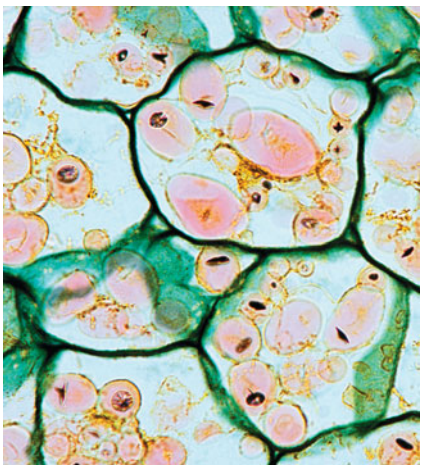
▲ Round and wrinkled peas in a pod.

to form sucrose 6-phosphate. The final step is the hydrolysis of sucrose 6-phosphate to form sucrose.

Inorganic phosphate (P_i) is produced in the sucrose synthesis pathway by the reactions catalyzed by fructose 1,6-*bis*phosphatase and sucrose phosphate phosphatase. Pyrophosphate (PP_i) is produced in the reaction catalyzed by UDP–glucose pyrophosphorylase. The pathway consumes one ATP equivalent (as UTP). Sucrose synthesis and glycogen synthesis require an activated glucose molecule in the form of UDP–glucose whereas starch biosynthesis uses ADP–glucose.

The first metabolically irreversible step in the sucrose biosynthesis pathway is the hydrolysis of fructose 1,6-*bis*phosphate to yield fructose 6-phosphate and P_i. The activity of fructose 1,6-*bis*phosphatase is inhibited by the allosteric modulator fructose 2,6-*bis*phosphate (Figure 12.9)—a molecule we encountered in our examinations of glycolysis and gluconeogenesis. In plants, the level of fructose 2,6-*bis*phosphate is regulated by several metabolites that reflect the suitability of conditions for sucrose synthesis.

Sucrose is taken up by non-photosynthetic cells where it is degraded by sucrase (invertase) to glucose and fructose that supply energy via glycolysis and the citric acid cycle (Section 11.6A). These hexoses can also be converted to starch in those tissues that store carbohydrate for future use. In root cells, for example, sucrose is converted to hexose monomers and these sugars are taken up by specialized organelles called **amyloplasts**. Amyloplasts are modified chloroplasts that lack the photosynthesis complexes but retain the enzymes for starch synthesis. In some plants, such as potatos, turnips, and carrots, the root cells can store huge reservoirs of starch.



▲ Amyloplasts in potato cells.



▲ Potatoes are an excellent source of starch. French fries are served in Québec with gravy and cheese curds. The dish is called poutine.

15.6 Additional Carbon Fixation Pathways

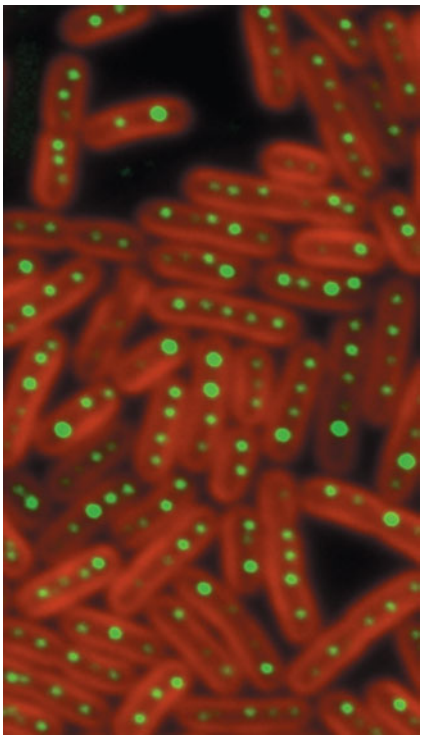
As mentioned earlier, one of the most important problems with carbon fixation is the inefficiency of Rubisco, especially the oxygenation reaction that greatly limits crop yields (Section 15.4C) . Different species have evolved a variety of ways of overcoming this problem.

A. Compartmentalization in Bacteria

Bacteria avoid the problems of photorespiration by confining Rubisco to specialized compartments called carboxysomes. Carboxysomes are surrounded by a protein coat that is impermeable to oxygen. Rubisco is localized to carboxysomes and so is the enzyme carbonic anhydrase that converts bicarbonate ($HCO_3^{\ominus}$) to CO_2 (see Section 2.10 and Figure 7.1). The advantage of compartmentalization is that Rubisco is supplied with an abundant source of CO_2 while protecting it against O_2, thus avoiding the inefficiencies of photorespiration.

B. The C_4 Pathway

Several plant species avoid wasteful photorespiration by means of secondary pathways for carbon fixation. The net effect of these secondary pathways is to increase the local



▲ **Carboxysomes.**
Cyanobacteria (*Synechococcus elongatus*) cells are stained with a fluorescent dye showing thylakoid membranes (red) and carboxysomes (green).

concentration of CO_2 relative to O_2 in those cells where Rubisco is active. One of these pathways is called the C_4 pathway because it involves four-carbon intermediates. C_4 plants tend to grow at high temperatures and high light intensities. They include such economically important species as maize (corn), sorghum, and sugarcane, and many of the most troublesome weeds. The avoidance of photorespiration by tropical plants is essential because the ratio of oxygenation to carboxylation by Rubisco increases with temperature.

The C_4 pathway concentrates CO_2 and delivers it to cells in the interior of the leaf where the Calvin cycle is active. The initial product of carbon fixation is a four-carbon acid (C_4) rather than a three-carbon acid as in the Calvin cycle. The C_4 pathway occurs in two different cell types within the leaf. First, CO_2 is hydrated to bicarbonate that reacts with the C_3 compound phosphoenolpyruvate to form a C_4 acid in mesophyll cells (near the leaf exterior). This reaction is catalyzed by an isozyme of phosphoenolpyruvate (PEP) carboxylase (Section 13.6). Next, the C_4 acid is transported to bundle sheath cells in the interior of the leaf where it is decarboxylated. Because they are not directly exposed to the atmosphere, the bundle sheath cells have a much lower O_2 concentration than mesophyll cells. The released CO_2 is fixed by the action of Rubisco and incorporated into the Calvin cycle. Phosphoenolpyruvate is regenerated from the remaining C_3 product. Figure 15.29 outlines the sequence of C_4 pathway reactions.

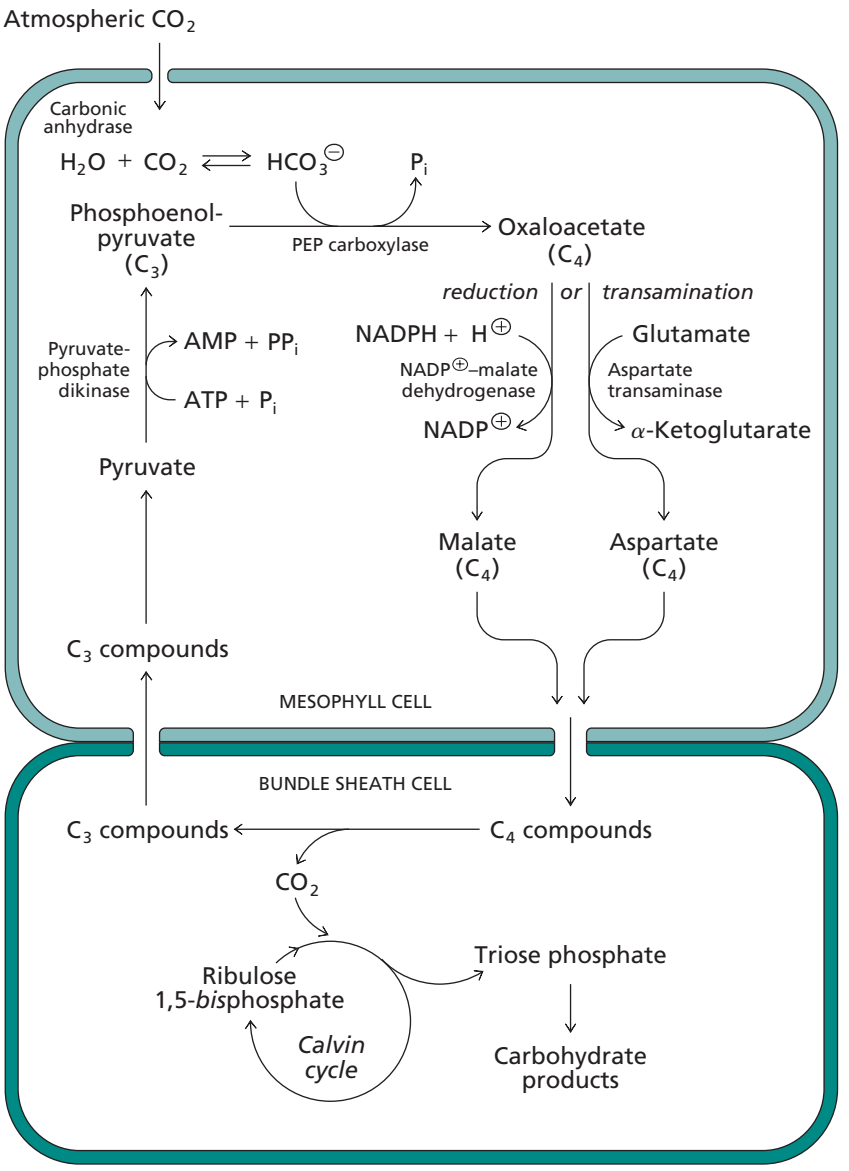


Figure 15.29 ▸
C_4 pathway. CO_2 is hydrated to bicarbonate (HCO_3^-) in the mesophyll cytosol. Bicarbonate reacts with phosphoenolpyruvate in a carboxylation reaction catalyzed by phosphoenolpyruvate (PEP) carboxylase, a cytosolic enzyme that has no oxygenase activity. Depending on the species, the oxaloacetate produced is either reduced or transaminated to form a four-carbon carboxylic acid or amino acid, which is transported to an adjacent bundle sheath cell and decarboxylated. The released CO_2 is fixed by the Rubisco reaction and enters the RPP cycle. The remaining three-carbon compound is converted back to the CO_2 acceptor, phosphoenolpyruvate.

The cell walls of internal bundle sheath cells are impermeable to gases. The decarboxylation of C_4 acids in these cells greatly increases the CO_2 concentration and creates a high ratio of CO_2 to O_2. The oxygenase activity of Rubisco is minimized because there is an insignificant amount of Rubisco in mesophyll cells and the ratio of CO_2 to O_2 is extremely high in bundle sheath cells. As a result, C_4 plants have essentially no photorespiration activity. Although there is an extra energy cost to form phosphoenolpyruvate for C_4 carbon assimilation, the absence of photorespiration gives C_4 plants a significant advantage over C_3 plants.



▲ **Field of Dreams.** These baseball players were probably studying the biochemistry of carbon fixation in the corn field.

C. Crassulacean Acid Metabolism (CAM)

Succulent plants, such as many species of cactus, grow primarily in arid environments where water loss can be a serious problem. A large amount of water can be lost from the leaf tissues during carbon fixation since the cells must be exposed to atmospheric CO_2 and water can evaporate from the surface. These plants minimize water loss during photosynthesis by assimilating carbon at night. The pathway is called Crassulacean acid metabolism because it was first discovered in the family Crassulaceae.

The surface of the leaf in terrestrial vascular plants is often covered with an impermeable waxy coating and CO_2 passes through structures called stomata to reach photosynthetic cells. **Stomata** are formed by two adjacent cells on the surface of the leaf. These guard cells define the entrance to a cavity lined with cells containing chloroplasts. The aperture between the guard cells changes in response to ion fluxes and the resulting osmotic uptake of water. The flux of ions across the guard cells is regulated by conditions that affect CO_2 fixation such as temperature and the availability of water. In the heat of the day, CAM plants keep their stomata closed to minimize water loss. At night, mesophyll cells take up CO_2 through open stomata. Water loss through the stomata is much lower at cooler nighttime temperatures than during the day. CO_2 is fixed by the PEP carboxylase reaction, and the oxaloacetate formed is reduced to malate (Figure 15.30).



▲ **Cactus is a CAM plant.**

Malate is stored in a large central vacuole in order to maintain a nearly neutral pH in the cytosol since the cellular concentration of this acid can reach 0.2 M by the end of the night. The vacuoles of CAM plants generally occupy more than 90% of the total volume of the cell. Malate is released from the vacuole and decarboxylated during the day when ATP and NADPH are formed by photosynthesis. Thus, the large pool of malate accumulated at night supplies CO_2 for carbon assimilation during the day. Leaf stomata are tightly closed when malate is decarboxylated so that neither water nor CO_2 can escape from the leaf and the level of cellular CO_2 can be much higher than the level of atmospheric CO_2. As in C_4 plants, the higher internal CO_2 concentration greatly reduces photorespiration.

In CAM plants the phosphoenolpyruvate required for malate formation is derived from starch via glycolysis. The phosphoenolpyruvate formed by malate decarboxylation (either directly by PEP carboxykinase or via malic enzyme and pyruvate phosphate dikinase) is converted to starch via gluconeogenesis and stored in the chloroplast.

CAM is analogous to C_4 metabolism in that the C_4 acid formed by the action of PEP carboxylase is subsequently decarboxylated to supply CO_2 to the Calvin cycle. In the C_4 pathway the carboxylation and decarboxylation phases of the cycle are spatially separated in distinct cell types whereas in CAM they are temporally separated in day and night cycles.

An important regulatory feature of the CAM pathway is the inhibition of PEP carboxylase by malate and low pH. PEP carboxylase is effectively inhibited during the day when the cytosolic concentration of malate is high and pH is low. This inhibition prevents futile cycling of CO_2 and malate by PEP carboxylase and avoids competition between PEP carboxylase and Rubisco for CO_2.

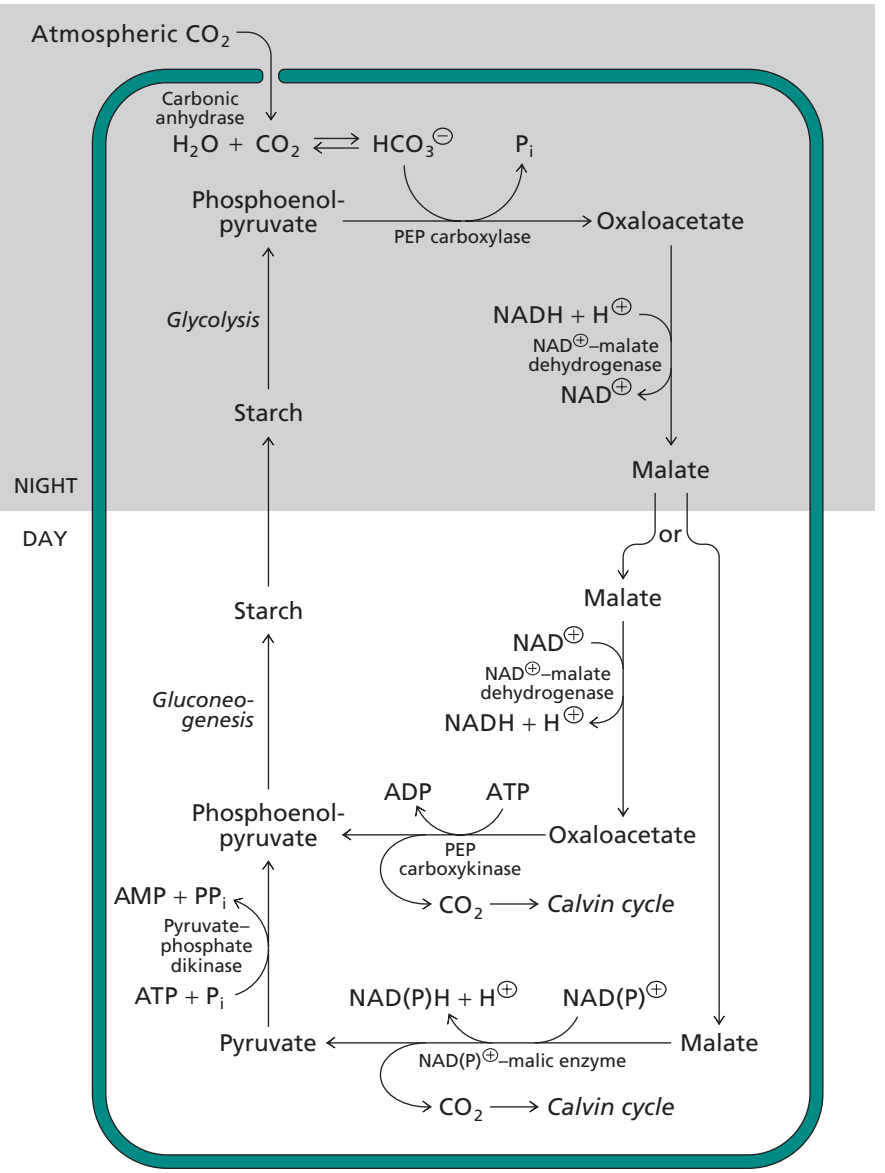


Figure 15.30 ▸
Crassulacean acid metabolism (CAM). At night, CO_2 is taken up, and PEP carboxylase and $NAD^{\oplus}$–malate dehydrogenase catalyze the formation of malate. The phosphoenolpyruvate required for malate synthesis is derived from starch. The next day, when NADPH and ATP are formed by the light reactions, the decarboxylation of malate increases the cellular concentration of CO_2 that can be fixed by the Calvin cycle. The decarboxylation of malate occurs by either of two pathways, depending on the species, and yields phosphoenolpyruvate, which is subsequently converted to starch through gluconeogenesis.

Summary

1. Chlorophyll is the major light-gathering pigment in photosynthesis. When chlorophyll molecules absorb a photon of light, an electron is promoted to a higher-energy molecular orbital. This electron can be transferred to an electron transfer chain giving rise to an electron-deficient chlorophyll molecule.
2. Accessory pigments transfer energy to the special pair of chlorophyll molecules by resonance energy transfer.
3. Photosystem II (PSII) complexes contain a type II reaction center. Electrons are transferred from the special pair of chlorophyll molecules to a short electron transfer chain consisting of a chlorophyll, a pheophytin, a bound quinone, and a mobile quinone.
4. In some bacteria QH_2 molecules from PSII bind to the cytochrome bc_1 complex. Electrons are transferred to cytochrome *c* and this process is coupled to the transfer of protons across the membrane via the Q cycle. Cytochrome *c* then binds to PSII and transfers electrons back to the electron-deficient special pair in a cyclic process of electron transfer. The resulting proton gradient drives ATP synthesis.
5. Photosystem I (PSI) complexes contain a type I reaction center. The electron transfer chain consists of two chlorophylls, a phylloquinone, three [Fe–S] clusters, and ferredoxin (or flavodoxin).
6. Reduced ferredoxin is the substrate for ferredoxin: $NADP^{\oplus}$ reductase (FNR), and NADPH is the product of photosystem I photosynthesis in a noncyclic electron transfer. In some cases, electrons are passed from ferredoxin to the cytochrome bc_1 complex and back to PSI via cytochrome *c* in a cyclic process of electron transfer.
7. Cyanobacteria, and chloroplasts, contain coupled photosystems consisting of PSI, PSII, and cytochrome *bf*—a photosynthetic version of cytochrome bc_1. When PSII absorbs a photon of light, electrons are transferred from PSII to cytochrome *bf* and plastocyanin. Plastocyanin resupplies electrons to PSI. When PSI absorbs a photon of light, excited electrons are used to synthesize NADPH.

In coupled photosystems, PSII is associated with an oxygen evolving complex (OEC) that catalyzes the oxidation of water to O_2 and supplies electrons to the PSII special pair.

8. The Z-scheme depicts electron flow during photosynthesis in terms of the change in reduction potentials of the various components of the electron transfer chains.
9. Photosynthesis complexes are concentrated in thylakoid membranes in cyanobacteria. Chloroplasts contain a complex internal membrane system of thylakoid membranes.
10. The Calvin cycle is responsible for fixing CO_2 into carbohydrates. The key enzyme is ribulose 1,5-*bis*phosphate carboxylase–oxygenase (Rubisco). Rubisco is an inefficient enzyme that catalyzes carboxylation of ribulose 1,5-*bis*phosphate. It also catalyzes an oxygenation reaction.
11. Sucrose and starch are the main products of photosynthetic carbohydrate synthesis in plants.
12. Additional carbon-fixation pathways in some plants serve to increase the concentration of CO_2 at the site of the Calvin cycle reactions.

Problems

1. In plants the transport of a single pair of electrons from P680 to NADPH is coupled to the accumulation of six protons in the lumen. This will result in production of 1.5 molecules of ATP (Section 14.11). Assuming that NADPH ≈ 2.5 ATP, this means that in photosynthesis transport of a pair of electrons through the complexes produces 1.5 + 2.5 = 4 ATP equivalents. Why is this process so much more efficient than respiratory electron transport?
2. The dragonfish is a deepwater species that flashes a red bioluminescent light to illuminate its prey. Although the visual pigments normally present in the retina of fish are not sensitive enough to pick up the red light, the dragonfish retina contains other pigments, derived from chlorophyll, that absorb at 667 nm. Suggest how these chlorophyll pigments might act as a photosensitizer to aid the dragonfish to detect prey using its own red light beacon, which other fish cannot see.
3. (a) Ribulose 1,5-*bis*phosphate carboxylase–oxygenase (Rubisco) has been called the "enzyme that feeds the world." Explain the basis for this statement.
   (b) Rubisco has also been accused of being the world's most incompetent enzyme and the most inefficient enzyme in primary metabolism. Explain the basis for this statement.
4. You frequently see photosynthesis plus the Calvin cycle described as

$$6\,CO_2 + 6\,H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6\,O_2$$

   Write a similar equation for the reactions in purple bacteria and in green sulfur bacteria.
5. (a) Some photosynthesis bacteria use H_2S as a hydrogen donor and produce elemental sulfur, whereas others use ethanol and produce acetaldehyde. Write the net reactions for photosynthesis for these bacteria.
   (b) Why is no oxygen produced by these bacteria?
   (c) Write a general equation for the photosynthetic fixation of CO_2 to carbohydrate using H_2A as the hydrogen donor.
6. Can a suspension of chloroplasts in the dark synthesize glucose from CO_2 and H_2O? If not, what must be added for glucose synthesis to occur? Assume that all the components of the Calvin cycle are present.
7. (a) How many photons are absorbed for every O_2 molecule produced in photosynthesis?
   (b) How many photons must be absorbed to generate enough NADPH reducing power for the synthesis of one molecule of a triose phosphate?
8. The herbicide 3-(3, 4-dichlorophenyl)-1,1-dimethylurea (DCMU) blocks photosynthetic electron transport from PSII to the cytochrome *bf* complex.
   (a) When DCMU is added to isolated chloroplasts, will both O_2 evolution and photophosphorylation cease?
   (b) If an external electron acceptor that reoxidizes P680* is added, how will this affect O_2 production and photophosphorylation?
9. (a) The luminal pH of chloroplasts suspended in a solution of pH 4.0 reaches pH 4.0 within a few minutes. Explain why there is a burst of ATP synthesis when the pH of the external solution is quickly raised to 8.0 and ADP and P_i are added.
   (b) If ample ADP and P_i are present, why does ATP synthesis cease after a few seconds?
10. Cyclic electron transport may occur simultaneously with noncyclic electron transport under certain conditions in chloroplasts. Is any ATP, O_2, or NADPH produced by cyclic electron transport?
11. A plant has been genetically engineered to contain a smaller percentage than normal of unsaturated lipids in the thylakoid membranes of the chloroplasts. This genetically changed plant has an improved tolerance to higher temperatures and also shows improved rates of photosynthesis and growth at 40°C. What major components of the photosynthesis system might be affected by changing the lipid composition of the thylakoid membranes?
12. A compound was added to isolated spinach chloroplasts and the effect on photosynthetic photophosphorylation, proton uptake, and noncyclic electron transport determined. Addition of the compound resulted in an inhibition of photosynthetic photophosphorylation (ATP synthesis), inhibition of proton uptake, and an enhancement in noncylic electron transport. Suggest a mechanism for the compound.
13. How many molecules of ATP (or ATP equivalents) and NADPH are required to synthesize (a) one molecule of glucose via photosynthetic CO_2 fixation in plants and (b) one glucose residue incorporated into starch?
14. After one complete turn of the Calvin cycle, where will the labeled carbon atoms from $^{14}CO_2$ appear in (a) glyceraldehyde 3-phosphate (b) fructose 6-phosphate, and (c) erythrose 4-phosphate?
15. (a) How many additional ATP equivalents are required to synthesize glucose from CO_2 in C_4 plants than are required in C_3 plants?
   (b) Explain why C_4 plants fix CO_2 much more efficiently than C_3 plants despite the extra ATP needed.
16. Explain how the following changes in metabolic conditions alter the Calvin cycle: (a) an increase in stromal pH, and (b) a decrease in stromal concentration of Mg^{2+}.

Selected Readings

Pigments

Armstead, I., Donnison, I., Aubry, S., Harper, J., Hörtensteiner, S., James, C., Mani, J., Moffet, M., Ougham, H., Roberts, L., Thomas, A., Weeden, N., Thomas, H., and King, I. (2007). Cross-species identification of Mendel's I locus. *Science* 315:73.

Sato Y., Morita R., Nishimura M., Yamaguchi H., and Kusaba M. (2007). Mendel's green cotyledon gene encodes a positive regulator of the chlorophyll-degrading pathway. *Proc. Natl. Acad. Sci. (USA)* 104:14169–14174.

Photosynthetic Electron Transport

Allen, J. F. (2004). Cytochrome b_6f: structure for signalling and vectorial metabolism. *Trends in Plant Sci.* 9:130–137.

Allen, J. P., and Williams, J. C. (2010). The evolutionary pathway from anoxygenic to oxygenic photosynthesis examined by comparison of the properties of photosystem II and bacterial reaction centers. *Photosynth. Res.* Published online May 7, 2010: Doi 10.1007/s11120-010-9552-x

Amunts, A., Toporik, H., Borovikova, A. B., and Nelson, N. (2010). Structure determination and improved model of plant photosystem I. *J. Biol. Chem*. 285:3478–3486.

Barber, J., Nield, J., Morris, E. P., and Hankamer, B. (1999). Subunit positioning in photosystem II revisited. *Trends Biochem. Sci.* 24:43–45.

Cramer, W. A., Zhang, H., Yan, J., Kurisu, G., and Smith, J. L. (2004). Evolution of photosynthesis: time-independent structure of the cytochrome b_6f complex. *Biochem*. 43:5921–5929.

Cramer, W. A., Zhang, H., Yan, J., Kurisu, G., and Smith, J. L. (2006). Transmembrane traffic in the cytochrome b_6f complex. *Annu. Rev. Biochem.* 75:769–790.

Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., and Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. *Science* 303:1831–1838.

Golbeck, J. H. (1992). Structure and function of photosystem I. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 43:293–324.

Kühlbrandt, W., Wang, D. N., and Fujiyoshi, Y. (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature* 367: 614–621.

Leslie, M. (2009). On the origin of photosynthesis. *Science* 323:1286–1287.

Müller, M. G., Slavov, C., Luthra, R., Redding, K. E., and Holzwarth, A. R. (2010). Independent initiation of primary electron transfer in the two branches of the photosystem I reaction center. *Proc. Natl. Acad. Sci. (USA)* 107:4123–4128.

Nugent, J. H. A. (1996). Oxygenic photosynthesis. Electron transfer in photosystem I and photosystem II. *Eur. J. Biochem*. 237:519–531.

Rhee, K.-H., Morris, E. P., Barber, J., and Kühlbrandt, W. (1998). Three-dimensional structure of the plant photosystem II reaction centre at 8 Å resolution. *Nature* 396:283–286.

Staehlin, L. A., and Arntzen, C. J., eds. (1986). *Photosynthesis III: Photosynthetic Membranes and Light Harvesting Systems*. Vol. 19 of *Encyclopedia of Plant Physiology* (New York: Springer-Verlag).

Photophosphorylation

Bennett, J. (1991). Phosphorylation in green plant chloroplasts. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42:281–311.

Photosynthetic Carbon Metabolism

Andrews, T. J. and Whitney, S. M. (2003). Manipulating ribulose *bis*phosphate carboxylase/oxygenase in the chloroplasts of higher plants. *Arch. Biochem. Biophys*. 414:159–169.

Bassham, J. A., and Calvin, M. (1957). *The Path of Carbon in Photosynthesis* (Englewood Cliffs, NJ: Prentice Hall).

Edwards, G. E., and Walker, D. (1983). C_3 C_4: *Mechanisms and Cellular and Environmental Regulation of Photosynthesis* (Berkeley: University of California Press).

Hartman, F. C., and Harpel, M. R. (1994). Structure, function, regulation, and assembly of D-ribulose-1,5-*bis*phosphate carboxylase/oxygenase. *Annu. Rev. Biochem.* 63:197–234.

Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010). Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science* 327:1258–1261.

Schnarrenberger, C., and Martin, W. (1997). The Calvin cycle—a historical perspective. *Photosynthetica* 33:331–345. An outline of the advances made since the 1950s.





Lipid Metabolism

The synthesis of lipids is an essential part of cellular metabolism since lipids are crucial components of cell membranes. In this chapter we describe the pathways for synthesis of the major lipids that were described in Chapter 9. The most important of these pathways is fatty acid synthesis since fatty acids are required in triacylglycerols. Other important biosynthesis pathways include cholesterol synthesis, eicosanoid synthesis, and the synthesis of sphingolipids.

Lipids can also be degraded as a normal part of cellular metabolism. The most important catabolic pathway is that of fatty acid oxidation (β-oxidation). In this pathway, long-chain fatty acids are broken down to acetyl CoA. The opposing pathways of fatty acid biosynthesis and fatty acid oxidation provide another example of how cells handle energy production and utilization in a manner that's compatible with the fundamentals of thermodynamics.

The catabolic pathways of lipid metabolism are part of basic fuel metabolism in animals. Triacylglycerols and glycogen are the two major forms of stored energy. Glycogen can supply ATP for muscle contraction for only a fraction of an hour. Sustained intense work, such as the migration of birds or the effort of marathon runners, is fueled by the metabolism of triacylglycerols. Triacylglycerols are anhydrous and their fatty acids are more reduced than amino acids or monosaccharides—this makes them very efficient at storing energy for use later on (Section 9.3). Triacylglycerols are oxidized when the energy demand increases. In most cases, fat is only used when other energy sources, such as glucose, are unavailable.

We will begin by examining the fundamental pathways of lipid metabolism—the ones that are present in all living species. Where necessary, we'll point out the differences between the bacterial and the eukaryotic pathways. These differences are minor. We then go on to describe the absorption and utilization of dietary lipids in mammals, including the hormonal regulation of lipid metabolism.

Derangements of this complicated mechanism of formation and metabolism of lipids are in many cases responsible for the genesis of some of our most important diseases, especially in the cardiovascular field. A detailed knowledge of the mechanisms of lipid metabolism is necessary to deal with these medical problems in a rational manner.

—S. Bergström, presentation speech on awarding the 1964 Nobel Prize in Physiology or Medicine to Konrad Bloch and Feodor Lynen

16.1 Fatty Acid Synthesis

Fatty acids are synthesized by the repetitive addition of two-carbon units to the growing end of a hydrocarbon chain. The growing chain is covalently attached to acyl carrier

Top: Whereas the polar bear lives off its stored fat for much of the year, the bird uses its fat stores for long flights.

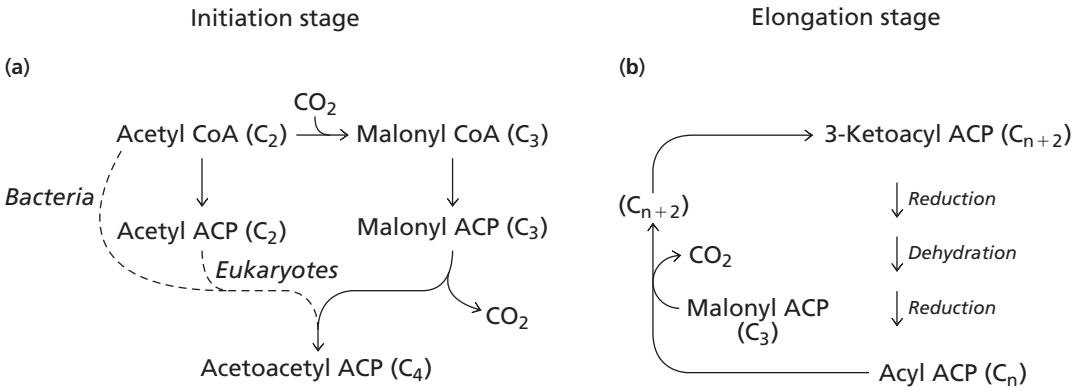
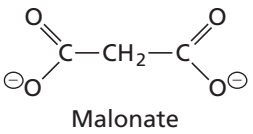


Figure 16.1 ►
Outline of fatty acid synthesis.



protein (ACP), a protein coenzyme (Section 7.6). The linkage is a thioester as in acetyl CoA. An overview of fatty acid synthesis is shown in Figure 16.1.

The first steps in the fatty acid synthesis pathway are the production of acetyl ACP and malonyl ACP from acetyl CoA. (Malonic acid, or malonate, is the name of the standard C_3 dicarboxylic acid.) The initiation step involves a condensation of acetyl and malonyl groups to give a four-carbon precursor and CO_2. This precursor serves as the primer for fatty acid synthesis. In the elongation stage, the acyl group attached to ACP (acyl ACP) is extended by two-carbon units donated by malonyl ACP. The product of the initial condensation (3-ketoacyl ACP) is modified by two reduction reactions and a dehydration reaction to produce a longer acyl ACP. Acyl ACP then serves as the substrate for additional condensation reactions.

Fatty acid synthesis takes place in the cytosol of all species. In adult mammals it occurs largely in liver cells and adipocytes. Some fatty acid synthesis takes place in specialized cells such as mammary glands during lactation.

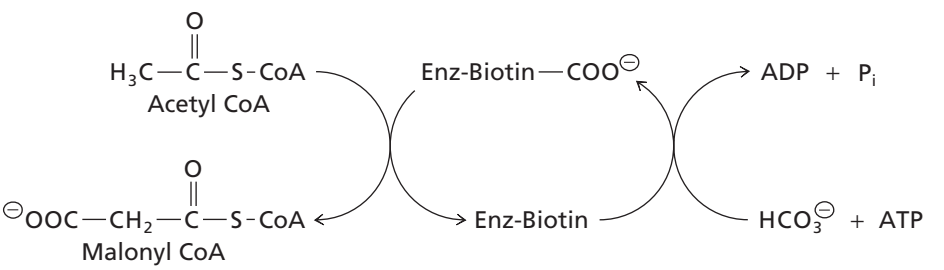
The regulation of fatty acid metabolism is described in Section 16.9.

A. Synthesis of Malonyl ACP and Acetyl ACP

Malonyl ACP is the substrate for fatty acid biosynthesis. It is synthesized in two steps, the first of which is the carboxylation of acetyl CoA in the cytosol to form malonyl CoA (Figure 16.2). The carboxylation reaction is catalyzed by the biotin-dependent enzyme acetyl-CoA carboxylase using a mechanism similar to the reaction catalyzed by pyruvate carboxylase (Figure 7.20). The ATP-dependent activation of $HCO_3^{\ominus}$ forms carboxybiotin. This reaction is followed by the transfer of activated CO_2 to acetyl CoA, forming malonyl CoA. These reactions are catalyzed in eukaryotes by a bifunctional enzyme and the biotin moiety is on a flexible arm that moves between the two active sites. The bacterial version of acetyl-CoA carboxylase is a multisubunit enzyme complex containing biotin carboxylase, biotin carboxylase carrier protein, and a heterodimeric transcarboxylase. In all species, acetyl-CoA carboxylase is the key regulatory enzyme of fatty acid synthesis and the carboxylation reaction is metabolically irreversible.

The second step in the synthesis of malonyl ACP is the transfer of the malonyl moiety from coenzyme A to ACP. This reaction is catalyzed by malonyl CoA:ACP transacylase (Figure 16.3). A similar enzyme called acetyl CoA:ACP transacylase converts acetyl CoA to the acetyl ACP. In most species these are separate enzymes with specificity for malonyl CoA or acetyl CoA but in mammals the two activities are combined in a bifunctional enzyme, malonyl–acetyl transferase (MAT) that's part of a larger complex (see below).

Figure 16.2 ►
Carboxylation of acetyl CoA to malonyl CoA, catalyzed by acetyl-CoA carboxylase.



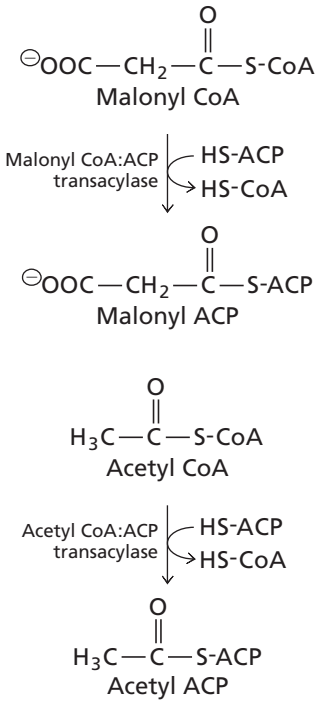
B. The Initiation Reaction of Fatty Acid Synthesis

The synthesis of long-chain fatty acids begins with the formation of a four-carbon unit attached to ACP. This molecule, called acetoacetyl ACP, is formed by condensation of a two-carbon substrate (acetyl CoA or acetyl ACP) and a three-carbon substrate (malonyl ACP) with the loss of CO_2. The reaction is catalyzed by 3-ketoacyl ACP synthase (KAS).

There are several versions of KAS in bacterial cells. One form of the enzyme (KAS III) is used in the initiation reaction and other versions (KAS I, KAS II) are used in subsequent elongation reactions. Bacterial KAS III uses acetyl CoA for the initial condensation reaction with malonyl ACP (Figure 16.4).

A two-carbon unit from acetyl CoA is transferred to the enzyme where it is covalently bound via a thioester linkage. The enzyme then catalyzes the transfer of this two-carbon unit to the end of malonyl ACP creating a four-carbon intermediate and releasing CO_2. Eukaryotic versions of 3-ketoacyl ACP synthase carry out the same reaction except that they use acetyl ACP as the initial substrate instead of acetyl CoA.

Recall that synthesis of malonyl CoA involves ATP-dependent carboxylation of acetyl CoA (Figure 16.2). This strategy of first carboxylating and then decarboxylating a compound results in a favorable free energy change for the process at the expense of ATP consumed in the carboxylation step. A similar strategy is seen in mammalian gluconeogenesis where pyruvate (C_3) is first carboxylated to form oxaloacetate (C_4) and then oxaloacetate is decarboxylated to form the C_3 molecule phosphoenolpyruvate (Section 12.1).

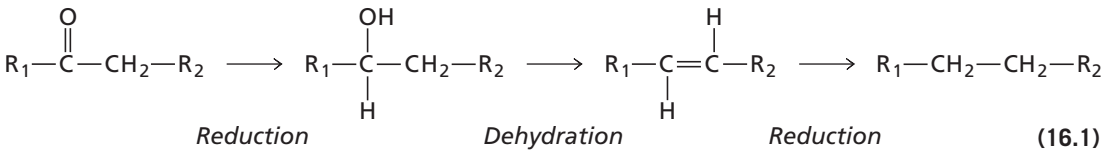


▲ **Figure 16.3**
Synthesis of malonyl ACP from malonyl CoA and acetyl ACP from acetyl CoA.

C. The Elongation Reactions of Fatty Acid Synthesis

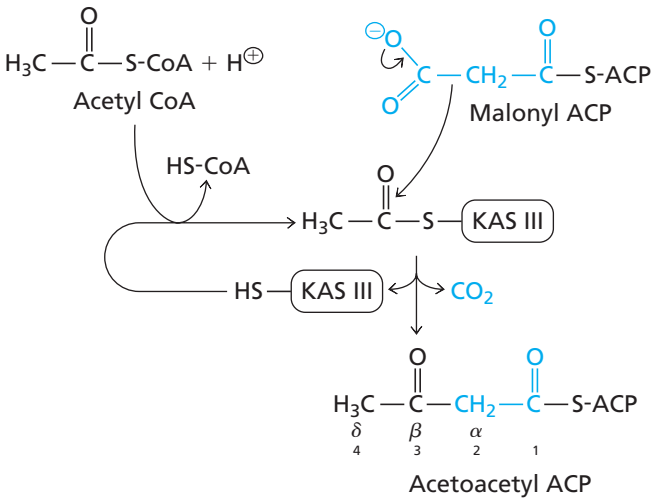
Acetoacetyl ACP contains the smallest version of a 3-ketoacyl moiety. The "3-keto-" in the name of this molecule refers to the presence of a keto group at the C-3 position. In the older terminology this carbon atom was the β-carbon and the product was called a β-ketoacyl moiety. The condensation enzyme is also called β-ketoacyl ACP synthase.

In order to prepare for subsequent condensation reactions, this oxidized 3-ketoacyl moiety has to be reduced by the transfer of electrons (and protons) to the C-3 position. Three separate reactions are required,



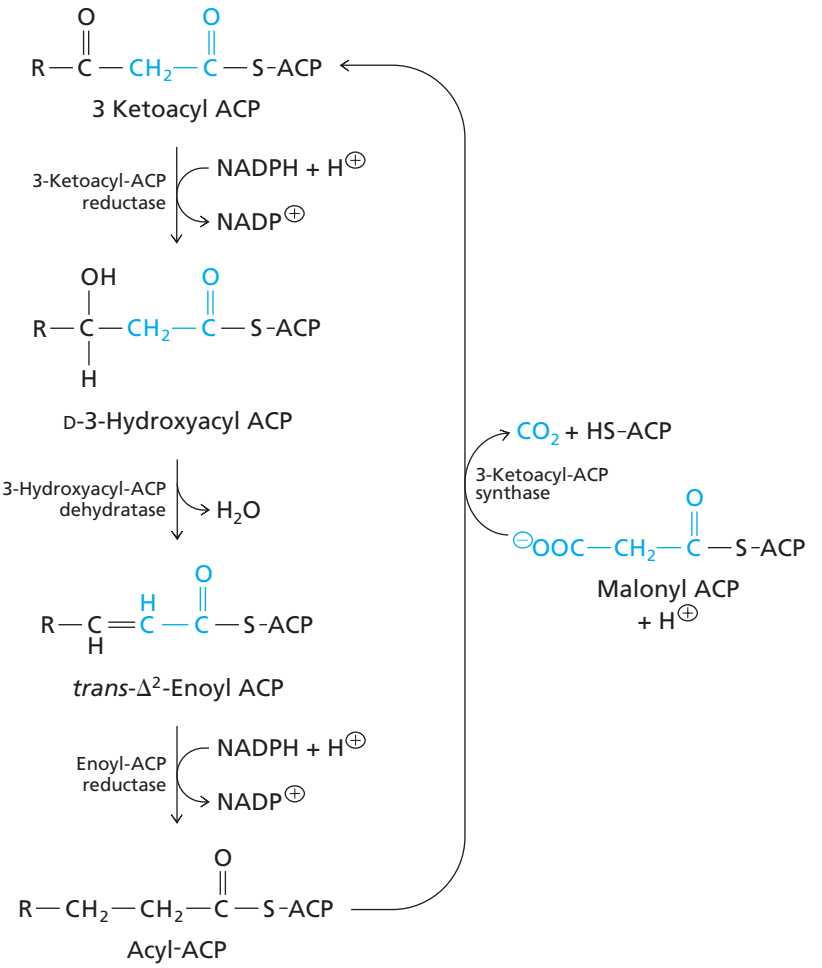
(16.1)

The ketone is reduced to the corresponding alcohol in the first reduction. The second step is the removal of water by a dehydratase producing a C═C double bond. Finally, a



◀ **Figure 16.4**
Synthesis of acetoacetyl ACP in bacteria.

Figure 16.5 ▶
The elongation stage of fatty acid synthesis. R represents —CH_3 in acetoacetyl ACP or [—CH_2—CH_2]$_n$—CH_3 in other 3-ketoacyl ACP molecules.



KEY CONCEPT

Malonyl ACP, formed from acetyl CoA, is the precursor for all fatty acid synthesis.

second reduction adds hydrogens to create the reduced acyl group. This is a common oxidation–reduction strategy in biochemical pathways. We have seen an example of the reverse reactions in the citric acid cycle where succinate is oxidized to oxaloacetate (Figure 13.5).

The specific reactions of the elongation cycle are shown in Figure 16.5. The first reduction is catalyzed by 3-ketoacyl ACP reductase (KR). The full name of the dehydratase enzyme is 3-hydroxyacyl ACP dehydratase (DH). The second reduction step is catalyzed by enoyl ACP reductase (ER). Note that during synthesis the D isomer of the β-hydroxy intermediate is formed in an NADPH-dependent reaction. We will see in Section 16.7 that the L isomer is formed during the degradation of fatty acids.

The final product of the reduction, dehydration, and reduction steps is an acyl ACP that is two carbons longer. This acyl ACP becomes the substrate for the elongation forms of 3-ketoacyl ACP synthase (KAS I and KAS II). All species use malonyl ACP as the carbon donor in the condensation reaction. The elongation reactions are repeated many times resulting in longer and longer fatty acid chains.

The end products of saturated fatty acid synthesis are 16- and 18-carbon fatty acids. Larger chain lengths cannot be accommodated in the binding site of the condensing enzyme. The completed fatty acid is released from ACP by the action of a thioesterase (TE) that catalyzes a cleavage reaction regenerating HS–ACP. For example, palmitoyl ACP is a substrate for a thioesterase that catalyzes formation of palmitate and HS–ACP.

$$\text{Palmitoyl-ACP} \xrightarrow[\text{Thioesterase}]{H_2O} \text{Palmitate } (C_{16}) + \text{HS-ACP} + H^{\oplus} \qquad \textbf{(16.2)}$$

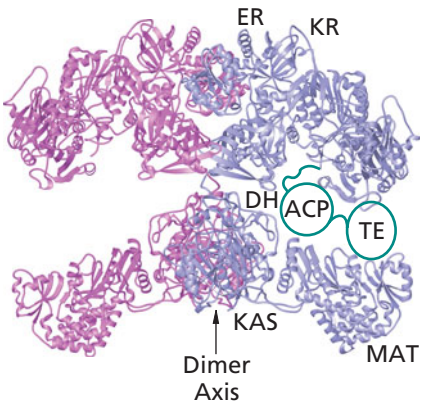
The overall stoichiometry of palmitate synthesis from acetyl CoA and malonyl CoA is

$$\text{Acetyl CoA} + 7\ \text{Malonyl CoA} + 14\ \text{NADPH} + 20\ \text{H}^{\oplus} \longrightarrow \text{Palmitate} + 7\ \text{CO}_2 + 14\ \text{NADP}^{\oplus} + 8\ \text{HS–CoA} + 6\ \text{H}_2\text{O} \quad (16.3)$$

In bacteria, each reaction in fatty acid synthesis is catalyzed by a discrete monofunctional enzyme. This type of pathway is known as a type II fatty acid synthesis system (FAS II). In fungi and animals, the various enzymatic activities are localized to individual domains in a large multifunctional enzyme and the complex is described as a type I fatty acid synthesis system (FAS I).

The large mammalian polypeptide is about 2500 amino acid residues in length (Mr = 270 kDa). Fatty acid synthase is a dimer where the two monomers are tightly bound, creating an enzyme with two sites where the fatty acids are synthesized on either side of the dimer axis (Figure 16.6). The bottom part of the enzyme in Figure 16.6 contains the condensing activities of malonyl/acetyl transferase (MAT) and 3-ketoacyl ACP synthase (KAS) that are responsible for adding a new two-carbon unit to the growing chain. These enzymes attach the fatty acid to a bound ACP phosphopantetheine prosthetic group (ACP) that is positioned on a flexible loop. The ACP-bound fatty acid visits the active sites of the modifying activities: 3-ketoacyl ACP reductase (KR), 3-hydroxyacyl ACP dehydratase (DH), and enoyl ACP reductase (ER). The fatty acid chain is eventually released by a thioesterase (TE) activity.

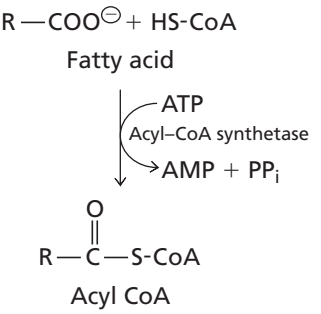
The structures of the ACP domain and the TE domain are not resolved in the crystal structure because they are tethered to the main part of the enzyme by a short stretch of residues that are intrinsically disordered (Section 4.7D). These flexible domains must be free to move during the reaction.



▲ **Figure 16.6**
Mammalian fatty acid synthase. The structure of the pig (*Sus scrofa*) enzyme is shown. It is a large dimer consisting of the following enzyme activities: malonyl/acetyl transferase (MAT), 3-ketoacyl ACP synthase (KAS), 3-ketoacyl ACP reductase (KR), 3-hydroxyacyl ACP dehydratase (DH), enoyl ACP reductase (ER), and thioesterase (TE). The fatty acid chain is attached to a bound ACP cofactor (ACP). The structures of the ACP and TE domains are not resolved because they are bound to a flexible tether. [PDB 2VZ9]

D. Activation of Fatty Acids

The thioesterase reaction (Reaction 16.2) results in release of free fatty acids but subsequent modifications of these fatty acids require an activation step where they are converted to thioesters of coenzyme A in an ATP-dependent reaction catalyzed by acyl-CoA synthetase (Figure 16.7). The pyrophosphate released in this reaction is hydrolyzed to two molecules of phosphate by the action of pyrophosphatase. As a result, two phosphoanhydride bonds, or two ATP equivalents, are consumed to form the CoA thioesters of fatty acids. Bacteria generally have a single acyl-CoA synthetase but in mammals there are at least four different acyl-CoA synthetase isoforms. Each of the distinct enzymes is specific for a particular fatty acid chain length: short ($<C_6$), medium (C_6 to C_{12}), long ($>C_{12}$), or very long ($>C_{16}$). The mechanism of the activation reaction is the same as that for the synthesis of acetyl CoA from acetate and CoA (Figure 10.13). Activation of fatty acids is required for their incorporation into membrane lipids (Section 16.2).



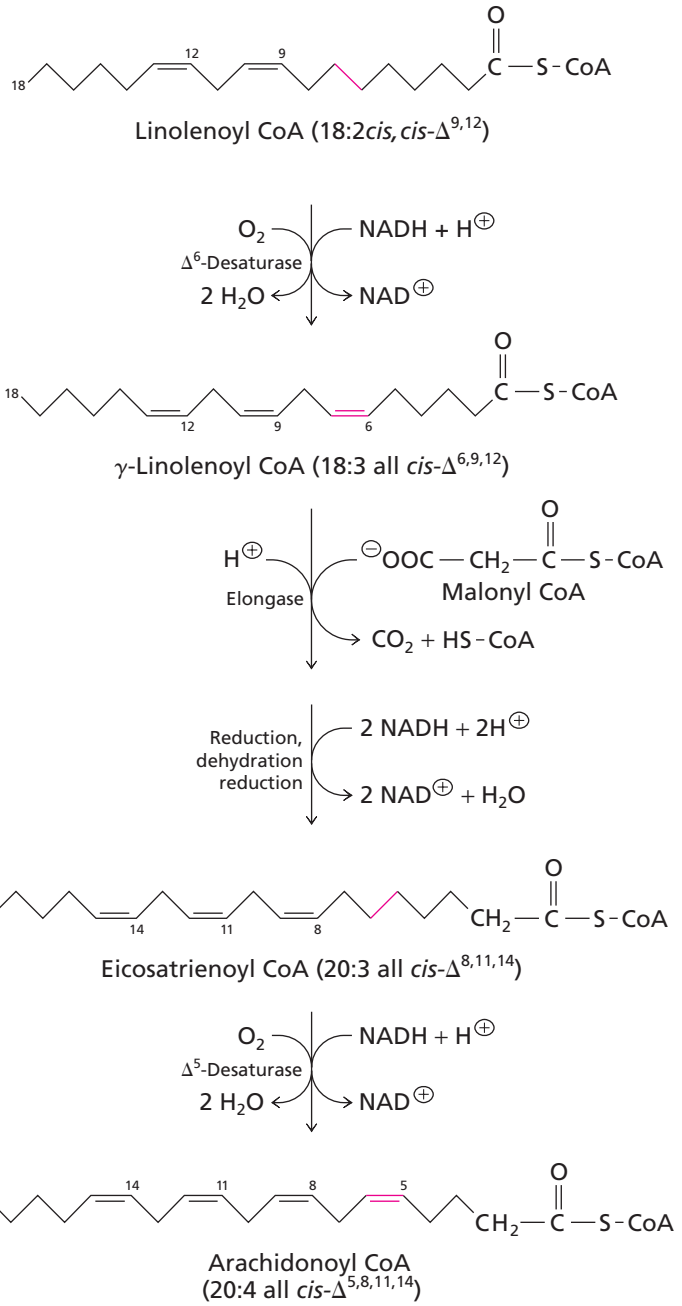
▲ **Figure 16.7**
Activation of fatty acids.

E. Fatty Acid Extension and Desaturation

The fatty acid synthase pathway cannot make fatty acids that are longer than 16 or 18 carbons (C_{16} or C_{18}). Longer fatty acids are made by extending palmitoyl CoA or stearoyl CoA in separate extension reactions. The enzymes that catalyze such extensions are known as elongases and they use malonyl CoA (not malonyl ACP) as the source of the two-carbon extension unit. An example of an elongase reaction is shown below in step 2 of Figure 16.8. Long chain fatty acids such as C_{20} and C_{22} fatty acids are common but C_{24} and C_{26} fatty acids are rare.

Unsaturated fatty acids are synthesized in both bacteria and eukaryotes but the pathways are quite different. In type II fatty acid synthesis systems (bacteria) a double bond is added to the growing chain when it reaches a length of ten carbon atoms. The reaction is catalyzed by specific enzymes that recognize the C_{10} intermediate. For example, 3-hydroxydecanoyl–ACP dehydratase specifically introduces a double bond at the 2 position just as in the normal dehydratase reaction during fatty acid synthesis (Figure 16.5). However, the specific C_{10} dehydratase creates a *cis*-2-decanoyl ACP and not the *trans* configuration that serves as a substrate for enoyl ACP reductase.

Figure 16.8 ►
Elongation and desaturation reactions in the conversion of linolenoyl CoA to arachidonoyl CoA.



The nomenclature of unsaturated fatty acids is described in Section 9.2.

Subsequent elongation of this unsaturated fatty acid proceeds by the normal fatty acid synthase pathway except that a specific 3-ketoacyl–ACP synthase enzyme recognizes the unsaturated fatty acid in the condensation reaction. The final products will be 16:1 Δ^8 and 18:1 Δ^{10} unsaturated fatty acids. These products can be further modified to create polyunsaturated fatty acids (PUFAs) in bacteria. The chains can be extended by elongase enzymes and additional double bonds are introduced by a class of enzymes called desaturases. Bacteria contain a huge variety of PUFAs that serve to increase the fluidity of membranes when species encounter low temperatures (Section 9.9). For example, many species of marine bacteria synthesize 20:5 and 22:6 PUFAs. Up to 25% of the membrane fatty acids are large polyunsaturated fatty acids in these species.

The introduction of a double bond during synthesis of fatty acids is not possible in eukaryotes since they employ a type I fatty acid synthase. This fatty acid synthase contains a single 3-ketoacyl–ACP synthase (KAS) activity that is part of a large multifunctional protein. The eukaryotic KAS active site does not recognize unsaturated fatty acid intermediates and could not extend them if they were created at the C_{10} step as in bacteria.

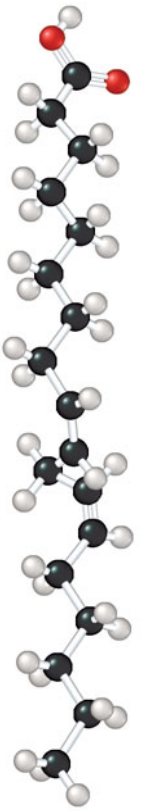
Consequently, eukaryotes synthesize unsaturated fatty acids entirely by using desaturases that act on the completed fatty acid derivatives palmitoyl CoA and stearoyl CoA.

Most eukaryotic cells contain various desaturases that catalyze the formation of double bonds as far as 15 carbons away from the carboxyl end of a fatty acid. For example, palmitoyl CoA is oxidized to its 16:1 Δ^9 analog that can be hydrolyzed to form the common fatty acid palmitoleate. Polyunsaturated fatty acids are synthesized by the sequential action of different, highly specific desaturases. In most cases, the double bonds are spaced at 3-carbon intervals as in synthesis of α-linolenate in plants.

$$\text{18:0 (stearoyl CoA)} \longrightarrow \text{18:1}\Delta^9 \longrightarrow \text{18:2}\Delta^{9,12}\text{(linolenoyl CoA)} \longrightarrow \text{18:3}\Delta^{9,12,15}(\alpha\text{-linolenoyl CoA}) \quad (16.4)$$

Mammalian cells do not contain a desaturase that acts beyond the C-9 position and they are not able to synthesize linoleate or α-linolenate. However, PUFAs with double bonds at the 12 position are absolutely essential for survival since they are precursors for synthesis of important eicosanoids such as prostaglandins. Because they lack a Δ^{12} desaturase, mammals must obtain linoleate from the diet. This is an essential fatty acid in the human diet. Deficiencies of α-lineolate are rare since most food contains adequate quantities. Plants, for example, are rich sources of PUFAs. Nevertheless, the composition of many "vitamin" supplements will include linoleic acid.

Mammals can convert dietary linoleate (activated to linolenoyl CoA) to arachidonoyl CoA (20:4) by a series of desaturation and elongation reactions (Figure 16.8). (Arachidonate derived from phospholipids is a precursor of eicosanoids, Section 16.3.) This pathway illustrates typical examples of elongase and desaturase activity in the synthesis of complex PUFAs. The intermediate γ-linolenoyl CoA (18:3) in the arachidonate pathway can undergo elongation and desaturation to produce C_{20} and C_{22} polyunsaturated fatty acids. Note that the double bonds of polyunsaturated fatty acids are not conjugated but are interrupted by a methylene group. Thus, a Δ^9 double bond, for example, directs insertion of the next double bond to the Δ^6 position or the Δ^{12} position.



▲ **Linoleate.** Linoleate is an essential component of the human diet.

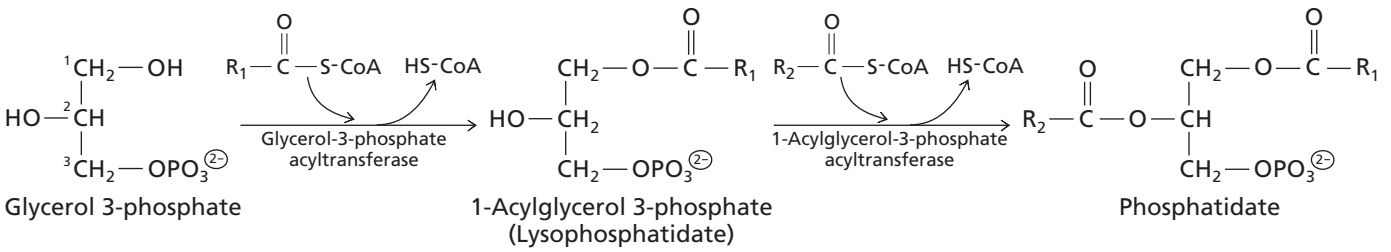
In addition to the essential fatty acids, mammalian diets must supply a number of essential vitamins (Chapter 7) and essential amino acids (Chapter 17).

16.2 Synthesis of Triacylglycerols and Glycerophospholipids

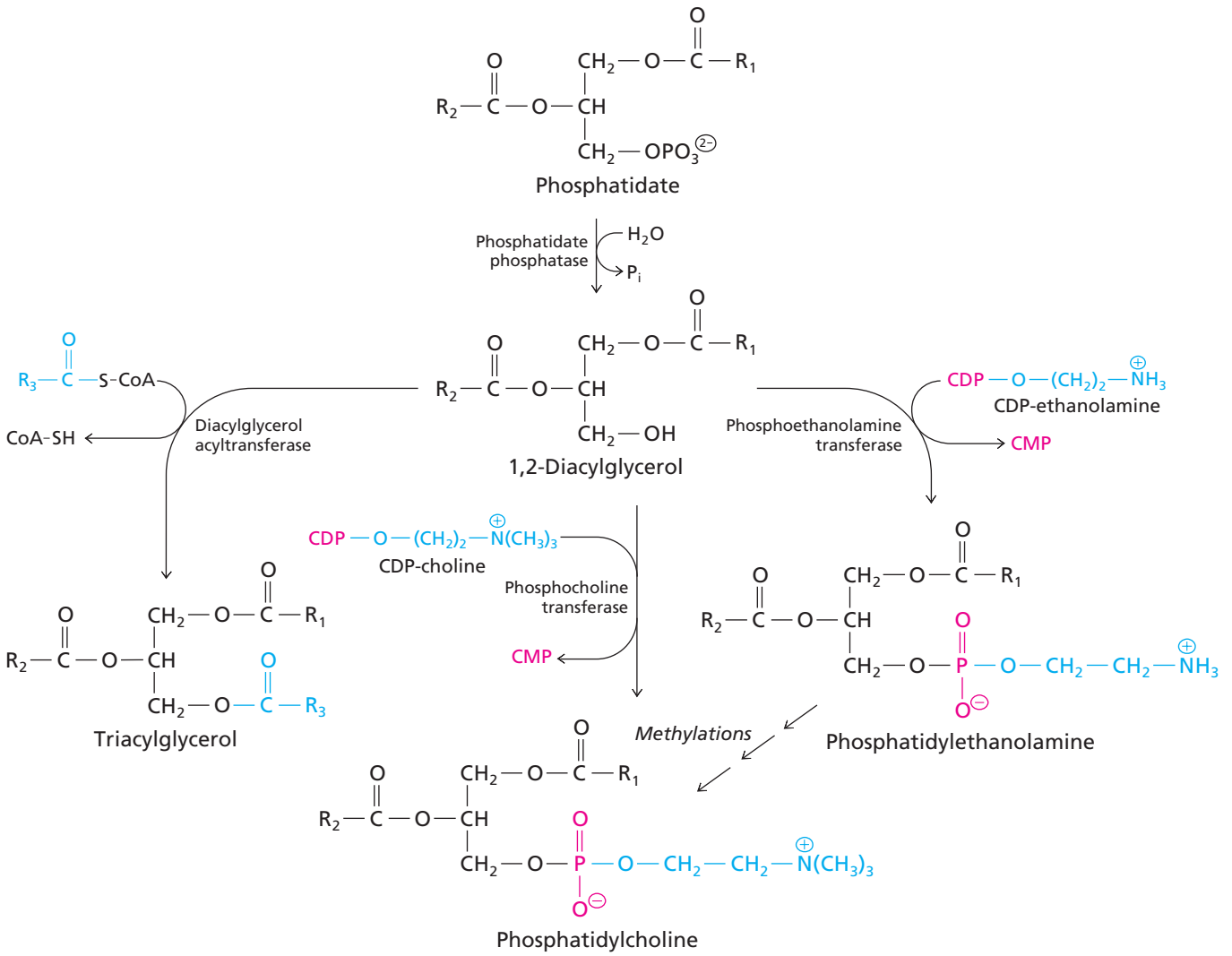
Most fatty acids are found in esterified forms as triacylglycerols or glycerophospholipids (Sections 9.3 and 9.4). Phosphatidate is an intermediate in the synthesis of triacylglycerols and glycerophospholipids. It is formed by transferring the acyl groups from fatty acid CoA molecules to the C-1 and C-2 positions of glycerol 3-phosphate (Figure 16.9). Glycerol 3-phosphate is synthesized from dihydroxyacetone phosphate in a reduction reaction catalyzed by glycerol 3-phosphate dehydrogenase. We encountered this enzyme when we discussed NADH shuttle mechanisms in Chapter 14 (Section 14.12).

In the older biochemistry literature triacylglycerols were called triglycerides (Section 9.3).

The lipid synthesis reactions are catalyzed by two separate acyltransferases that use fatty acyl CoA molecules as the acyl group donors. The first acyltransferase is glycerol-3-phosphate acyltransferase. It catalyzes esterification at C-1 of glycerol 3-phosphate to form 1-acylglycerol 3-phosphate (lysophosphatidate) and it exhibits a preference for saturated fatty acyl chains. The second acyltransferase is 1-acylglycerol-3-phosphate acyltransferase and it catalyzes esterification at C-2 of 1-acylglycerol 3-phosphate. This enzyme



▲ **Figure 16.9**
Formation of phosphatidate. Glycerol 3-phosphate acyltransferase catalyzes esterification at C-1 of glycerol 3-phosphate. It has a preference for saturated acyl chains. 1-Acylglycerol-3-phosphate acyltransferase catalyzes esterification at C-2 and has a preference for unsaturated acyl chains.



▲ **Figure 16.10**
Synthesis of triacylglycerols and neutral phospholipids. The formation of triacylglycerols, phosphatidylcholine, and phosphatidylethanolamine proceeds via a diacylglycerol intermediate. A cytosine-nucleotide derivative donates the polar head groups of the phospholipids. Three enzymatic methylation reactions, in which *S*-adenosylmethionine is the methyl-group donor, convert phosphatidylethanolamine to phosphatidylcholine.

prefers unsaturated chains. The product of the two reactions is a phosphatidate, one of a family of molecules whose specific properties depend on the attached acyl groups.

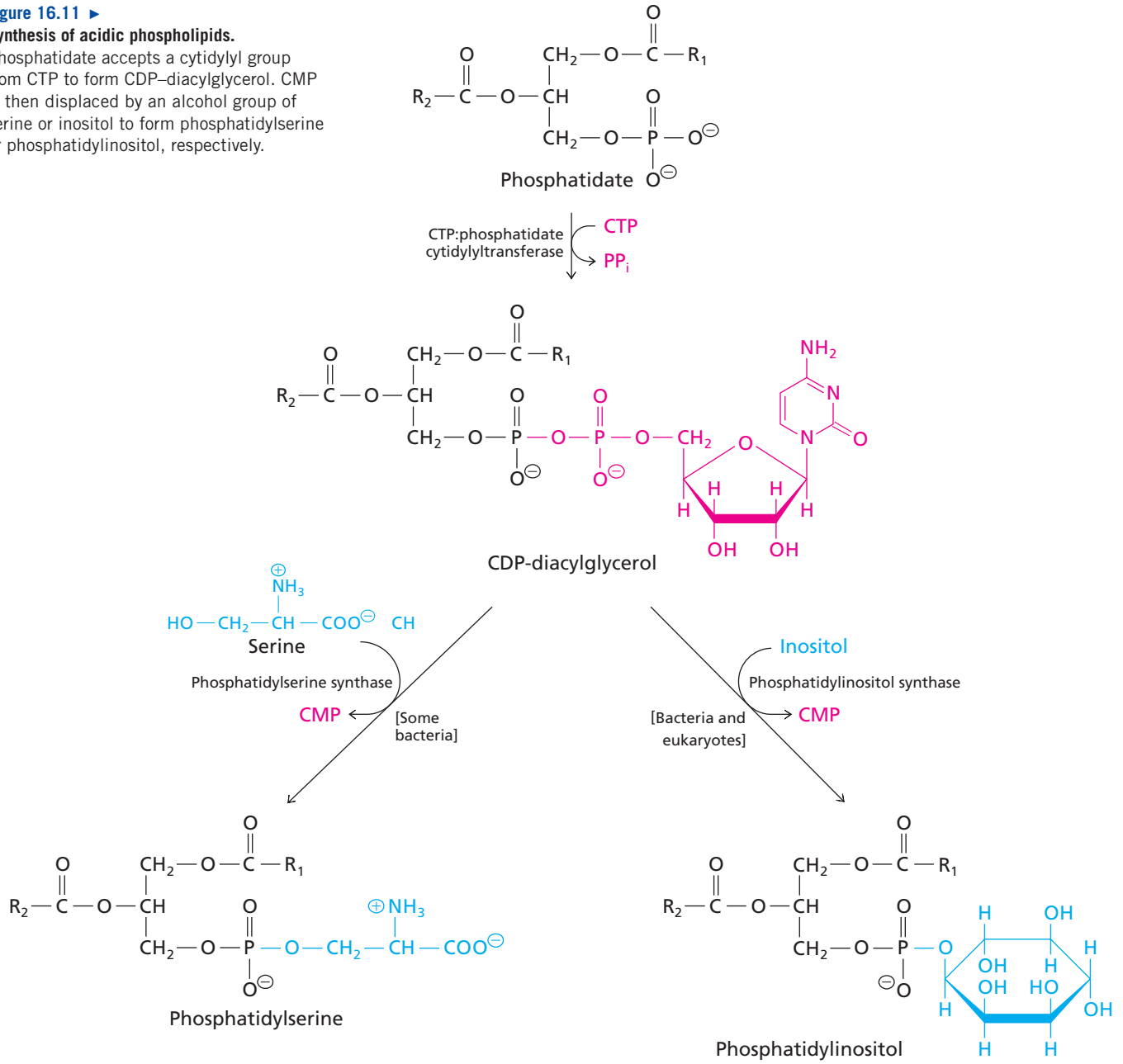
The formation of triacylglycerols and neutral phospholipids from phosphatidate begins with a dephosphorylation catalyzed by phosphatidate phosphatase (Figure 16.10). The product of this reaction is a 1,2-diacylglycerol that can be directly acylated to form a triacylglycerol. Alternatively, 1,2-diacylglycerol can react with a nucleotide–alcohol derivative, such as CDP–choline or CDP–ethanolamine (Section 7.3), to form phosphatidylcholine or phosphatidylethanolamine, respectively. These derivatives are formed from CTP by the general reaction

$$\text{CTP} + \text{Alcohol phosphate} \longrightarrow \text{CDP–alcohol} + \text{PP}_i \qquad \textbf{(16.5)}$$

Phosphatidylcholine can also be synthesized by methylation of phosphatidylethanolamine by *S*-adenosylmethionine (Section 7.3).

Phosphatidate is also the precursor of acidic phospholipids. In this pathway, phosphatidate is first activated by reacting with CTP to form CDP–diacylglycerol with the release of pyrophosphate (Figure 16.11). In some bacteria, the displacement of CMP by serine produces phosphatidylserine. In both prokaryotes and eukaryotes, displacement of CMP by inositol produces phosphatidylinositol. Phosphatidylinositol can be converted to phosphatidylinositol 4-phosphate (PIP) and phosphatidylinositol 4,5-*bis*phosphate

Figure 16.11 ▶
Synthesis of acidic phospholipids.
Phosphatidate accepts a cytidylyl group from CTP to form CDP–diacylglycerol. CMP is then displaced by an alcohol group of serine or inositol to form phosphatidylserine or phosphatidylinositol, respectively.



(PIP_2) through successive ATP-dependent phosphorylation reactions. Recall that PIP_2 is the precursor of the second messengers inositol 1,4,5-*tris*phosphate (IP_3) and diacylglycerol (Section 9.11D).

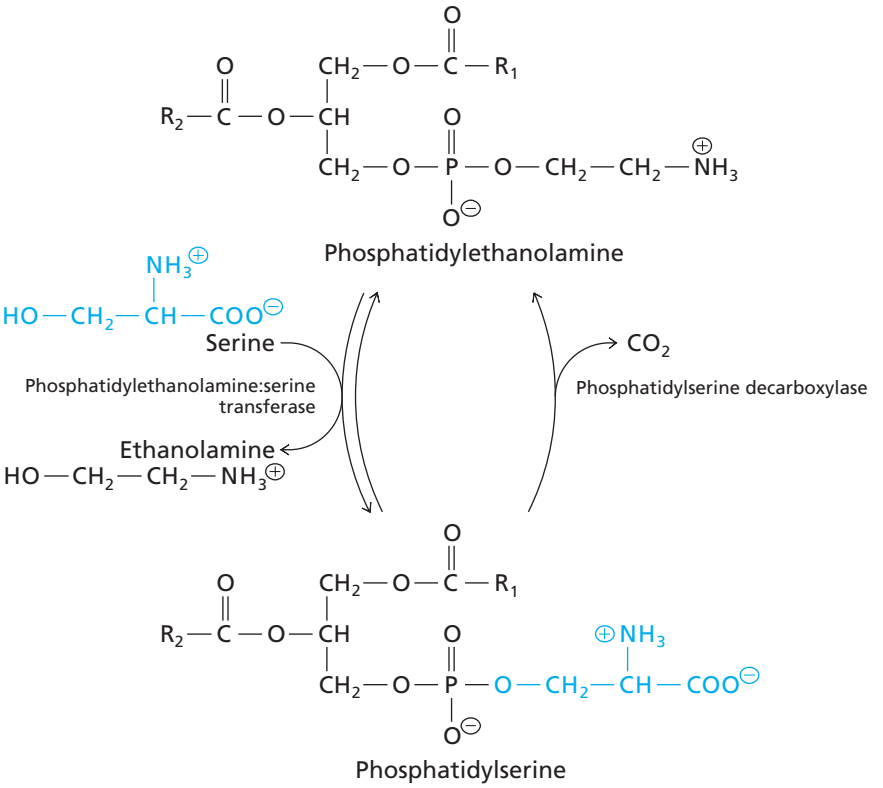
Most eukaryotes use a different pathway for the synthesis of phosphatidylserine. It is formed from phosphatidylethanolamine via the reversible displacement of ethanolamine by serine, catalyzed by phosphatidylethanolamine:serine transferase (Figure 16.12). Phosphatidylserine can be converted back to phosphatidylethanolamine in a decarboxylation reaction catalyzed by phosphatidylserine decarboxylase.

16.3 Synthesis of Eicosanoids

There are two general classes of eicosanoids: prostaglandins + thromboxanes, and leukotrienes. Arachidonate (20:4 $\Delta^{5,8,11,14}$) is the precursor of many eicosanoids. Recall that arachidonate is synthesized from linoleoyl CoA (18:2 $\Delta^{9,12}$) in a pathway that requires a Δ^6 desaturase, an elongase, and a Δ^5 desaturase as shown in Figure 16.8.

Prostaglandins are synthesized by the cyclization of arachidonate in a reaction catalyzed by a bifunctional enzyme called prostaglandin endoperoxide H synthase (PGHS).

Figure 16.12 ►
Interconversions of phosphatidylethanolamine and phosphatidylserine.



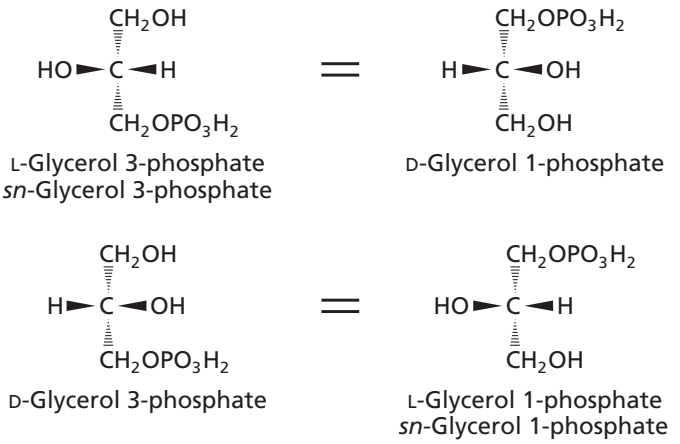
The enzyme is bound to the inner surface of the endoplasmic reticulum through a cluster of hydrophobic α helices that penetrate one of the lipid bilayers (Figure 16.13). The cyclooxygenase (COX) activity of the enzyme catalyzes the formation of a hydroperoxide (prostaglandin G_2). The PGHS enzyme contains a second active site for a hydroperoxidase activity that rapidly converts the unstable hydroperoxide to prostaglandin H_2 (Figure 16.14). This product is converted to various short-lived regulatory molecules including prostacyclin, prostaglandins, and thromboxane A_2. Unlike hormones, which are

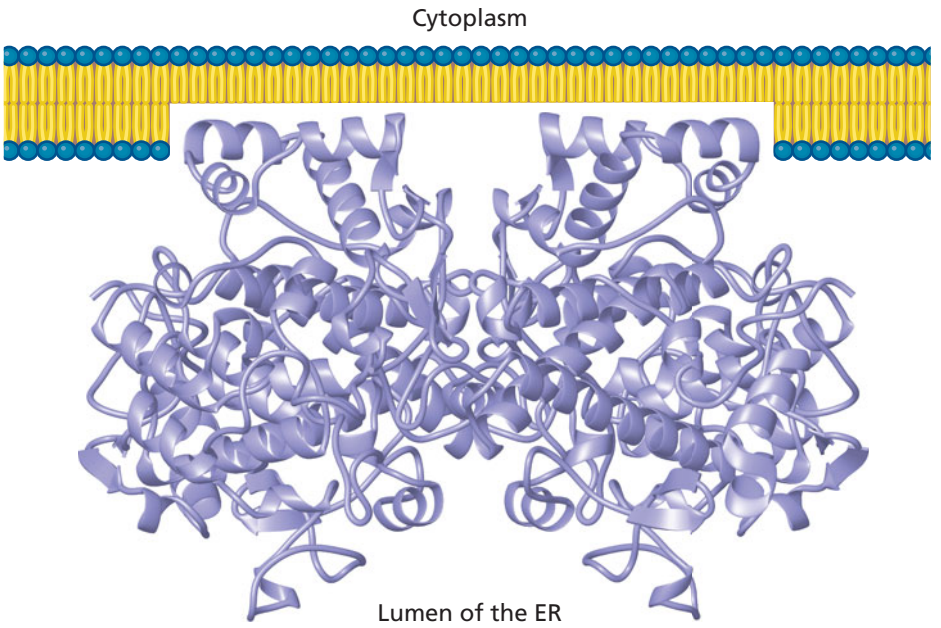
BOX 16.1 *sn*-GLYCEROL 3-PHOSPHATE

One of the precursors for synthesis of triacylglycerols is glycerol 3-phosphate shown as a Fischer projection in Figure 16.9. This molecule could also be accurately drawn upside down as glycerol 1-phosphate. This changes the stereochemical naming convention from L to D. Similarly, D-glycerol 3-phosphate and L-glycerol 1-phosphate are different names for the same molecule.

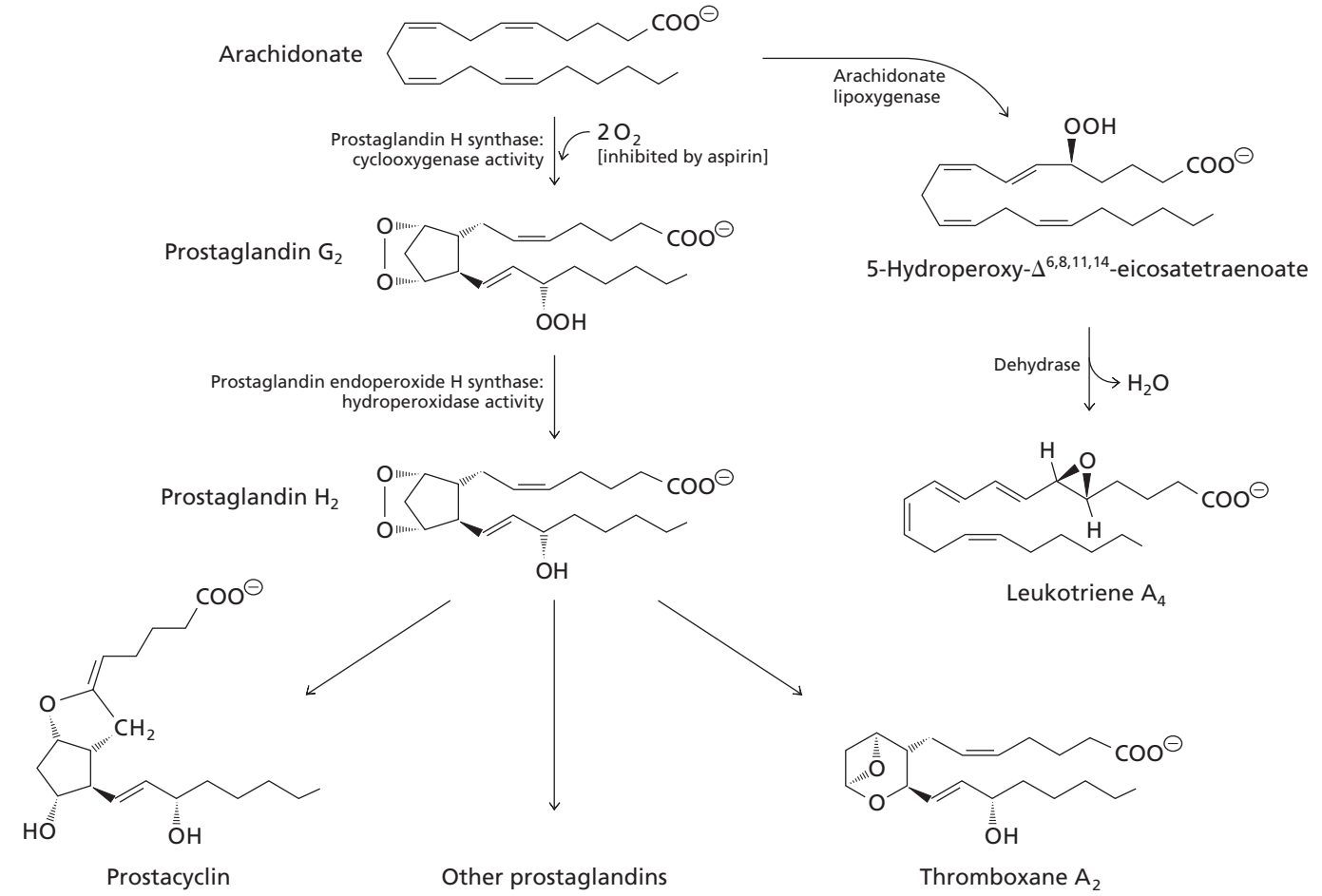
Having different names for the same molecule could lead to confusion since the glycerol phosphate precursor is a prochiral molecule meaning that modified lipids will have different stereochemical names depending on whether you start with L-glycerol 3-phosphate or D-glycerol 1-phosphate. In order to avoid this, a new convention is introduced to number the carbon atoms. In a Fischer projection where the hydroxyl group on C-2 is on the left, the "top" carbon atom becomes C-1 and the "bottom" one is C-3. Thus, L-glycerol 3-phosphate becomes *sn*-glycerol 3-phosphate where "*sn*" stands for *Stereochemical numbering.*

The accurate name for the triglyceride precursor is *sn*-glycerol 3-phosphate in most cases. In archaebacteria the precursor is *sn*-glycerol 1-phosphate (Box 9.5).





◀ **Figure 16.13**
Prostaglandin endoperoxide H synthase (PGHS, COX-1). This enzyme is a dimer bound to the inner membrane of the endoplasmic reticulum. The active sites of the cyclooxygenase and hydroperoxidase activities are located in the large cleft at the bottom of the enzyme. [from sheep, *Ovis aries*, PDB 1PRH]



▲ **Figure 16.14**
Major pathways for the formation of eicosanoids. The prostaglandin H synthase (PGHS) pathway leads to prostaglandin H_2 that can be converted to prostacyclin, thromboxane A_2 and a variety of prostaglandins. The lipoxygenase pathway shown produces leukotriene A_4 a precursor of some other leukotrienes. The cyclooxygenase activity of PGHS is inhibited by aspirin.



▲ **The bark of willow trees is a natural source of salicylates.**

produced by glands and travel in the blood to their sites of action, eicosanoids typically act in the immediate neighborhood of the cell in which they are produced. For example, thromboxane A_2 is produced by blood platelets and it leads to platelet aggregation and blood clots and constriction of the smooth muscles in arterial walls causing localized changes in blood flow. The uterus produces contraction-triggering prostaglandins during labor. Eicosanoids also mediate pain sensitivity, inflammation, and swelling.

Recall that linoleate must be supplied in the human diet, usually from plants, in order to support the synthesis of arachidonate and eicosanoids. One of the reasons why linoleate is essential is because it's required for synthesis of prostaglandins and prostaglandins are necessary for survival.

Aspirin blocks production of some eicosanoids and thus relieves the symptoms of pain and reduces fever. The active ingredient of aspirin, acetylsalicylic acid, irreversibly inhibits COX activity by transferring an acetyl group to an active-site serine residue of the bifunctional enzyme. By blocking the activity of COX, aspirin prevents the formation of a variety of eicosanoids that are synthesized after the COX reaction. Aspirin was first developed as a marketable drug in 1897 but other salicylates have long been used in the treatment of pain. The ancient Greeks, for example, used the bark of willow trees for pain relief. Willow bark is a natural source of salicylates.

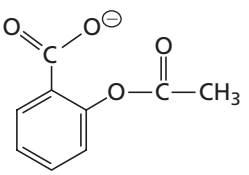
The second class of eicosanoids are the products of reactions catalyzed by lipoxygenases. In Figure 16.14, arachidonate lipoxygenase is shown catalyzing the first step in the pathway leading to leukotriene A_4. (The term *triene* refers to the presence of three conjugated double bonds.) Further reactions produce other leukotrienes, such as the compounds once called the "slow-reacting substances of anaphylaxis" (allergic response) that are responsible for the occasionally fatal effects of exposure to antigens.

BOX 16.2 THE SEARCH FOR A REPLACEMENT FOR ASPIRIN

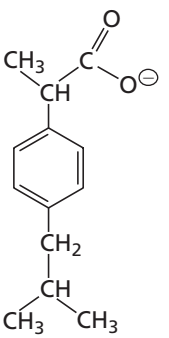
Most natural salicylates have serious side effects. They cause inflamation of the mouth, throat, and stomach and they taste horrible. Aspirin avoids most of these side effects, which is why it became such a popular drug when it was first introduced. However, aspirin can cause dizziness, ringing in the ears, and bleeding or ulcers of the stomach lining. There are two different forms of PGHS (also called COX after their cyclooxygenase activity). COX-1 is a constitutive enzyme that regulates secretion of mucin in the stomach, thus protecting the gastric wall. COX-2 is an inducible enzyme that promotes inflammation, pain, and fever. Aspirin inhibits both isozymes.

There are many other nonsteroidal anti-inflammatory drugs (NSAIDS) that inhibit COX activity. Aspirin is the only one that inhibits by covalent modification of the enzyme. The others act by competing with arachidonate for binding to the COX active site. Ibuprofen (Advil®), for example, binds rapidly, but weakly, to the active site and its inhibition is readily reversed when drug levels drop. Acetaminophen (Tylenol®) is an effective inhibitor of COX activity in intact cells.

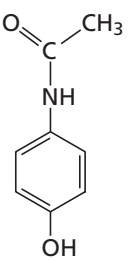
Physicians would like to have a drug that selectively inhibits COX-2 and not COX-1. Such a compound would not cause stomach irritation. A number of specific COX-2 inhibitors have been synthesized and many are currently available for patients. These drugs, while expensive, are important for patients with arthritis who must take pain killers on a regular basis. In some cases, the new NSAIDS have been associated with increased risk of cardiovascular disease and they have been taken off the market (e.g., Vioxx®). X-ray crystallographic studies of COX-2 and its interaction with these inhibitors has aided the search for even better replacements for aspirin.



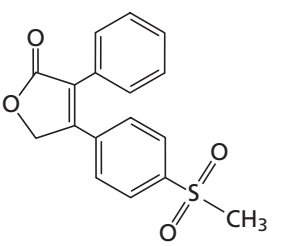
Aspirin



Ibuprofen



Acetaminophen



Rofecoxib (Vioxx®)
(COX-2 specific NSAID)

16.4 Synthesis of Ether Lipids

Ether lipids have an ether linkage in place of one of the usual ester linkages (Section 9.4). The pathway for the formation of ether lipids in mammals begins with dihydroxyacetone phosphate (Figure 16.15). First, an acyl group from fatty acyl CoA is esterified to the oxygen atom at C-1 of dihydroxyacetone phosphate producing 1-acyldihydroxyacetone phosphate. Next, a fatty alcohol displaces the fatty acid to produce 1-alkyldihydroxyacetone phosphate. The keto group of this compound is then reduced by NADPH to form 1-alkylglycero-3-phosphate. This reduction is followed by esterification at C-2 of the glycerol residue to produce 1-alkyl-2-acylglycero-3-phosphate. The subsequent reactions—dephosphorylation and addition of a polar head group (either choline or ethanolamine)—are the same as those shown earlier in Figure 16.10. Plasmalogens, which contain a vinyl ether linkage at C-1 of the glycerol backbone (Figure 9.9), are formed from alkyl ethers by oxidation of the alkyl ether linkage. This reaction is catalyzed by an oxidase that requires NADH and O_2. The oxidase is similar to the acyl-CoA desaturases (Figure 16.8) that introduce double bonds into fatty acids.

In eukaryotes, ether lipids are not as common as the glycerophospholipids containing ester linkages although some species and some tissues have membranes that are enriched in plasmalogens. Ether lipids are more common in bacteria, especially in archaebacteria where the majority of membrane lipids are ether lipids (Box 9.5).

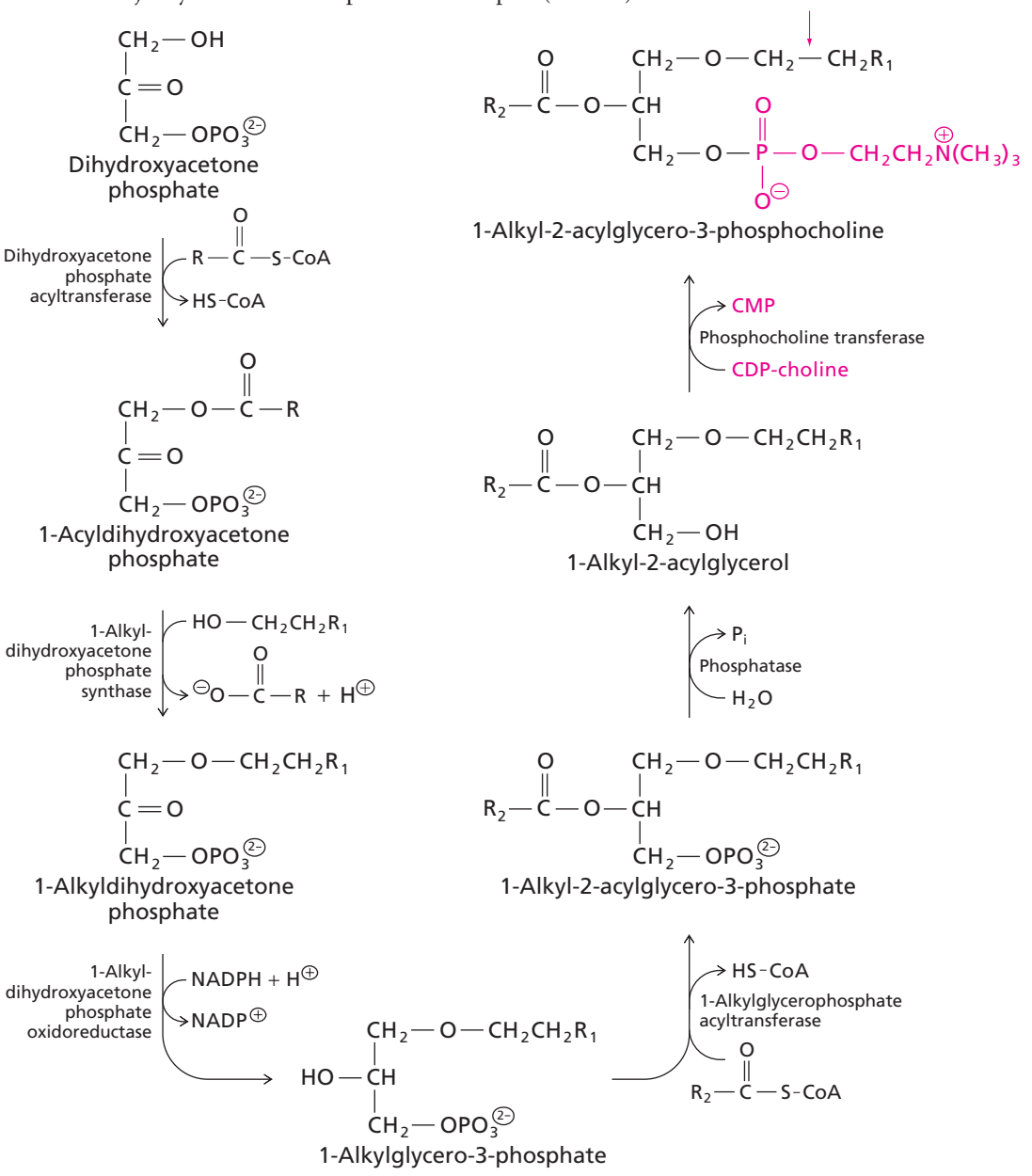


Figure 16.15 ▼
Synthesis of ether lipids. Plasmalogens are synthesized from ether lipids by the formation of a double bond at the position marked with a red arrow.

16.5 Synthesis of Sphingolipids

Sphingolipids are membrane lipids that have sphingosine (a C_{18} unsaturated amino alcohol) as their structural backbone (Figure 9.10). In the first step of sphingolipid biosynthesis, serine (a C_3 unit) condenses with palmitoyl CoA, producing 3-ketosphinganine and CO_2 (Figure 16.16). Reduction of 3-ketosphinganine by NADPH produces sphinganine. Next, a fatty acyl group is transferred from acyl CoA to the amino group of sphinganine in an *N*-acylation reaction. The product of this reaction is dihydroceramide, or ceramide without the characteristic double bond between C-4 and C-5 of a typical sphingosine. This double bond is introduced in a reaction catalyzed by dihydroceramide Δ^4-desaturase, an enzyme that is similar to other desaturases that we have encountered. The final product is ceramide (*N*-acylsphingosine).

Ceramide is the source of all the other sphingolipids. It can react with phosphatidylcholine to form sphingomyelin or with a UDP–sugar to form a cerebroside. Complex sugar–lipid conjugates, gangliosides, can be formed by reaction of a cerebroside with additional UDP-sugars and CMP-*N*-acetylneuraminic acid (Figure 9.12). Gangliosides are found in the outer leaflet of the plasma membrane, as are most glycolipids.

16.6 Synthesis of Cholesterol

The steroid cholesterol is an important component of many membranes (Section 9.8) and a precursor of steroid hormones and bile salts in mammals. All the carbon atoms in cholesterol come from acetyl CoA, a fact that emerged from early radioisotopic labeling experiments. Squalene, a C_{30} linear hydrocarbon, is an intermediate in the biosynthesis of the 27-carbon cholesterol molecule. Squalene is formed from 5-carbon units related to isoprene. The stages in the cholesterol biosynthesis pathway are

$$\text{Acetate } (C_2) \longrightarrow \text{Isoprenoid } (C_5) \longrightarrow \text{Squalene } (C_{30}) \longrightarrow \text{Cholesterol } (C_{27}) \quad \textbf{(16.6)}$$

A. Stage 1: Acetyl CoA to Isopentenyl Diphosphate

Mitochondrial isozymes of acetoacetyl-CoA thiolase and HMG-CoA synthase are involved in the synthesis of ketone bodies (Section 16.11).

The first step in cholesterol synthesis is sequential condensation of three molecules of acetyl CoA. These condensation steps are catalyzed by acetoacetyl-CoA thiolase and HMG-CoA synthase. The product, HMG CoA, is then reduced to mevalonate in a reaction catalyzed by HMG-CoA reductase (Figure 16.17). This is the first committed step in cholesterol synthesis. Mevalonate is converted to the C_5 compound isopentenyl diphosphate by two phosphorylations followed by decarboxylation. The conversion of three molecules of acetyl CoA to isopentenyl diphosphate requires energy in the form of three ATP and two NADPH. In addition to its role in cholesterol synthesis, isopentenyl diphosphate is an important donor of isoprenyl units for many other biosynthesis reactions.

Many species of bacteria have a completely different, mevalonate-independent pathway for synthesis of isopentyl diphosphate. The initial precursors in this pathway are glyceraldehyde 3-phosphate + pyruvate and not acetyl CoA. The mevalonate-independent pathway is more ancient than the mevalonate-dependent pathway shown here.

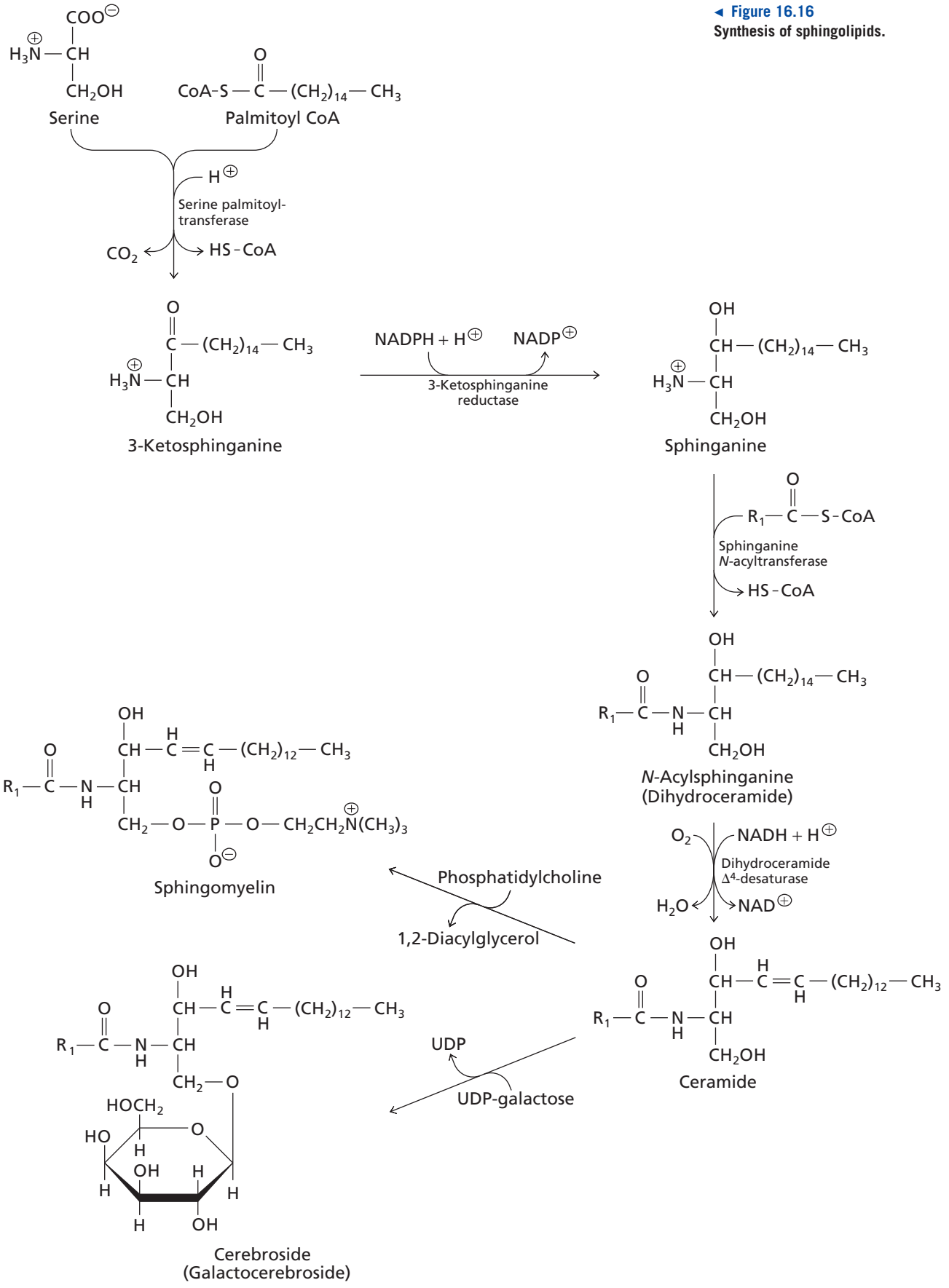
B. Stage 2: Isopentenyl Diphosphate to Squalene

KEY CONCEPT

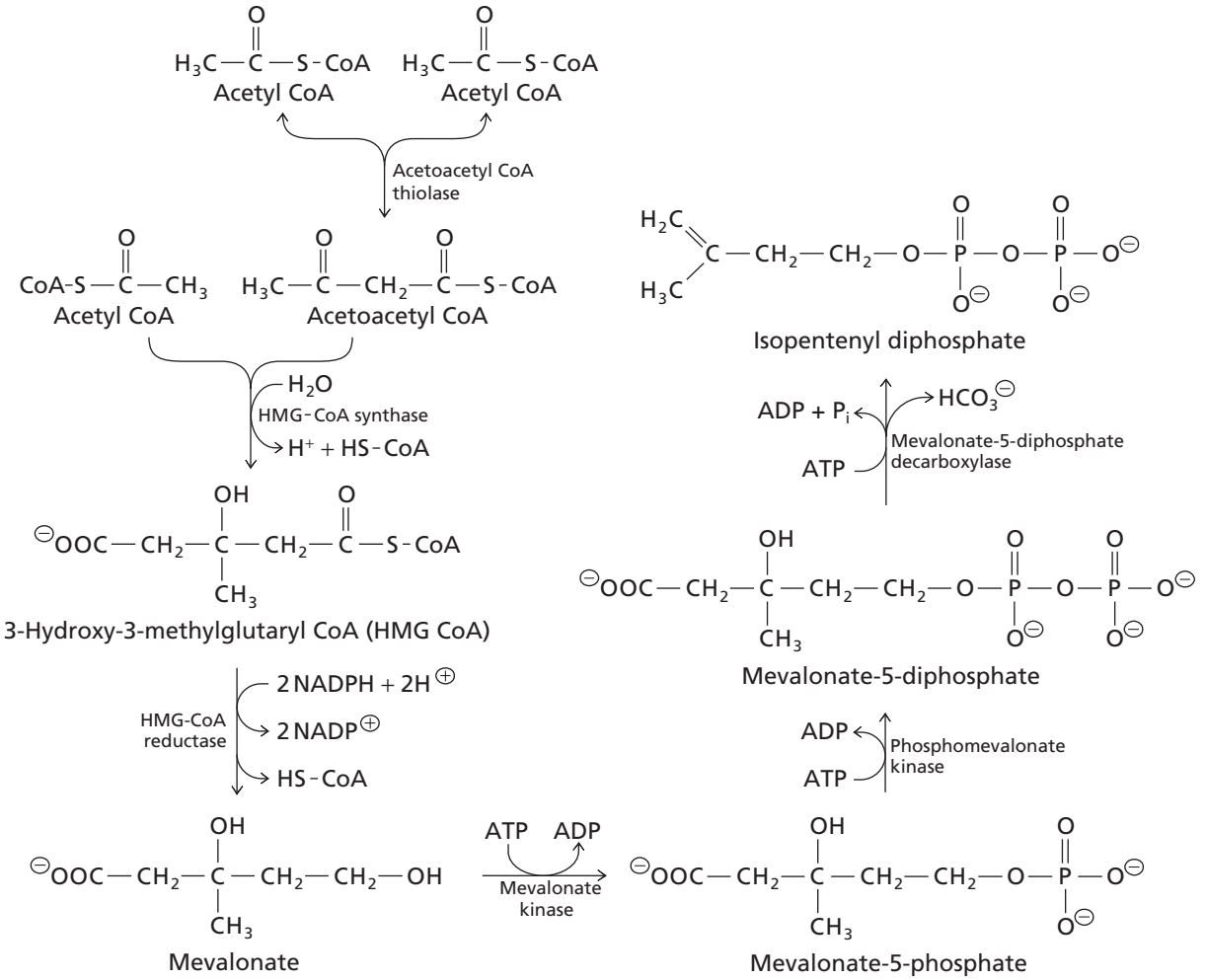
Isopentenyl diphosphate is the precursor for synthesis of all isoprenoids.

Isopentenyl diphosphate is converted to dimethylallyl diphosphate by a specific isomerase called isopentenyl diphosphate isomerase (IDI). The two isomers are then joined in a head-to-tail condensation reaction catalyzed by prenyl transferase (Figure 16.18). The products of this reaction are a C_{10} molecule (geranyl diphosphate) and pyrophosphate. A second condensation reaction, also catalyzed by prenyl transferease, produces the important C_{15} intermediate, farnesyl diphosphate. The condensation of isoprenyl units produces a characteristic branched hydrocarbon with regularly spaced double bonds at the branch position. These isoprene units (Figure 9.13) are present in a number of important cofactors.

Two molecules of farnesyl diphosphate are joined in a head-to-head condensation reaction to form the C_{30} molecule squalene. Pyrophosphate, whose hydrolysis drives reaction equilibria toward completion, is produced in three steps in the squalene synthesis pathway. Note that all double bonds in squalene are *trans*.



◀ **Figure 16.16**
Synthesis of sphingolipids.



▲ **Figure 16.17**
Stage I of cholesterol synthesis: formation of isopentenyl diphosphate. The condensation of three acetyl CoA molecules leads to HMG CoA, which is reduced to mevalonate. Mevalonate is then converted to the five-carbon molecule isopentenyl diphosphate via two phosphorylations and one decarboxylation.

C. Stage 3: Squalene to Cholesterol

The steps between squalene and the first fully cyclized intermediate, lanosterol, include the addition of a hydroxyl group followed by a concerted series of cyclizations to form the four-ring steroid nucleus (Figure 16.19). Lanosterol accumulates in appreciable quantities in cells that are actively synthesizing cholesterol. The conversion of lanosterol to cholesterol occurs via two pathways, both involving many steps.

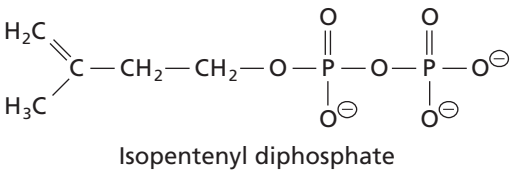
D. Other Products of Isoprenoid Metabolism

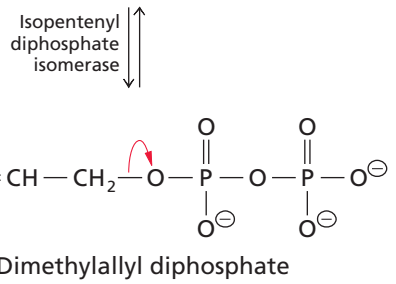
A multitude of isoprenoids are synthesized from cholesterol or its precursors. Isopentenyl diphosphate, the C_5 precursor of squalene, is the precursor of a large number of other products, such as quinones; the lipid vitamins A, E, and K; carotenoids; terpenes; the side chains of some cytochrome heme groups; and the phytol side chain of chlorophyll (Figure 16.20). Many of these isoprenoids are made in bacteria, which do not synthesize

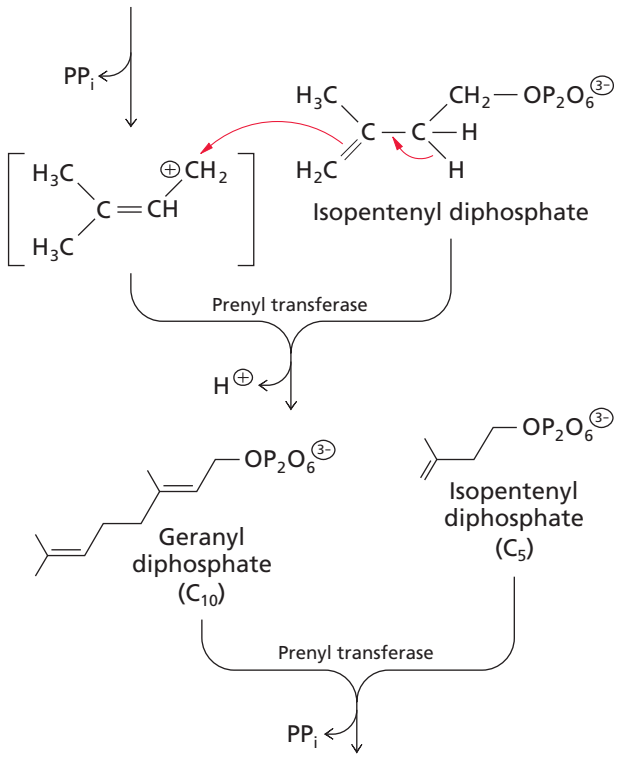


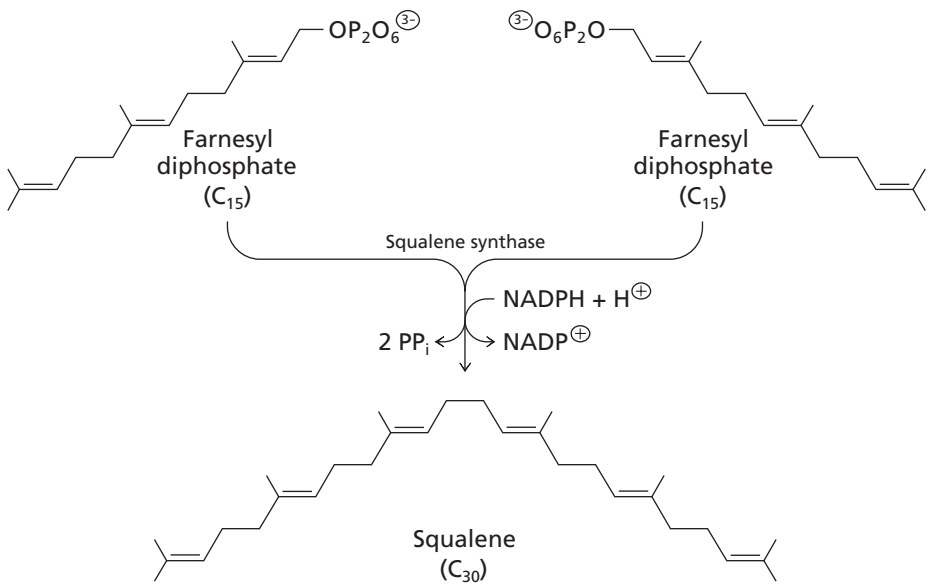


◀ **Konrad Bloch (1912–2000)** (top) and **Feodor Lynen (1911–1979)** (bottom) received the Nobel Prize in Physiology or Medicine in 1964 "for their discoveries concerning the mechanism and regulation of the cholesterol and fatty acid metabolism".

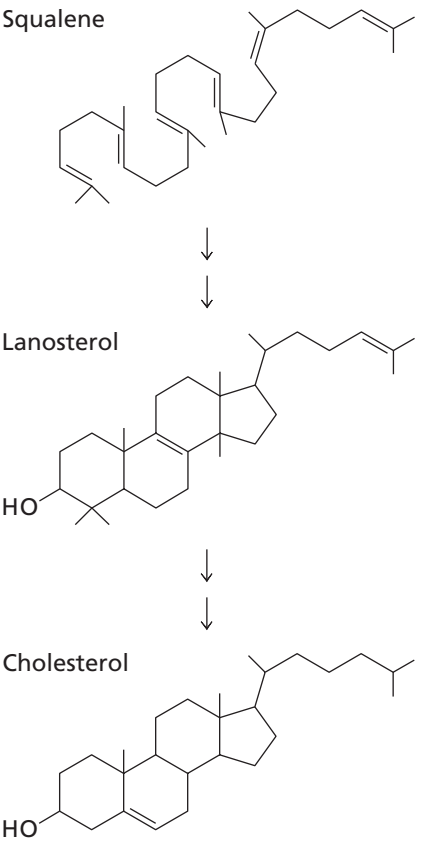








◀ **Figure 16.18**
Condensation reactions in the second stage of cholesterol synthesis.



▲ **Figure 16.19**
Final stage of cholesterol synthesis: squalene to cholesterol. The conversion of lanosterol to cholesterol requires up to 20 steps.

BOX 16.3 REGULATING CHOLESTEROL LEVELS

The HMG-CoA reductase reaction appears to be the principal site for the regulation of cholesterol synthesis. HMG-CoA reductase has three regulatory mechanisms—covalent modification, repression of transcription, and control of degradation. Short-term control is effected by covalent modification: HMG-CoA reductase is an interconvertible enzyme that is inactivated by phosphorylation. This phosphorylation is catalyzed by an unusual AMP-activated protein kinase that can also catalyze the phosphorylation and concomitant inactivation of acetyl-CoA carboxylase (Section 16.9). The action of the kinase appears to decrease the ATP-consuming synthesis of both cholesterol and fatty acids when AMP levels rise. The amount of HMG-CoA reductase in cells is also closely regulated. Cholesterol (endogenous cholesterol delivered by plasma lipoproteins or dietary cholesterol delivered by chylomicrons) can repress transcription of the gene that encodes HMG-CoA reductase. In addition, high levels of cholesterol and its derivatives increase the rate of degradation of HMG-CoA reductase, possibly by increasing the rate of transport of the membrane-bound enzyme to the site of its degradation.

Lowering of serum cholesterol levels decreases the risk of coronary heart disease. A number of drugs called statins are potent competitive inhibitors of HMG-CoA reductase. Statins are often used as part of the treatment of hypercholesterolemia because they can effectively lower blood cholesterol levels. Another useful approach is to bind bile salts in the intestine to resin particles, to prevent their reabsorption. More cholesterol must then be converted to bile salts. Inhibition of HMG-CoA reductase may not be the most desirable method for controlling cholesterol levels because mevalonate is needed for the synthesis of important molecules such as ubiquinone.

► **Structure of HMG-CoA and two common statins.**

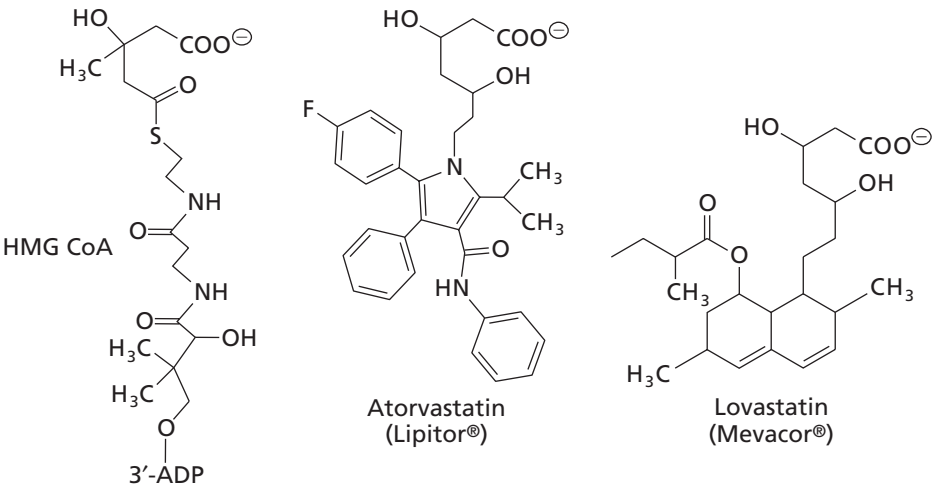
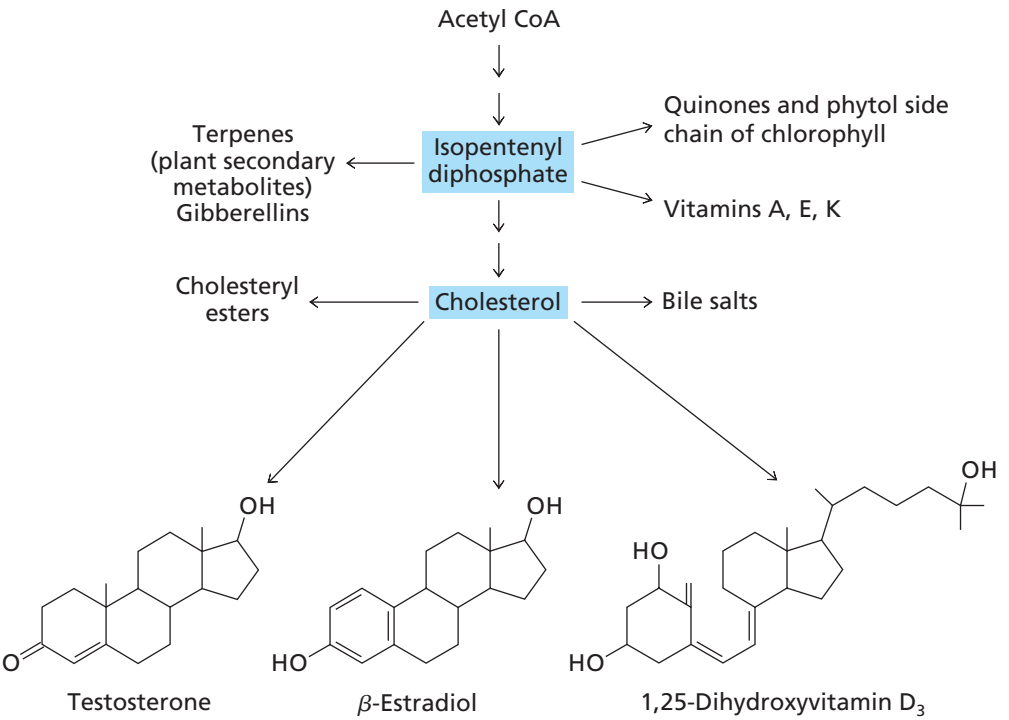


Figure 16.20 ►
Other products of isopentenyl diphosphate and cholesterol metabolism.



cholesterol. The two pathways for the biosynthesis of isopentyl diphosphate (Section 16.6A) are much more ancient than the more recent cholesterol biosynthesis pathway.

Cholesterol is the precursor of bile salts, which facilitate intestinal absorption of lipids; vitamin D that stimulates Ca^{2+} uptake from the intestine; steroid hormones such as testosterone and β-estradiol that control sex characteristics; and steroids that control salt balance. The principal product of steroid synthesis in mammals is cholesterol itself, which modulates membrane fluidity and is an essential component of the plasma membrane of animal cells.

16.7 Fatty Acid Oxidation

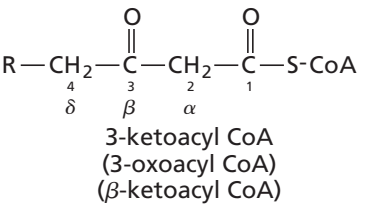
Fatty acids, released from triacylglycerols (Section 16.9), are oxidized by a pathway that degrades them by removing two-carbon units at each step. The two-carbon fragments are transferred to coenzyme A to form acetyl CoA, and the remainder of the fatty acid re-enters the oxidative pathway. This degradative process is called the β-oxidation pathway because the β-carbon atom (C-3) of the fatty acid is oxidized. Fatty acid oxidation is divided into two stages: activation of fatty acids and degradation to two-carbon fragments (as acetyl CoA). The NADH and ubiquinol (QH_2) produced by the oxidation of fatty acids can be oxidized by the respiratory electron transport chain, and the acetyl CoA can enter the citric acid cycle.

Acetyl CoA can be completely oxidized by the citric acid cycle to yield energy (in the form of ATP) that can be used in other biochemical pathways. The carbon atoms from fatty acids can also be used as substrates for amino acid synthesis since several of the intermediates in the citric acid cycle are diverted to amino acid biosynthesis pathways (Section 13.6). In those organisms that possess a glyoxylate pathway (Section 13.7), acetyl CoA from fatty acid oxidation can be used to synthesize glucose via the gluconeogenesis pathway.

The oxidation of fatty acids occurs as part of the normal turnover of membrane lipids. Thus, bacteria, protists, fungi, plants, and animals all have a β-oxidation pathway. In addition to its role in normal cellular metabolism, fatty acid oxidation is a major component of fuel metabolism in animals. A significant percentage of dietary food consists of membrane lipids and fat and this rich course of energy is exploited by oxidizing fatty acids. In this section we describe the basic biochemical pathways of fatty acid oxidation. In the following sections we will discuss the role of fatty acid oxidation in mammalian fuel metabolism.

KEY CONCEPT

β-Oxidation is an ancient and ubiquitous pathway for degradation of fatty acids.



▲ **3-ketoacyl CoA, 3-oxoacyl CoA, β-ketoacyl CoA**

A. Activation of Fatty Acids

The activation of fatty acids for oxidation is catalyzed by acyl-CoA synthetase (Figure 16.7). This is the same activation step that is required for the synthesis of polyunsaturated fatty acids and complex lipids.

B. The Reactions of β-Oxidation

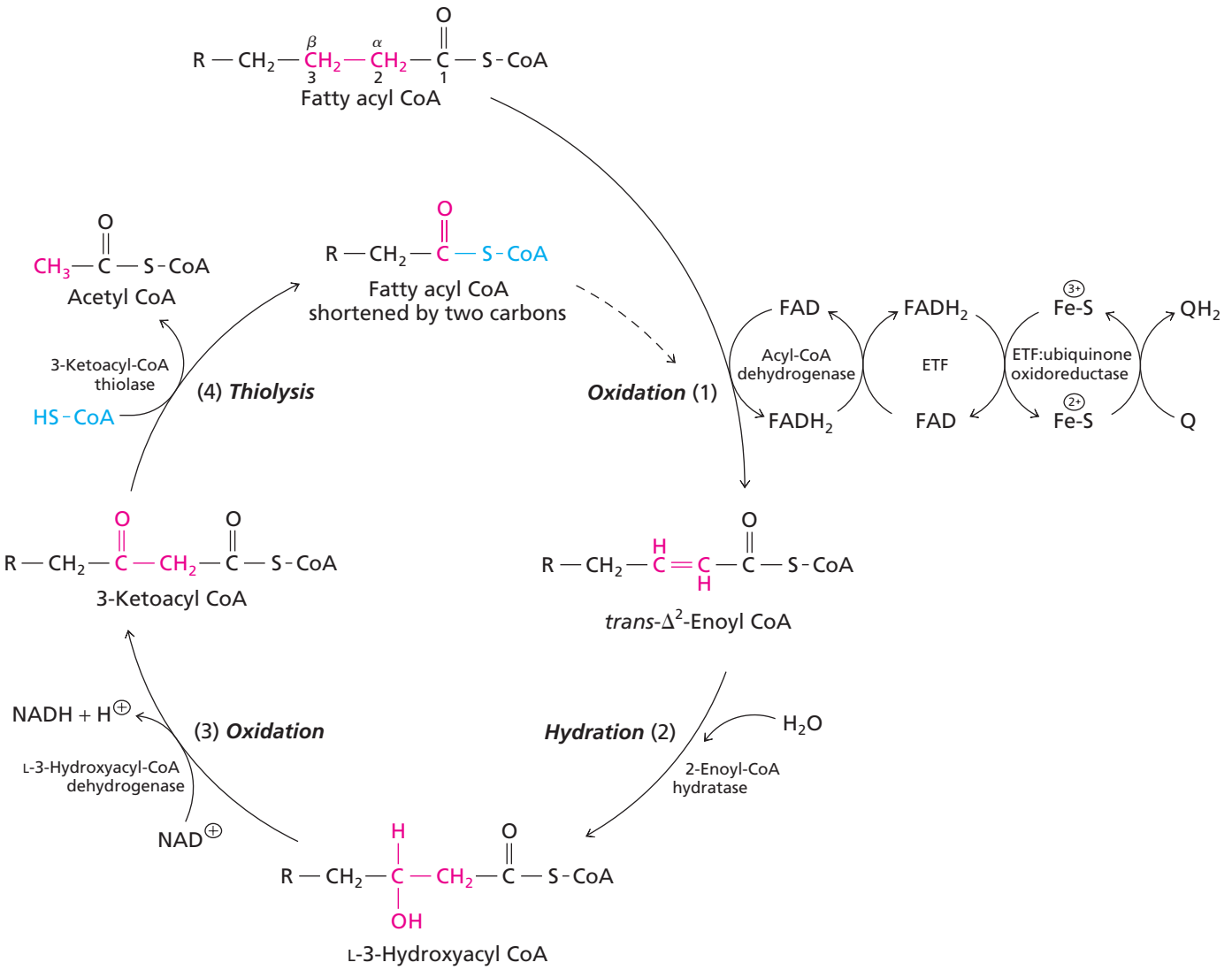
In eukaryotes, β-oxidation takes place in mitochondria and in specialized organelles called peroxisomes. In bacteria, the reactions take place in the cytosol. Four steps are required to produce acetyl CoA from fatty acyl CoA: oxidation, hydration, further oxidation, and thiolysis (Figure 16.21). We focus first on the oxidation of a saturated fatty acid with an even number of carbon atoms.

In the first oxidation step, acyl-CoA dehydrogenase catalyzes the formation of a double bond between the C-2 and C-3 atoms of the acyl group forming *trans* 2-enoyl CoA. There are several separate acyl-CoA dehydrogenase isozymes, each with a different chain length preference: short, medium, long, or very long.

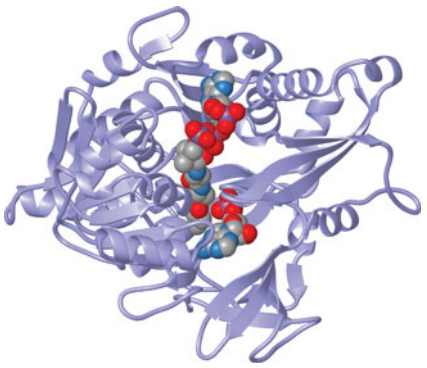
When the double bond is formed, electrons from fatty acyl CoA are transferred to the FAD prosthetic group of acyl-CoA dehydrogenase and then to another FAD prosthetic group bound to a mobile, water-soluble, protein coenzyme called electron



▲ **Bear bile.** In Vietnam bears are kept in captivity—often under deplorable conditions—and bile is extracted from their stomachs on a regular basis. Bear bile is thought to be an effective remedy for fever and poor eyesight.



▲ **Figure 16.21**
***β*-oxidation of saturated fatty acids.** One round of *β*-oxidation consists of four enzyme-catalyzed reactions. Each round generates one molecule each of QH_2, NADH, acetyl CoA, and a fatty acyl CoA molecule two carbon atoms shorter than the molecule that entered the round. (ETF is the electron-transferring flavoprotein, a water-soluble protein coenzyme.)



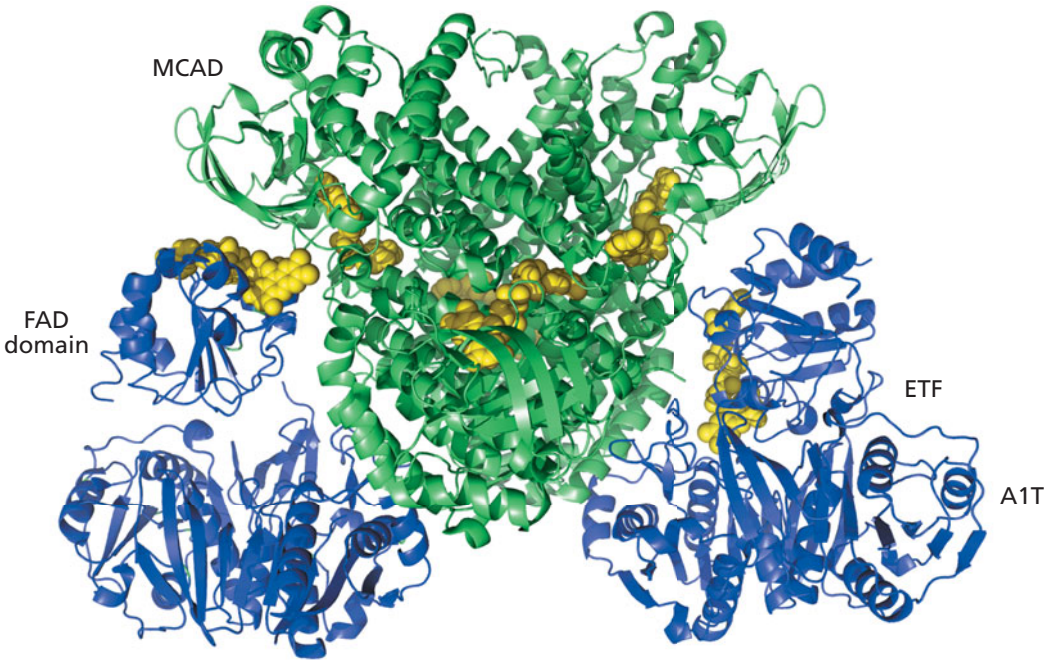
▲ **Human medium chain acyl-CoA synthetase.** The products of the reaction, AMP and acyl CoA, are bound in the active site. The enzyme is a dimer but only one subunit is shown. [PDB 3EQ6]

transferring flavoprotein (ETF, Figure 16.22). (ETF also accepts electrons from several other flavoproteins that are not involved in fatty acid metabolism.) Electrons are then passed to Q in a reaction catalyzed by ETF:ubiquinone oxidoreductase. This enzyme is embedded in the membrane and QH_2 from fatty acid oxidation enters the pool of QH_2 that can be oxidized by the membrane-associated electron transport system.

The second step is a *hydration* reaction. Water is added to the unsaturated *trans* 2-enoyl CoA produced in the first step to form the L isomer of 3-hydroxyacyl CoA. The enzyme is 2-enoyl-CoA hydratase.

The third step is a second *oxidation* catalyzed by L-3-hydroxyacyl-CoA dehydrogenase. This production of 3-ketoacyl CoA from 3-hydroxyacyl CoA is an $NAD^{\oplus}$-dependent reaction. The resulting reducing equivalents (NADH) can be used directly in biosynthesis pathways or they can be oxidized by the membrane-associated electron transport system.

Finally, in Step 4, the nucleophilic sulfhydryl group of HS–CoA attacks the carbonyl carbon of 3-ketoacyl CoA in a reaction catalyzed by 3-ketoacyl-CoA thiolase. This enzyme, also called thiolase II, is related to the acetoacyl-CoA thiolase (thiolase I) that we encountered



▲ **Figure 16.22**
Model of the medium chain acyl-CoA dehydrogenase (MCAD) bound to ETF. The MCAD subunits are colored green and the ETF subunits are colored blue. Bound FADs are represented as space-filling molecules (yellow). The model is based on the structure in PDB entry 2A1T containing a mutant protein that blocks movement of the FAD domain of ETF. The left-hand side of the dimer shows the probable position of the FAD domain during transfer of electrons from MCAD to ETF and the right-hand side shows the position of the FAD domain in free, unbound ETF. The flexibility of the FAD domain as it shifts from one position to another is responsible for its lack of resolution in the wild-type ETF: MCAD crystal structure. (Toogood et al., 2004; Toogood et al., 2005)

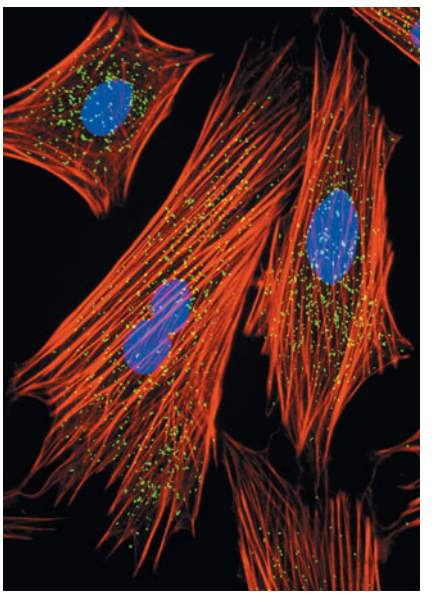
in the isopentenyl diphosphate pathway (Section 16.6A). Acetoacyl-CoA thiolase is specific for acetoacetyl CoA, while 3-ketoacyl-CoA thiolase acts on long chain fatty acid derivatives. The release of acetyl CoA leaves a fatty acyl CoA molecule shortened by two carbons. This acyl CoA molecule is a substrate for another round of the four reactions and the metabolic spiral continues until the entire molecule has been converted to acetyl CoA.

As the fatty acyl chain becomes shorter, the first step is catalyzed by acyl-CoA dehydrogenase isozymes with preferences for shorter chains. Interestingly, the first three reactions of fatty acid oxidation are chemically parallel to three steps of the citric acid cycle. In these reactions, an ethylene group ($—CH_2CH_2—$, as in succinate) is oxidized to a two-carbon unit containing a carbonyl group ($—COCH_2—$, as in oxaloacetate). The steps are the reverse of the reactions in the fatty acid synthesis pathway (Section 16.1C).

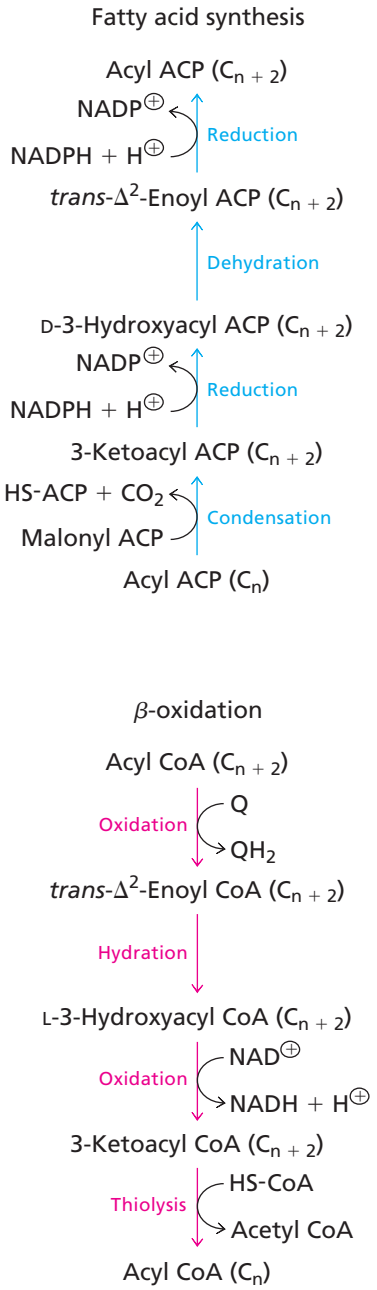
In eukaryotes, fatty acid oxidation also occurs in peroxisomes. In fact, peroxisomes are the only site of fatty acid β-oxidation in most eukaryotes (but not mammals). In peroxisomes, the initial oxidation step is catalyzed by acyl-CoA oxidase—an enzyme that is homologous to the acyl-CoA dehydrogenase that catalyzes the first oxidation in mitochondria. The peroxisomal enzyme transfers electrons to O_2 to form hydrogen peroxide (H_2O_2).

$$\text{Fatty acyl CoA} + O_2 \xrightarrow{\text{Acyl-CoA oxidase}} \textit{trans}\text{-}\Delta^2\text{-Enoyl CoA} + H_2O_2 \qquad \textbf{(16.7)}$$

In bacterial and mitochondrial β-oxidation the product of the first oxidation step is QH_2 that can be used in the respiratory electron transport chain. This results in synthesis of ATP—each QH_2 molecule is equivalent to 1.5 molecules of ATP (Section 14.11). There is no membrane-associated electron transport system in peroxisomes and this is why a different type of oxidation–reduction takes place in peroxisomes. It also means that fewer ATP equivalents are produced during peroxisomal β-oxidation. In mammals, where both mitochondrial and peroxisomal pathways exist, the peroxisomal β-oxidation pathway degrades very long chain fatty acids, branched fatty acids, long chain dicarboxylic acids, and possibly



▲ **Peroxisomes**. Indian Muntjac (*Muntiacus muntjak*) fibroblast cells were stained with green reagent to show peroxisomes. Actin fibers are stained red and nuclear DNA is purple. The small peroxisomes are scattered throughout the cytoplasm. [http://www.microscopyu.com/staticgallery/fluorescence/muntjac.html]



▲ **Figure 16.23**
Fatty acid synthesis and β-oxidation.

trans unsaturated fatty acids producing smaller, more polar compounds that can be excreted. Most of the common fatty acids are degraded in mitochondria.

C. Fatty Acid Synthesis and β-Oxidation

Fatty acid synthesis involves carbon-carbon bond formation (condensation) followed by reduction, dehydration, and reduction steps in preparation for the next condensation reaction. The reverse reactions—oxidation, hydration, oxidation, and carbon-carbon bond cleavage—are part of the degradation pathway of β-oxidation. We compare the two pathways in Figure 16.23.

The active thioesters in fatty acid oxidation are CoA derivatives whereas the intermediates in fatty acid synthesis are bound as thioesters to acyl carrier protein (ACP). In both cases, the acyl groups are attached to phosphopantetheine. Synthesis and degradation both proceed in two-carbon steps. However, oxidation results in a two-carbon product, acetyl CoA, whereas synthesis requires a three-carbon substrate, malonyl ACP that transfers a two-carbon unit to the growing chain releasing CO_2. Reducing power for synthesis is supplied by NADPH, whereas oxidation depends on $NAD^{\oplus}$ and ubiquinone (via the electron-transferring flavoprotein). Finally, the intermediate in fatty acid synthesis is D-3-hydroxyacyl-ACP whereas the L isomer (L-3-hydroxyacyl-CoA) is produced during β-oxidation.

The biosynthesis and catabolic pathways are catalyzed by a completely different set of enzymes and the intermediates form separate pools due to the fact that they are bound to different cofactors (CoA and ACP). In eukaryotic cells the two opposing pathways are physically separated. The biosynthesis enzymes are found in the cytosol and the β-oxidation enzymes are confined to mitochondria and peroxisomes.

D. Transport of Fatty Acyl CoA into Mitochondria

Long-chain fatty acyl CoA formed in the cytosol cannot diffuse across the inner mitochondrial membrane into the mitochondrial matrix where the reactions of β-oxidation occur in mammals. A transport system, called the carnitine shuttle system, actively transports fatty acids into mitochondria (Figure 16.24). In the cytosol, the acyl group of

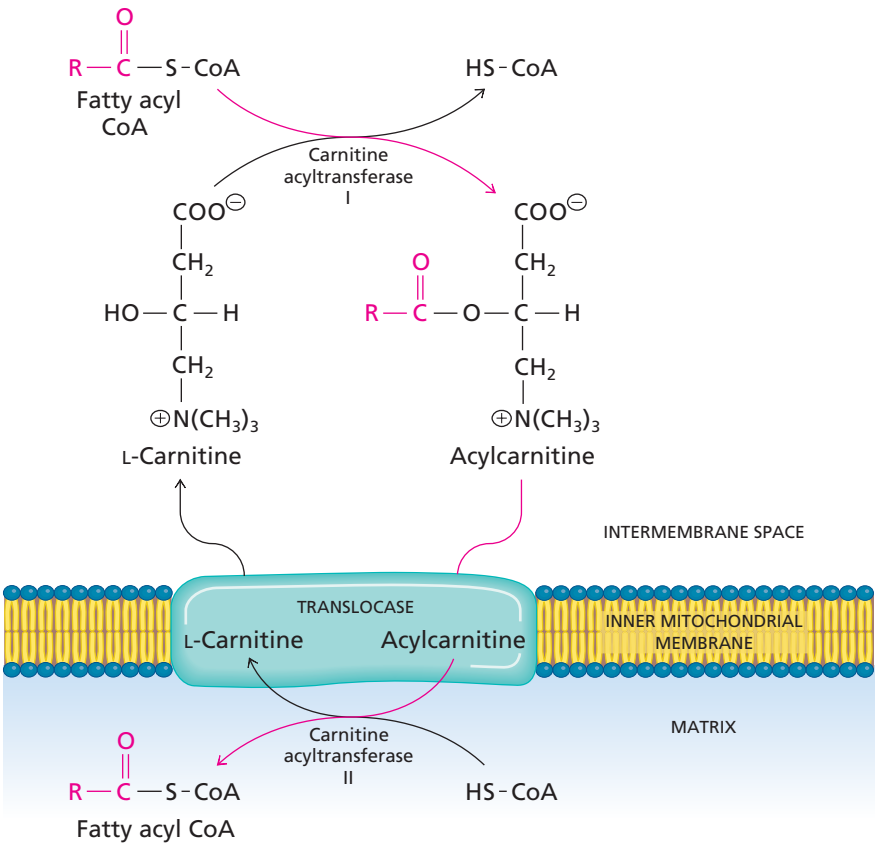
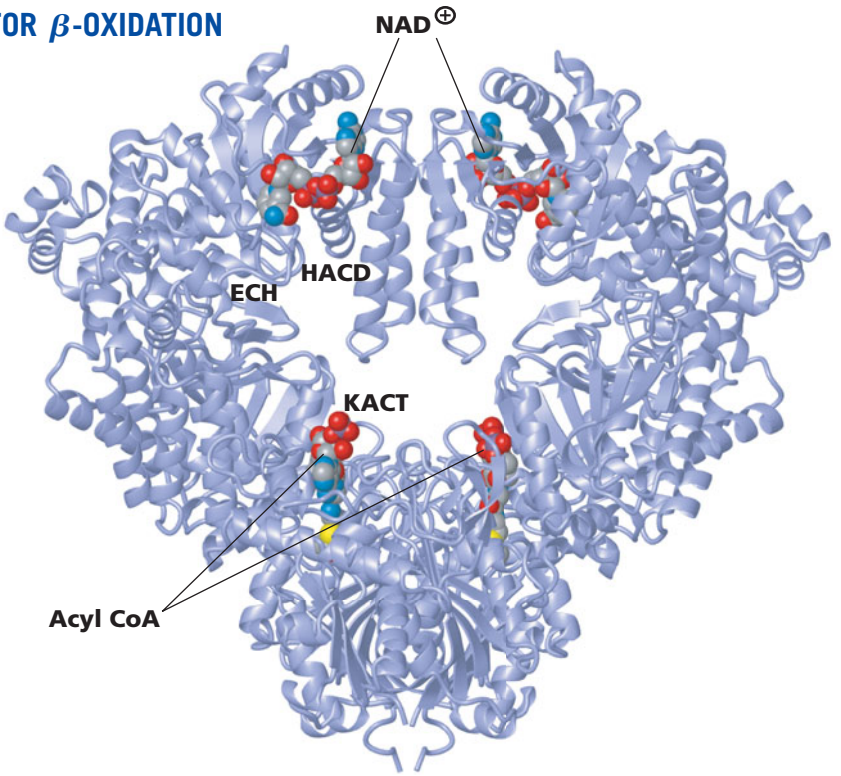


Figure 16.24 ►
Carnitine shuttle system for transporting fatty acyl CoA into the mitochondrial matrix. The path of the acyl group is traced in red.

BOX 16.4 A TRIFUNCTIONAL ENZYME FOR β-OXIDATION

Many species contain a trifunctional enzyme for β-oxidation. The 2-enoyl-CoA hydratase (ECH) and L-3-hydroxyacyl-CoA dehydrogenase (HACD) activities are located on a single polypeptide chain (α subunit). The 3-ketoacyl-CoA thiolase (KACT) activity is localized to the β subunit and the two subunits combine to form a protein with $\alpha_2\beta_2$ quaternary structure.

The structure of a bacterial enzyme is shown in the figure. During β-oxidation the product of the first reaction, *trans*-2-enoyl CoA, binds to the ECH site of the trifunctional enzyme. The substrate then undergoes the next three reactions within the cavity formed by the ECH, HACD, and KACT active sites in each half of the dimer. The two intermediates in the pathway are not released during these reactions because they are bound by their CoA termini. This is an example of metabolic channeling by a multienzyme complex.



▲ **Structure of the fatty acid β-oxidation multienzyme complex from the bacterium *Pseudomonas fragi*.** In this structure a molecule of acyl CoA is bound at each of the KACT sites. [PDB 1WDK]

fatty acyl CoA is transferred to the hydroxyl group of carnitine to form acylcarnitine in a reaction catalyzed by carnitine acyltransferase I, also called carnitine palmitoyltransferase I (CPTI). The enzyme is associated with the outer membrane of the mitochondria.

This reaction is a key site for regulation of the oxidation of intracellular fatty acids. The acyl ester acylcarnitine is a "high energy" molecule with a free energy of hydrolysis similar to that of a thiol ester. Acylcarnitine then enters the mitochondrial matrix in exchange for free carnitine via the carnitine:acylcarnitine translocase. In the mitochondrial matrix, the isozyme carnitine acyltransferase II catalyzes the reverse of the reaction catalyzed by carnitine acyltransferase I. The effect of the carnitine shuttle system is to remove fatty acyl CoA from the cytosol and regenerate fatty acyl CoA in the mitochondrial matrix.

The carnitine shuttle system is not used in most eukaryotes since fatty acid oxidation takes place in the peroxisomes. Fatty acids are transported into peroxisomes by a different mechanisms of course, no transport mechanism is required in prokaryotes since all these reactions take place in the cytoplasm.

E. ATP Generation from Fatty Acid Oxidation

The complete oxidation of fatty acids supplies more energy than the oxidation of an equivalent amount of glucose. As is the case in glycolysis, the energy yield of fatty acid oxidation can be estimated from the total theoretical yield of ATP (Section 13.5). As an example, let's consider the balanced equation for the complete oxidation of one molecule of stearate (C_{18}) by eight cycles of β-oxidation. Stearate is converted to stearoyl CoA at a cost of two ATP equivalents and the oxidation of steroyl CoA yields acetyl CoA and the reduced coenzymes QH_2 and NADH.

$$\text{Stearoyl CoA} + 8\ \text{HS–CoA} + 8\ \text{Q} + 8\ \text{NAD}^{\oplus} \longrightarrow 9\ \text{Acetyl CoA} + 8\ \text{QH}_2 + 8\ \text{NADH} + 8\ \text{H}^{\oplus} \qquad (16.8)$$

KEY CONCEPT

Unlike the pathways for gluconeogenesis and glycolysis, the pathways for the synthesis and degradation of fatty acids are completely different.

In Section 16.7D we compare the cost of fatty acid synthesis to the energy recovered in β-oxidation.

We can calculate the theoretical yield of 9 molecules of acetyl CoA by assuming that they enter the citric acid cycle where they are completely oxidized to CO_2. These reactions produce 10 ATP equivalents for each molecule of acetyl CoA. The net yield from oxidation of stearate is 120 ATP equivalents.

Eight cycles of β-oxidation yield

8 QH_2	$\approx$	12 ATP
8 NADH	$\approx$	20 ATP
9 molecules of acetyl CoA	$\approx$	90 ATP
activation of stearate	$\approx$	−2 ATP
Total	=	120 ATP

By comparison, the oxidation of glucose to CO_2 and water yields approximately 32 ATP molecules. Since stearate has 18 carbons and glucose has only six carbons, we normalize the yield of ATP from glucose by comparing the oxidation of three molecules of glucose: $3 \times 32 = 96$ ATP. This theoretical ATP yield is only 80% of the value for stearate. Fatty acids provide more energy per carbon atom than carbohydrates because carbohydrates are already partially oxidized. Furthermore, because fatty acid moieties are hydrophobic, they can be stored in large quantities as triacylglycerols without large amounts of bound water, as are found with carbohydrates. Anhydrous storage allows far more energy to be stored per gram.

We can also calculate the cost of synthesizing stearate in order to compare it to the energy recovered during β-oxidation. For this calculation we need to know the cost of synthesizing acetyl CoA from CO_2. This value (17 ATP equivalents) is obtained from the reactions of CO_2 fixation in plants (Section 15.4C).

8 acetyl CoA → 8 malonyl ACP		$\approx$	8 ATP
8 synthesis steps	16 NADPH	$\approx$	40 ATP
9 acetyl CoA	9 × 17	$\approx$	153 ATP
Total		=	201 ATP

The energy recovered in the degradation of stearate is about 60% (120/201) of the total theoretical energy required for its synthesis. This is a typical example of biochemical efficiency.

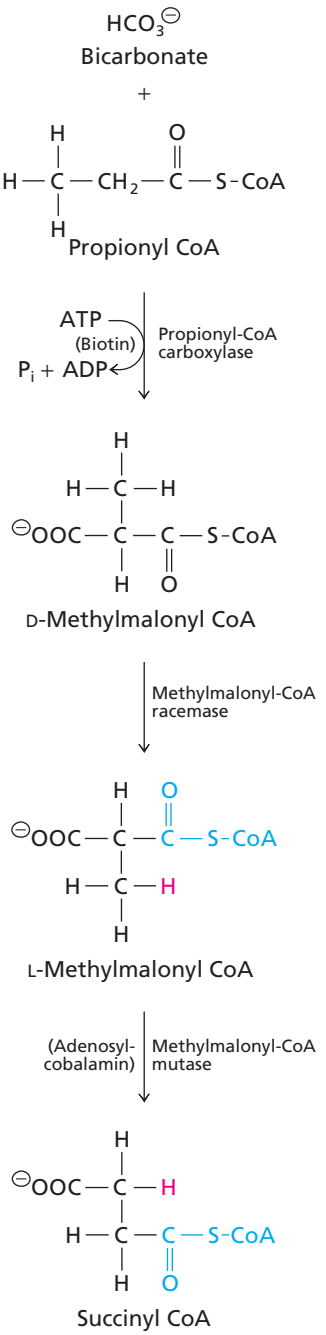
F. β-Oxidation of Odd-Chain and Unsaturated Fatty Acids

Most fatty acids have an even number of carbon atoms. Odd-chain fatty acids are synthesized by bacteria and by some other organisms. Odd-chain fatty acids are oxidized by the same sequence of reactions as even-chain fatty acids except that the product of the final thiolytic cleavage is propionyl CoA (CoA with a C_3 acyl group) rather than acetyl CoA (CoA with a C_2 acyl group). In mammals, propionyl CoA can be converted to succinyl CoA in a three step pathway (Figure 16.25).

The first reaction is catalyzed by propionyl-CoA carboxylase, a biotin-dependent enzyme that incorporates bicarbonate into propionyl CoA to produce D-methylmalonyl CoA. Methylmalonyl-CoA racemase catalyzes the conversion of D-methylmalonyl CoA to its L isomer. Finally, methylmalonyl-CoA mutase catalyzes the formation of succinyl CoA.

Methylmalonyl-CoA mutase is one of the few enzymes that require adenosylcobalamin as a cofactor. We learned in Section 7.12 that adenosylcobalamin-dependent enzymes catalyze intramolecular rearrangements in which a hydrogen atom and a substituent on an adjacent carbon atom exchange places. In the reaction catalyzed by methylmalonyl-CoA mutase, the —C(O)—S-CoA group exchanges with a hydrogen atom of a methyl group (Figure 7.28).

The succinyl CoA molecule formed by the action of methylmalonyl-CoA mutase is metabolized to oxaloacetate. Since oxaloacetate is a substrate for gluconeogenesis, the



▲ **Figure 16.25**
Conversion of propionyl CoA to succinyl CoA.

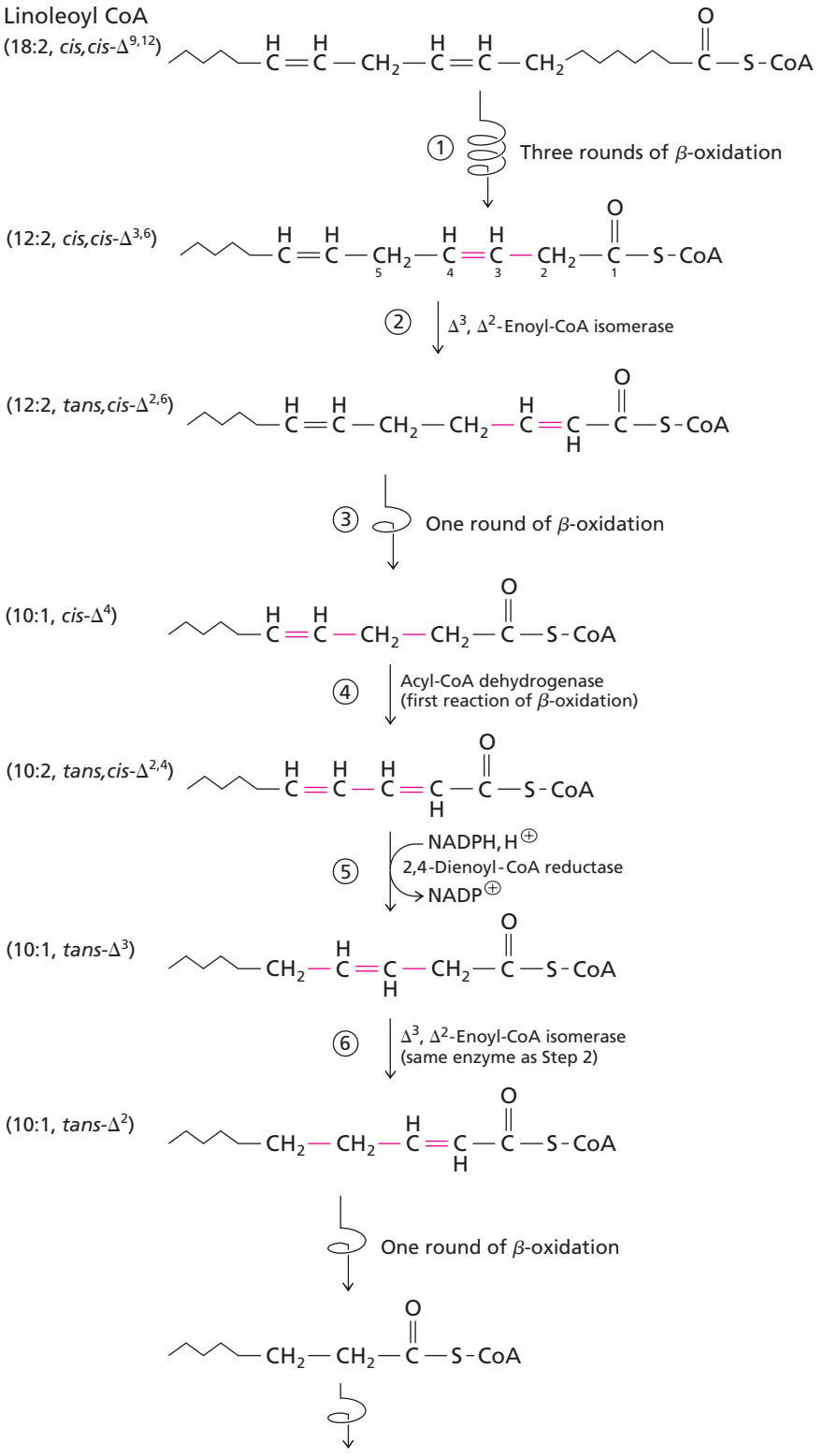


Figure 16.26 ►
Oxidation of linoleoyl CoA. Oxidation requires two enzymes: enoyl-CoA isomerase and 2,4-dienoyl-CoA reductase—in addition to the enzymes of the β-oxidation pathway.

propionyl group derived from the β-oxidation of an odd-chain fatty acid can be converted to glucose.

The oxidation of unsaturated fatty acids requires two enzymes in addition to those usually needed for the oxidation of saturated fatty acids. The oxidation of the Coenzyme A derivative of linoleate (18:2 *cis,cis* $\Delta^{9,12}$-octadecadienoate) illustrates the modified pathway (Figure 16.26).

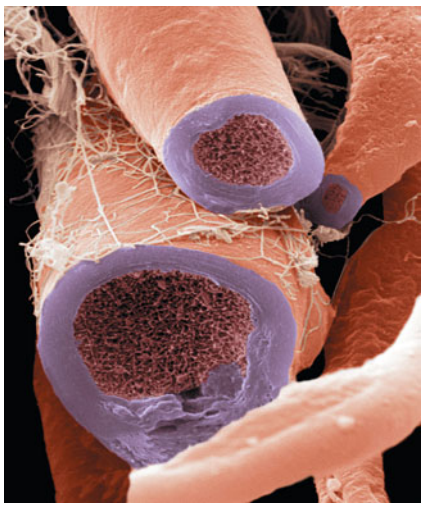


▲ **The camel's hump stores fat for energy production when food is scarce.** The hump of the camel contains fat that is used to supply energy. It does not store water. The ability of camels to go for long periods of time without water is due to completely different adaptations having nothing to do with fat metabolism. The camel shown here is the Arabian camel or dromedary, *Camelus dromedarius*.

Like all polyunsaturated fatty acids linoleoyl CoA has both odd-numbered and even-numbered double bonds (its double bonds are separated by a methylene group). Unsaturated fatty acids are normal substrates for the enzymes of the β-oxidation pathway until an odd-numbered double bond of the shortened fatty acid chain interferes with catalysis. In this example, three rounds of β-oxidation convert linoleoyl CoA to the C_{12} molecule 12:2 *cis*,*cis*-$\Delta^{3,6}$-dienoyl CoA (step 1). This molecule has a *cis*-3,4 double bond rather than the usual *trans*-2,3 double bond that would be produced during β-oxidation of saturated fatty acids. The *cis*-3,4 intermediate is not a substrate for 2-enoyl-CoA hydratase since the normal β-oxidation enzyme is specific for *trans* acyl CoAs and, in addition, the double bond is in the wrong position for hydration.

The inappropriate double bond is rearranged from Δ^3 to Δ^2 to produce the C_{12} molecule 12:2 *trans*,*cis*-$\Delta^{2,6}$-dienoyl CoA in a reaction catalyzed by Δ^3,Δ^2-enoyl-CoA isomerase (step 2). This product can re-enter the β-oxidation pathway and another round of β-oxidation can be completed resulting in the C_{10} molecule 10:1 *cis*-Δ^4-enoyl CoA (step 3). The first enzyme of the β-oxidation pathway, acyl-CoA dehydrogenase, acts on this compound, producing the C_{10} molecule 10:2 *trans*,*cis*-$\Delta^{2,4}$-dienoyl CoA. This resonance-stabilized diene resists hydration. 2,4-Dienoyl-CoA reductase catalyzes the NADPH-dependent reduction of the diene (step 5) to produce a C_{10} molecule with a single double bond (10:1 *trans*-Δ^3-enoyl CoA). This product (like the substrate in step 2) is acted on by Δ^3,Δ^2-enoyl-CoA isomerase to produce a compound that continues through the β-oxidation pathway. Note that the isomerase can convert both *cis*-Δ^3 and *trans*-Δ^3 double bonds to the *trans*-Δ^2 intermediate.

The oxidation of a monounsaturated fatty acid with a *cis* double bond at an odd-numbered carbon (e.g., oleate) requires the activity of the isomerase but not the reductase, in addition to the enzymes of β-oxidation. Oleoyl (18:1 *cis*-Δ^3) CoA undergoes three cycles of β-oxidation, forming three molecules of acetyl CoA and the CoA ester of the (12:1 *cis*-Δ^3) acid. Δ^3,Δ^2-Enoyl-CoA isomerase then catalyzes conversion of the 12-carbon enoyl CoA to a 12-carbon *trans*-Δ^2 enoyl CoA, which can undergo β-oxidation.



▲ **Myelin sheath.** These nerve fibers are coated with several layers of myelin membranes (colored purple) forming a protective sheath around the axons. Plasmalogens are important components of myelin membranes. The symptoms of multiple sclerosis (MS) are caused by degradation of myelin in the brain and spinal cord leading to loss of motor control.

16.8 Eukaryotic Lipids Are Made at a Variety of Sites

Eukaryotic cells are highly compartmentalized. The compartments can have quite different functions, and their surrounding membranes can have quite distinct phospholipid and fatty acyl constituents. Most lipid biosynthesis in eukaryotic cells occurs in the endoplasmic reticulum. Phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol, and phosphatidylserine, for example, are all synthesized in the ER. The biosynthesis enzymes are membrane bound with their active sites oriented toward the cytosol so that they have access to the water-soluble cytosolic compounds. The major phospholipids are incorporated into the ER membrane. From there they are transported to other membranes in the cell in vesicles that travel between the endoplasmic reticulum and Golgi apparatus and between the Golgi apparatus and various membrane target sites. Soluble transport proteins also participate in carrying phospholipids and cholesterol to other membranes.

Although the endoplasmic reticulum is the principal site of lipid metabolism in the cell, there are also lipid-metabolizing enzymes at other locations. For instance, membrane lipids can be tailored to give the lipid profile characteristic of individual cellular organelles. In the plasma membrane, acyltransferase activities catalyze the acylation of lysophospholipids. Mitochondria have the enzyme phosphatidylserine decarboxylase that catalyzes the conversion of phosphatidylserine to phosphatidylethanolamine. Mitochondria also contain the enzymes responsible for the synthesis of diphosphatidylglycerol (cardiolipin, Table 9.2), a molecule found uniquely in the inner membrane of the mitochondrion. Lysosomes possess various hydrolases that degrade phospholipids and sphingolipids. Peroxisomes possess enzymes involved in the early stages of ether

lipid synthesis. Defects in peroxisomal formation can lead to poor plasmalogen synthesis, with potentially fatal consequences.

The tissues of the central nervous system are especially prone to damage. In those tissues plasmalogens constitute a substantial portion of the lipids of the myelin sheath. Often, different subcellular locations have a different set of enzymes (isozymes) responsible for the biosynthesis of different, segregated pools of lipids, with each pool having its own biological function.

16.9 Lipid Metabolism is Regulated by Hormones in Mammals

Fatty acids are no longer oxidized in mitochondria when the energy supply is sufficient to meet the immediate needs of an organism. Instead, they are transported to adipose tissue where they are stored for future use when energy is needed (e.g., lack of food). This aspect of lipid metabolism is similar to the strategy in carbohydrate metabolism where excess glucose is stored in specialized cells as glycogen (animals) or starch (plants).

The mobilization and storage of lipids requires communication between different tissues. Hormones that circulate in the blood are ideally suited to act as signals between cells. Lipid metabolism must be coordinated with carbohydrate metabolism, so it's not surprising that the same hormones also affect the synthesis, degradation, and storage of carbohydrates.

Glucagon, epinephrine, and insulin are the principal hormonal regulators of fatty acid metabolism. Glucagon and epinephrine are present in high concentrations in the fasted state and insulin is present in high concentrations in the fed state. The concentration of circulating glucose must be maintained within fairly narrow limits at all times. In the fasted state, carbohydrate stores become depleted and synthesis of carbohydrates must occur to maintain the level of glucose in the blood. To further relieve pressure on the limited supply of glucose, fatty acids are mobilized to serve as fuel, and many tissues undergo regulatory transitions that decrease their use of carbohydrates and increase their use of fatty acids. The opposite occurs in the fed state when carbohydrates are used as fuel and precursors for fatty acid synthesis.

The key regulatory enzyme for fatty acid synthesis is acetyl-CoA carboxylase. High insulin levels after a meal inhibit the hydrolysis of stored triacylglycerols and stimulate the formation of malonyl CoA by acetyl-CoA carboxylase. Malonyl CoA allosterically inhibits carnitine acyltransferase I. As a result, fatty acids remain in the cytosol rather than being transported into mitochondria for oxidation. Regulation of fatty acid synthesis and degradation is reciprocal, with increased metabolism by one pathway balanced by decreased activity in the opposing pathway. In animals this regulation is achieved by hormones that indirectly affect the activities of the enzymes.

Hormone signaling pathways are described in Section 9.12.

Triacylglycerols are delivered to adipose tissue in the form of lipoproteins that circulate in blood plasma (Section 16.10B). When they arrive at adipose tissue the triglycerols are hydrolyzed to release fatty acids and glycerol that are then taken up by adipocytes. Hydrolysis is catalyzed by lipoprotein lipase (LPL), an extracellular enzyme bound to endothelial cells of the capillaries of adipose tissue. Following entry into adipocytes, the fatty acids are re-esterified for storage as triacylglycerols.

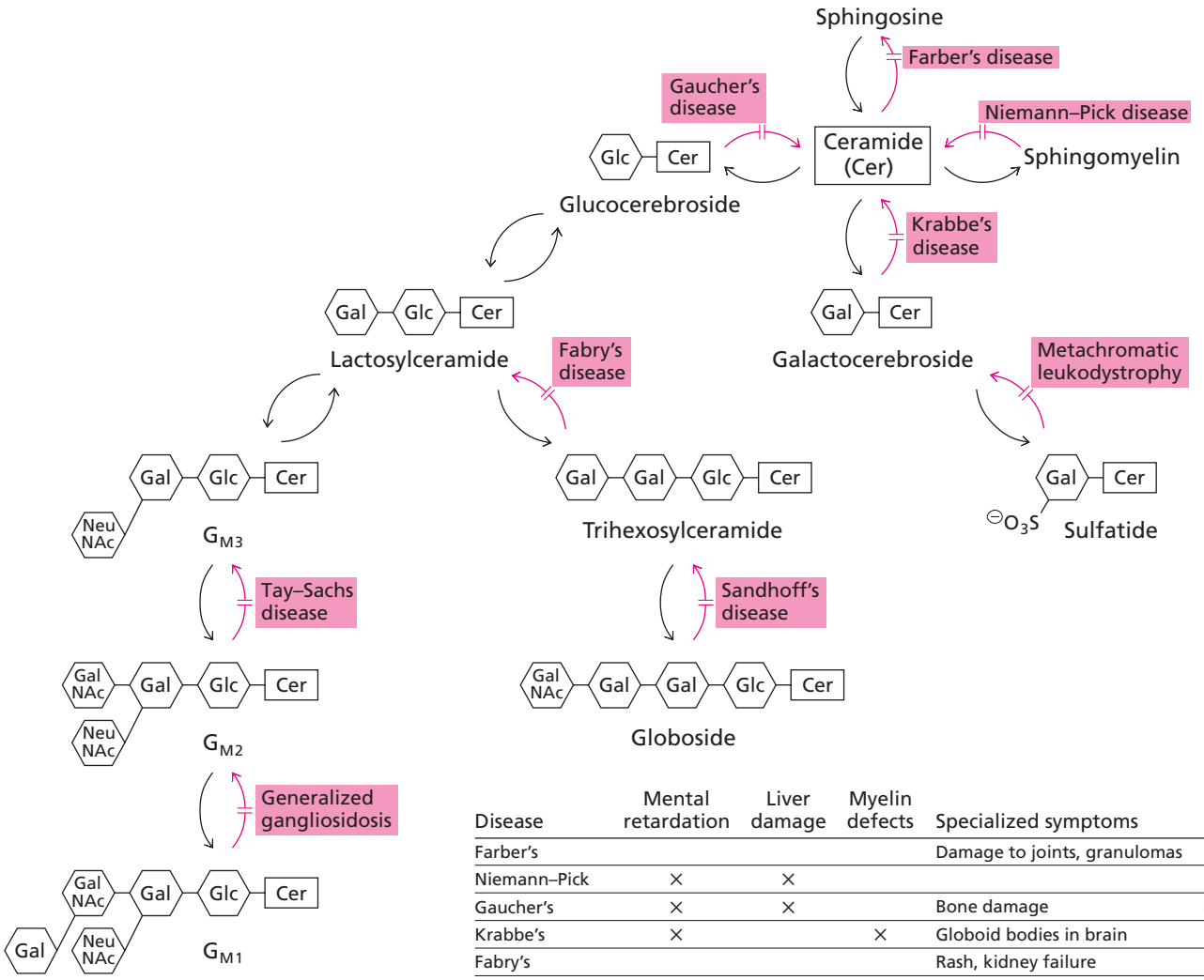
Subsequent mobilization, or release, of fatty acids from adipocytes depends on metabolic needs. A hormone-sensitive lipase in adipocytes catalyzes the hydrolysis of triacylglycerols to free fatty acids and monoacylglycerols. Although hormone-sensitive lipase can also catalyze the conversion of monoacylglycerols to glycerol and free fatty acids, a more specific and more active monoacylglycerol lipase probably accounts for most of this catalytic activity.

The hydrolysis of triacylglycerols is inhibited in the fed state by high concentrations of insulin. When carbohydrate stores are depleted and insulin concentrations are low, an increased concentration of epinephrine stimulates triacylglycerol hydrolysis. Epinephrine binds to the β-adrenergic receptors of adipocytes leading to activation of the

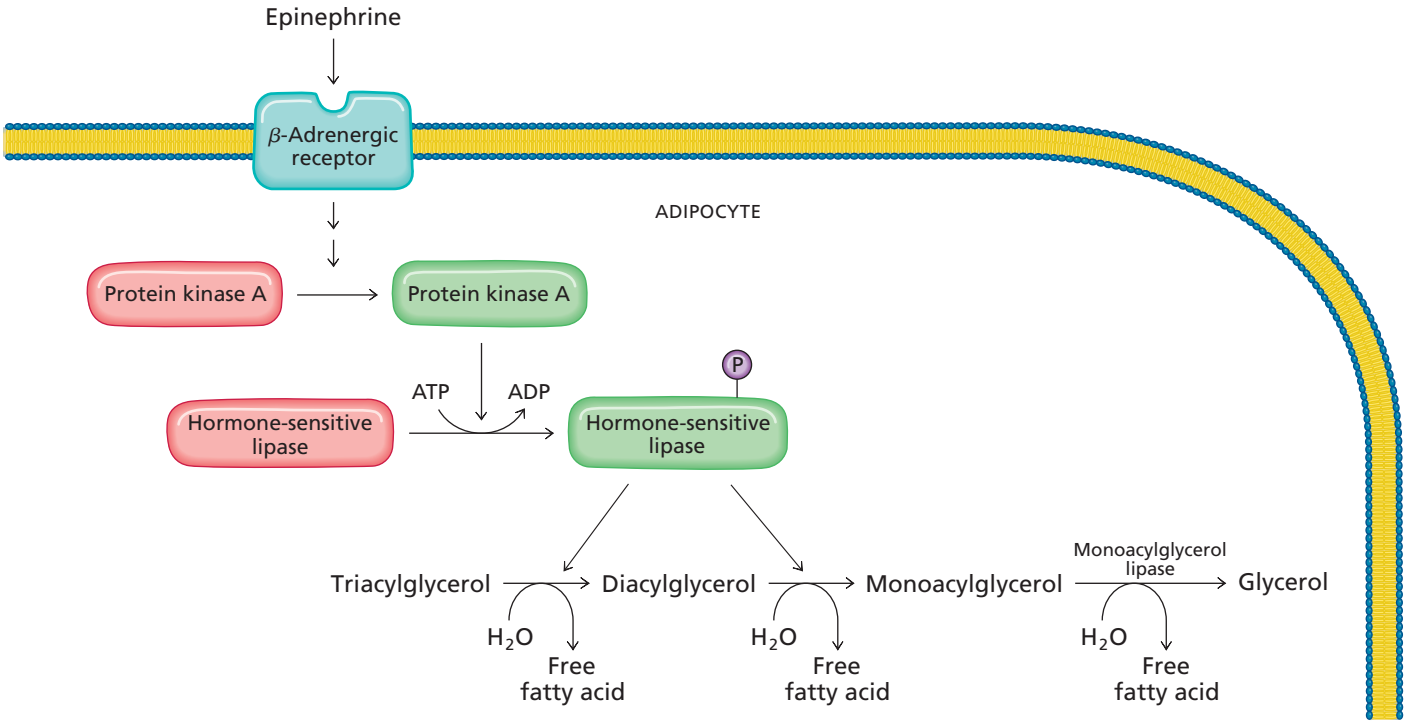
BOX 16.5 LYSOSOMAL STORAGE DISEASES

There are no metabolic diseases associated with defects in the sphingolipid biosynthesis pathways. It is likely that mutations in the genes for biosynthesis enzymes are lethal since sphingolipids are essential membrane components. In contrast, defects in the sphingolipid degradation pathway can have serious clinical consequences. Sphingolipid catabolism is largely carried out in the lysosomes of cells. Lysosomes contain a variety of glycosidases that catalyze the stepwise hydrolytic removal of sugars from the oligosaccharide chains of sphingolipids. There are certain inborn errors of metabolism in which a genetic defect leads to a deficiency in a particular degradative lysosomal enzyme resulting in lysosomal storage diseases. The accumulation of nondegradable lipid by-products can cause lysosomes to swell leading to cellular and ultimately tissue enlargement. This is particularly deleterious in central nervous tissue that has little room for expansion. Swollen lysosomes accumulate in the cell bodies of nerve cells and lead to neuronal death, possibly by leakage of lysosomal enzymes into the cell. As a result, blindness, mental retardation, and death can occur. In Tay–Sachs disease, for instance, there is a deficiency in hexosaminidase A, which catalyzes the removal of *N*-acetylgalactosamine from the oligosaccharide chain of gangliosides. If removal of this sugar does not occur, the disassembly of gangliosides is blocked, leading to a buildup of the nondegradable by-product, ganglioside G_{M2}. (The complete structure of ganglioside G_{M2} is shown in Figure 9.12.)

Schematic pathways for the formation and degradation of a variety of sphingolipids are shown in the accompanying figure. A number of defects in sphingolipid metabolism, whose clinical manifestations are termed *sphingolipidoses*, are identified there.



Disease	Mental retardation	Liver damage	Myelin defects	Specialized symptoms	Fatal
Farber's				Damage to joints, granulomas	×
Niemann–Pick	×	×			×
Gaucher's	×	×		Bone damage	Frequently
Krabbe's	×		×	Globoid bodies in brain	
Fabry's				Rash, kidney failure	
Metachromatic leukodystrophy	×		×	Paralysis, dementia	
Tay–Sachs	×			Blindness, seizures	×
Sandhoff's	×			Same as Tay–Sachs; progresses more rapidly	×
Generalized gangliosidosis	×	×		Bone damage	×



▲ **Figure 16.27**
Triacylglycerol degradation in adipocytes. Epinephrine initiates the activation of protein kinase A, which catalyzes the phosphorylation and activation of hormone-sensitive lipase. The lipase catalyzes the hydrolysis of triacylglycerols to monoacylglycerols and free fatty acids. The hydrolysis of monoacylglycerols is catalyzed by monoacylglycerol lipase.

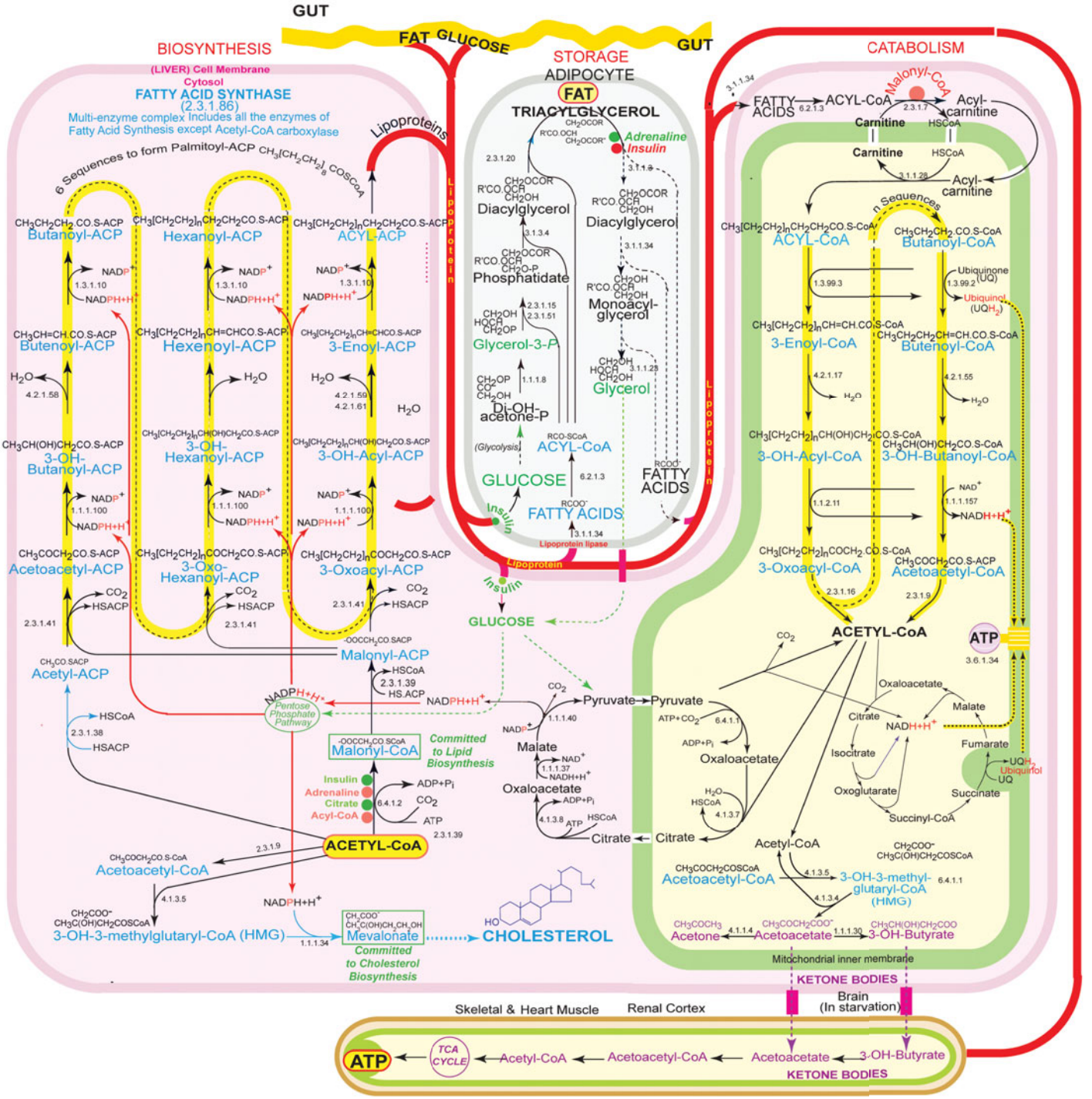
cAMP-dependent protein kinase A. Protein kinase A catalyzes the phosphorylation and activation of hormone-sensitive lipase (Figure 16.27).

Glycerol and free fatty acids diffuse through the adipocyte plasma membrane and enter the bloodstream. Glycerol is metabolized by the liver, where most of it is converted to glucose via gluconeogenesis. Free fatty acids are poorly soluble in aqueous solution and travel through blood bound to serum albumin (Section 16.9C). Fatty acids are carried to tissues such as heart, skeletal muscle, and liver, where they are oxidized in mitochondria to release energy. Fatty acids are a major source of energy during the fasting state (e.g., while we sleep).

At the same time, an increase in glucagon levels inactivates acetyl-CoA carboxylase, the enzyme that catalyzes the synthesis of malonyl CoA in the liver. The result is increased transport of fatty acids into mitochondria and greater flux through the β-oxidation pathway. The high concentrations of acetyl CoA and NADH that are produced by fatty acid oxidation decrease glucose and pyruvate oxidation by inhibiting the pyruvate dehydrogenase complex. Thus, not only are fatty acid oxidation and storage reciprocally regulated but fatty acid metabolism is also regulated so that storage is favored in times of plenty (such as immediately after feeding) and fatty acid oxidation proceeds when glucose must be spared.

Citrate—a precursor of cytosolic acetyl CoA—activates acetyl-CoA carboxylase *in vitro*, but the physiological relevance of this activation has not been fully established. Acetyl-CoA carboxylase is inhibited by fatty acyl CoA. The ability of fatty acid derivatives to regulate acetyl-CoA carboxylase is physiologically appropriate; an increased concentration of fatty acids causes a decrease in the rate of the first committed step of fatty acid synthesis. Acetyl CoA-carboxylase activity is also under hormonal control. Glucagon stimulates phosphorylation and concomitant inactivation of the enzyme in the liver, and epinephrine stimulates its inactivation by phosphorylation in adipocytes. Several protein kinases can catalyze phosphorylation and thus inhibition of acetyl-CoA carboxylase. The action of AMP activated protein kinase inactivates both fatty acid synthesis (by inhibiting the acetyl-CoA carboxylase step) and steroid synthesis in the presence of a high AMP/ATP ratio.

LIPID METABOLISM



ANABOLISM $8CH_3CO.SCoA + 14NADPH + 14H^+ + 7ATP \longrightarrow CH_3[CH_2CH_2]_7CO.SCoA + 14NADP^+ + 7HSCoA + 7ADP + 7P_i$

Acetyl-CoA → Palmitoyl-CoA

CATABOLISM $CH_3[CH_2CH_2]_7CO.SCoA + 7Ubiquinone + 7H_2O + 7NAD^+ + 7HSCoA \longrightarrow 8CH_3CO.SCoA + 7Ubiquinol + 7NADH + 7H^+$

Palmitoyl-CoA → Acetyl-CoA

COMPLETE (AEROBIC) OXIDATION OF PALMITOYL CoA

$CH_3[CH_2CH_2]_7CO.SCoA + 23O_2 + \sim(106\ ADP + 106P_i) \longrightarrow 16CO_2 + 119\ H_2O + HSCoA + \sim 106ATP$

This is a fascinating equation which explains how some animals, such as camels and polar bears can survive in the most adverse environments They can use fat ,not only as the sole source of energy, but also of water. The killer whale cannot utilise sea-water but creates its own from fat

ENZYMES

1.1.1.8	Glycerol-3-P-dehydrogenase	1.3.1.10	Enoyl-[ACP]-reductase	2.3.1.51	1-Acylglycerol-3-*P* O-acyl transferase	4.1.3.5	OH-Methylglutaryl-CoA synthase
1.1.1.34	HMG-CoA reductase	1.3.99.2	Butyryl-CoA dehydrogenase			4.1.3.7	Citrate synthase
1.1.1.35	3-OH-acyl-CoA dehydrogenase	1.3.99.3	Acyl-CoA dehydrogenase	3.1.1.3	Triacylglycerol lipase	4.1.3.8	ATP Citrate lyase
1.1.1.37	Malatedehydrogenase	2.3.1.7	Carnitine-O-acyltransferase	3.1.1.23	Acylglycerol lipase	4.2.1.17	Enoyl-CoA hydratase
1.1.1.40	Malatedehydrogenase (oxaloacetate	2.3.1.9	Acetyl-CoA-C-acetyl transferase	3.1.1.28	Acylcarnitine hydrolase	4.2.1.55	3-OH-Butyryl-CoA dehydratase
1.1.1.100	3-Oxoacyl-[ACP]	2.3.1.15	Glycerol-3-*P* O-acyl transferase	3.1.1.34	Lipoprotein lipase	4.2.1.58	Crotonyl-[ACP] hydratase
1.1.1.157	3-OH-butyryl-CoA	2.3.1.16	Acetyl-CoA C-acyltransferase	3.1.3.4	Phosphatidate phosphatase	4.2.1.59	3-OH-octanoyl-[ACP] dehydratase
1.1.1.211	Long-chain 3-OH-acyl-CoA	2.3.1.20	DiacylglycerolO-acyltransferase	4.1.1.4	Acetoacetate decarboxylase	4.2.1.61	3-OH-Palmitoyl-[ACP] dehydratase
1.2.4.1	Pyruvatedehydrogenase	2.3.1.38	[ACP] S-acyl transferase	4.1.1.9	Malonyl-CoA decarboxylase	6.2.1.3	Long-chain-fatty.-acid-CoA ligase
		2.3.1.39	[ACP] S-malonyl transferase	4.1.3.4	OH-Methylglutaryl-CoA lyase	6.4.1.1	Pyruvate carboxylase
						6.4.1.2	Acetyl-CoA carboxylase

16.10 Absorption and Mobilization of Fuel Lipids

The fatty acids and glycerol that mammals use as metabolic fuels are obtained from triacylglycerols in the diet and from adipocytes. The fats stored in adipocytes include fats synthesized from the catabolism of carbohydrates and amino acids. Free fatty acids occur only in trace amounts in cells—this is fortunate because, as anions, they are detergents and at high concentrations could disrupt cell membranes. We begin our study of lipid metabolism by examining the dietary uptake, transport, and mobilization of fatty acids in mammals.

A. Absorption of Dietary Lipids

Most lipids in the diets of mammals are triacylglycerols with smaller amounts of phospholipids and cholesterol. The digestion of dietary lipids occurs mainly in the small intestine, where suspended fat particles are coated with bile salts (Figure 16.28). Bile salts are amphipathic cholesterol derivatives synthesized in the liver, collected in the gallbladder, and secreted into the lumen of the intestine. Micelles of bile salts solubilize fatty acids and monoacylglycerols so that they can diffuse to and be absorbed by the cells of the intestinal wall. Lipids are transported through the body as complexes of lipid and protein known as lipoproteins.

Triacylglycerols are broken down in the small intestine by the action of lipases. These enzymes are synthesized as zymogens in the pancreas and secreted into the small intestine where they are activated. Pancreatic lipase catalyzes hydrolysis of the primary esters (at C-1 and C-3) of triacylglycerols releasing fatty acids and generating monoacylglycerols (Figure 16.29). A small protein called colipase helps bind the water-soluble lipase to the lipid substrates. Colipase also activates lipase by holding it in a conformation with an open active site. The fatty acids derived from dietary triacylglycerols are primarily long chain molecules.

Most of these bile salts recirculate through the lower parts of the small intestine, the hepatic portal blood, and then the liver. Bile salts circulate through the liver and intestine several times during the digestion of a single meal. Fatty acids are converted to fatty acyl CoA molecules within the intestinal cells. Three of these molecules can combine with glycerol, or two with a monoacylglycerol, to form a triacylglycerol. As described below, these water-insoluble triacylglycerols combine with cholesterol and specific proteins to form chylomicrons for transport to other tissues.

The fate of dietary phospholipids is similar to that of triacylglycerols. Pancreatic phospholipases secreted into the intestine catalyze the hydrolysis of phospholipids (Figure 9.8), which aggregate in micelles. The major phospholipase in the pancreatic secretion is phospholipase A_2, which catalyzes hydrolysis of the ester bond at C-2 of a glycerophospholipid to form a lysophosphoglyceride and a fatty acid (Figure 16.30). A model of phospholipase A_2 with a lipid substrate is shown in Figure 16.31. Lysophosphoglycerides are absorbed by the intestine and re-esterified to glycerophospholipids in intestinal cells.

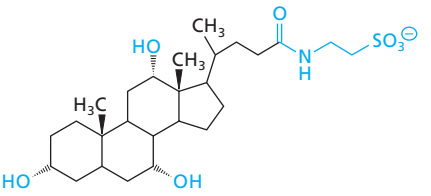
Lysophosphoglycerides are normally present in cells only at low concentrations. High concentrations can disrupt cellular membranes by acting as detergents. This occurs, for example, when snake venom phospholipase A_2 acts on phospholipids in red blood cells, causing lysis of erythrocyte membranes. This is probably what killed Cleopatra.

Unlike other types of dietary lipids, most dietary cholesterol is unesterified. Dietary cholesteryl esters are hydrolyzed in the lumen of the intestine by the action of an esterase. Free cholesterol, which is insoluble in water, is solubilized by bile-salt micelles for absorption. Most cholesterol reacts with acyl CoA to form cholesteryl esters (Figure 9.16) in the intestinal cells.

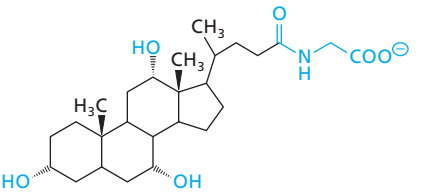
B. Lipoproteins

Triacylglycerols, cholesterol, and cholesteryl esters cannot be transported in blood or lymph as free molecules because they are insoluble in water. Instead, these lipids assemble with phospholipids and amphipathic lipid-binding proteins to form spherical

◀ **IUMBM–Nicholson metabolic chart for lipid metabolism in mammals.**
Designed by Donald Nicholson ©2002 IUBMB.

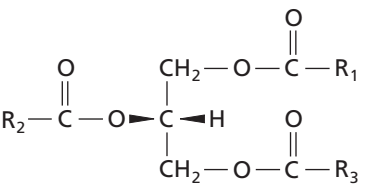


Taurocholate

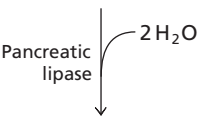


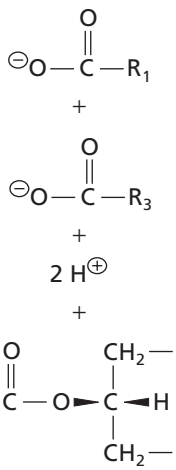
Glycocholate

▲ **Figure 16.28**
Bile salts. The cholesterol derivatives taurocholate and glycocholate are the most abundant bile salts in humans. Bile salts are amphipathic: the hydrophilic parts are shown in blue, and the hydrophobic parts are shown in black.



Triacylglycerol

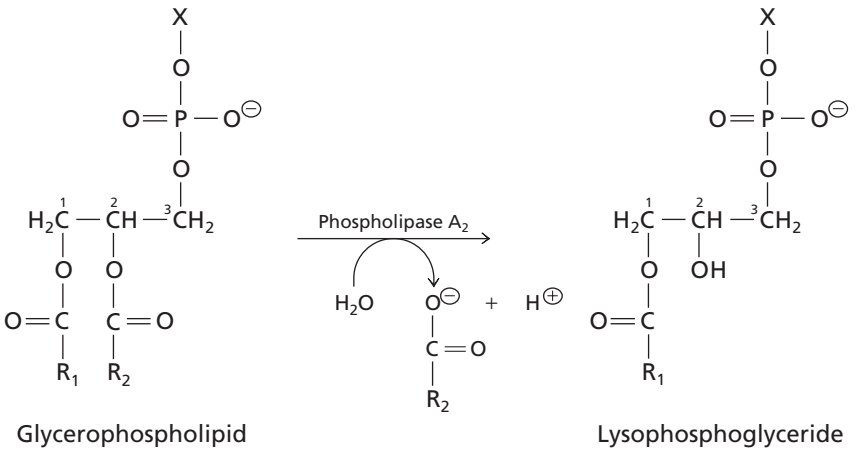


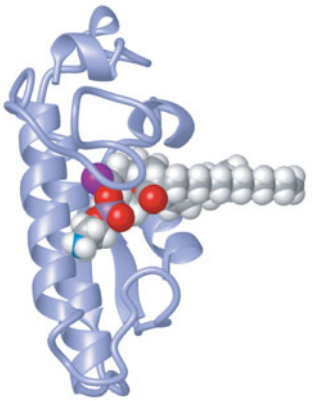


2-Monoacylglycerol

▲ **Figure 16.29**
Action of pancreatic lipase. Removal of the C-1 and C-3 acyl chains produces free fatty acids and a 2-monoacylglycerol. The intermediates, 1,2- and 2,3-diacylglycerol, are not shown.

Figure 16.30 ►
Action of phospholipase A_2. X represents a polar head group. R_1 and R_2 are long hydrophobic chains, making up much of the phospholipid molecule.





▲ Figure 16.31
Structure of phospholipase A_2 from cobra venom. Phospholipase A_2 catalyzes the hydrolysis of phospholipids at lipid–water interfaces. The model shows how a phospholipid substrate (dimyristoyl phosphatidylethanolamine, space-filling model) can fit into the active site of the water-soluble enzyme. A calcium ion (purple) in the active site probably helps bind the anionic head group. About half of the hydrophobic portion of the lipid would be buried in the lipid aggregate. Mammalian phospholipases are structurally similar to the venom enzyme. [PDB 1POB].

macromolecular particles known as lipoproteins. A lipoprotein has a hydrophobic core containing triacylglycerols and cholesteryl esters and a hydrophilic surface consisting of a layer of amphipathic molecules such as cholesterol, phospholipids, and proteins (Figure 16.32).

The largest lipoproteins are chylomicrons that deliver triacylglycerols and cholesterol from the intestine via the lymph and blood to tissues such as muscle (for oxidation) and adipose tissue (for storage) (Figure 16.33). Chylomicrons are present in blood only after a meal. The cholesterol-rich remnants of chylomicrons—having lost most of their triacylglycerol—deliver cholesterol to the liver. Liver cells are responsible for synthesizing most of the newly synthesized cholesterol that enters the bloodstream but almost all cell types make cholesterol for internal use. Lipoproteins deliver both dietary and liver-derived cholesterol to the rest of the body's cells. Cholesterol biosynthesis is regulated by hormones and by the levels of cholesterol in the blood.

Blood plasma contains several other types of lipoproteins. They are classified according to their relative densities and types of lipid (Table 16.1). Since proteins are more dense than lipids, the greater the protein content of a lipoprotein, the greater its density. Very low density lipoproteins (VLDLs) consist of approximately 98% lipid and only 2% protein. VLDLs are formed in the liver and carry lipids synthesized in the liver, or not needed by the liver, to other tissues such as adipose tissue. Lipases within capillaries of muscle and adipose tissue degrade VLDLs and chylomicrons. When VLDLs give up triacylglycerols to tissue cells their lipid content decreases and their remnants

BOX 16.6 EXTRA VIRGIN OLIVE OIL

Olive oil contains mostly triacylglycerols. If it has been produced by crushing olives with no additional chemical treatment, then it is called *virgin* olive oil according to the International Olive Oil Council (IOOC).

The quality of olive oil is often determined by the presence of free fatty acids that form when triacylglycerols break down during production. Virgin olive oil should have less than 2% free fatty acids (acidity) and *extra virgin* olive oil has less than 0.8% free fatty acids (acidity).



► **Extra virgin olive oil.** Extra virgin olive oil has less than 0.8% free fatty acids. http://www.examiner.com/fountain-of-youth-in-atlanta/extra-virgin-olive-oil-benefits

are degraded to intermediate density lipoproteins (IDLs). Of the IDLs formed during the breakdown of VLDLs, some are taken up by the liver and others are degraded to low density lipoproteins (LDLs). LDLs are enriched in cholesterol and cholesteryl esters and deliver these lipids to peripheral tissues. High density lipoproteins (HDLs) are formed as protein-rich particles in blood plasma. They pick up cholesterol from peripheral tissues, chylomicrons, and VLDL remnants and convert it into cholesterol esters. HDLs transport cholesterol and cholesteryl esters back to the liver. Cholesteryl esters from HDLs can be picked up by IDLs, which become LDLs.

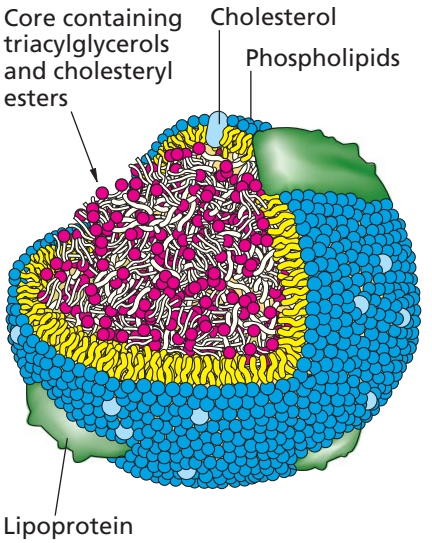
Large lipoprotein particles contain a number of different lipid binding proteins. These are often called apolipoproteins—the "apo-" prefix usually refers to polypeptides that bind to a tightly associated cofactor as described in Chapter 7. Two of these apolipoproteins are large, hydrophobic, monomeric proteins. ApoB-100 (M_r 513,000) is firmly bound to the outer layer of VLDLs, IDLs, and LDLs. The smaller apolipoproteins of VLDLs and IDLs are weakly bound and most dissociate during lipoprotein degradation, leaving apoB-100 as the major protein component of LDLs. ApoB-48 (M_r 241,000), which is present only in chylomicrons, is identical in primary structure to the N-terminal 48% of apoB-100.

ApoB-100 and apoB-48 form much of the amphipathic crust or shell over the hydrophobic lipoprotein core of their respective lipoproteins. ApoB-100 is the protein that attaches LDL to its cell surface receptor; apoB-48 lacks this property. The other apolipoproteins are smaller than apoB-48. They have a variety of functions, including modulating the activity of certain enzymes involved in lipid mobilization and interacting with cell surface receptors.

Cholesterol, an essential component of eukaryotic cell membranes, is delivered to peripheral tissues by LDLs. The lipoprotein particles bind to the LDL receptor on the cell surface. A complex between LDL and its receptor enters the cell by endocytosis and fuses with a lysosome. Lysosomal lipases and proteases degrade the LDL releasing cholesterol that is then incorporated into cell membranes or stored as cholesteryl esters. An abundance of intracellular cholesterol suppresses synthesis of HMG-CoA reductase, a key enzyme in the biosynthesis of cholesterol and it also inhibits synthesis of the LDL receptor. Individuals lacking LDL receptors suffer from familial hypercholesterolemia, a disease in which cholesterol accumulates in the blood and is deposited in the skin and in arteries. Such patients die of heart disease at an early age.

HDLs remove cholesterol from plasma and from cells of nonhepatic tissues returning it to the liver. They bind to a receptor called SR-B1 at the liver surface and transfer cholesterol and cholesterol esters into liver cells. The lipid depleted HDL particles return to the plasma. In the liver, the cholesterol can be converted to bile salts that are secreted into the gallbladder.

The buildup of lipid deposits in the arteries (atherosclerosis) is associated with increased risk of coronary heart disease that can lead to a heart attack. High levels of LDL ("bad" cholesterol) *increase* the chance of developing atherosclerosis. High levels of



▲ **Figure 16.32**
Structure of a lipoprotein. A core of neutral lipids, including triacylglycerols and cholesteryl esters, is coated with phospholipids in which apolipoproteins and cholesterol are embedded.



▲ **Chylomicrons.**

BOX 16.7 LIPOPROTEIN LIPASE AND CORONARY HEART DISEASE

Lipoprotein lipase (Section 16.9) is the enzyme that releases fatty acid from the triacylglcerols in lipoproteins. It plays an important role in clearing triacylglycerols from the blood plasma. High concentrations of triacylglycerols are associated with coronary heart disease.

The human population contains several variants (mutations) of the lipoprotein lipase (LPL) gene. Some of these are associated with decreased LPL activity. One example is the D9N variant where an asparagine residue substitutes for the normal aspartate residue at position 9. Individuals who carry this variant are more likely to suffer from coronary heart disease due to the buildup of triacyglycerol-containing lipoproteins in the blood plasma.

In the S447X variant a normal serine codon is mutated to a stop codon (X) at position 447. The result is a truncated protein that is shorter than the normal protein. About 17% of the population carries at least one copy of this variant gene and 1% of the population is homozygous for this variant. The S447X enzyme is more active than the wild-type enzyme and this results in lower triacylglycerol levels in plasma. Males (but not females) who carry this variant are less likely to suffer heart attacks. This is an example of a beneficial allele that has arisen in the human population.

[Online Mendelian Inheritance in Man (OMIM) MIM=609708]

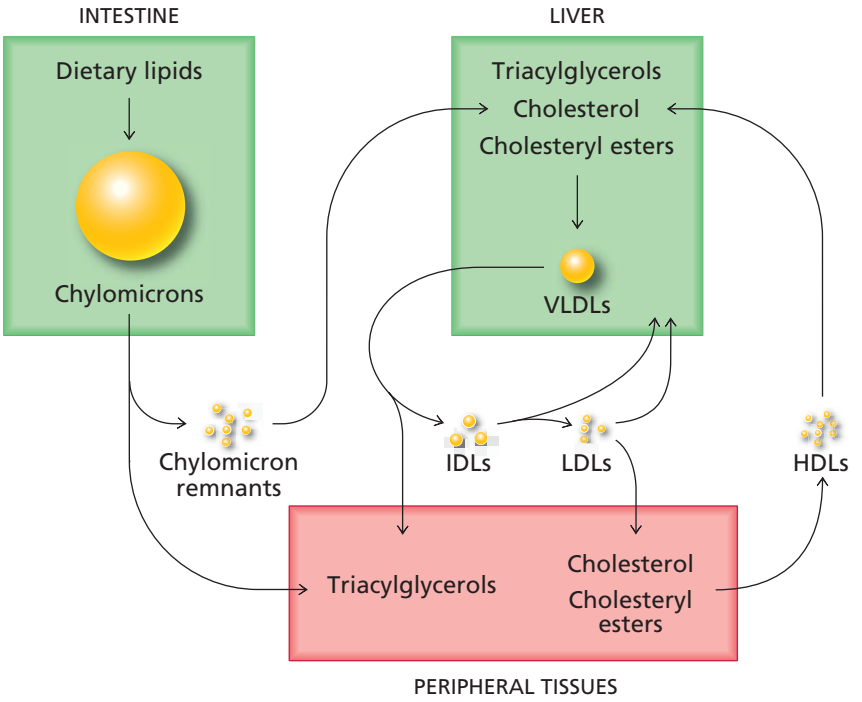
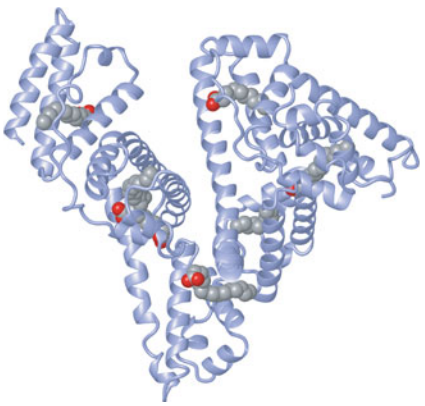


Figure 16.33 ►
Summary of lipoprotein metabolism. Chylomicrons formed in intestinal cells carry dietary triacylglycerols to peripheral tissues, including muscle and adipose tissue. Chylomicron remnants deliver cholesteryl esters to the liver. VLDLs assemble in the liver and carry endogenous lipids to peripheral tissues. When VLDLs are degraded (via IDLs), they pick up cholesterol and cholesteryl esters from HDLs and become LDLs, which carry cholesterol to nonhepatic tissues. HDLs deliver cholesterol from peripheral tissues to the liver.



▲ Figure 16.34
Human serum albumin. Seven bound molecules of palmitate are shown. [PDB 1E7H]

HDL ("good" cholesterol), on the other hand, are correlated with a *decrease* in the risk of having a heart attack. Statins (Box 16.4) block synthesis of cholesterol in the liver and lower LDL levels.

C. Serum Albumin

In addition to complex lipids such as cholesterol and triacylglycerols, free fatty acids are also transported in blood plasma. Fatty acids bind to serum albumin, an abundant plasma protein. This protein, especially the bovine version (bovine serum albumin, BSA) has been intensely studied for over 40 years. Recently, the structure of human serum albumin (HSA) in association with free fatty acids of various chain lengths (Figure 16.34) has been solved by X-ray crystallography.

HSA belongs to the all-α category of tertiary structures (Section 4.7, Figure 4.24a). There are seven distinct binding sites for palmitic acid (16:0) and other medium and long chain fatty acids. In most cases, the carboxylate end of the fatty acids interacts with the side chains of basic amino acid residues and the methylene tails fit into hydrophobic pockets that can accommodate chains of 10–18 carbons. HSA also binds many important drugs that are only sparingly soluble in water.

16.11 Ketone Bodies Are Fuel Molecules

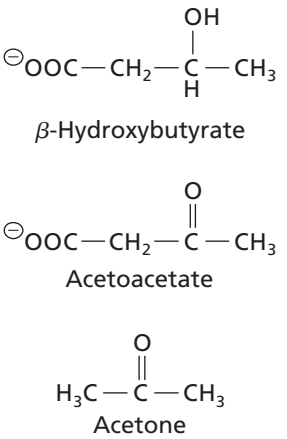
Most acetyl CoA produced in the liver from fatty acid oxidation is routed to the citric acid cycle but some of it can follow an alternate pathway. During periods of fasting, glycolysis is decreased and the gluconeogenic pathway is active. Under these conditions the

Table 16.1 Lipoproteins in human plasma

	Chylomicrons	VLDLs	IDLs	LDLs	HDLs
Molecular weight × 10^{-6}	>400	10–80	5–10	2.3	0.18–0.36
Density (g cm^{-3})	<0.95	0.95–1.006	1.006–1.019	1.019–1.063	1.063–1.210
Chemical composition (%)					
Protein	2	10	18	25	33
Triacylglycerol	85	50	31	10	8
Cholesterol	4	22	29	45	30
Phospholipid	9	18	22	20	29

pool of oxaloacetate molecules becomes temporarily depleted. The amount of acetyl CoA from enhanced β-oxidation exceeds the capacity of the citric acid cycle (recall that oxaloacetate reacts with acetyl CoA in the first step of the citric acid cycle). The excess acetyl CoA is used to form ketone bodies—β-hydroxybutyrate, acetoacetate, and acetone. As indicated by their structures (Figure 16.35), not all ketone bodies are ketones. The only quantitatively significant ketone bodies are β-hydroxybutyrate and acetoacetate; small amounts of acetone are produced by the nonenzymatic decarboxylation of acetoacetate, a β-keto acid.

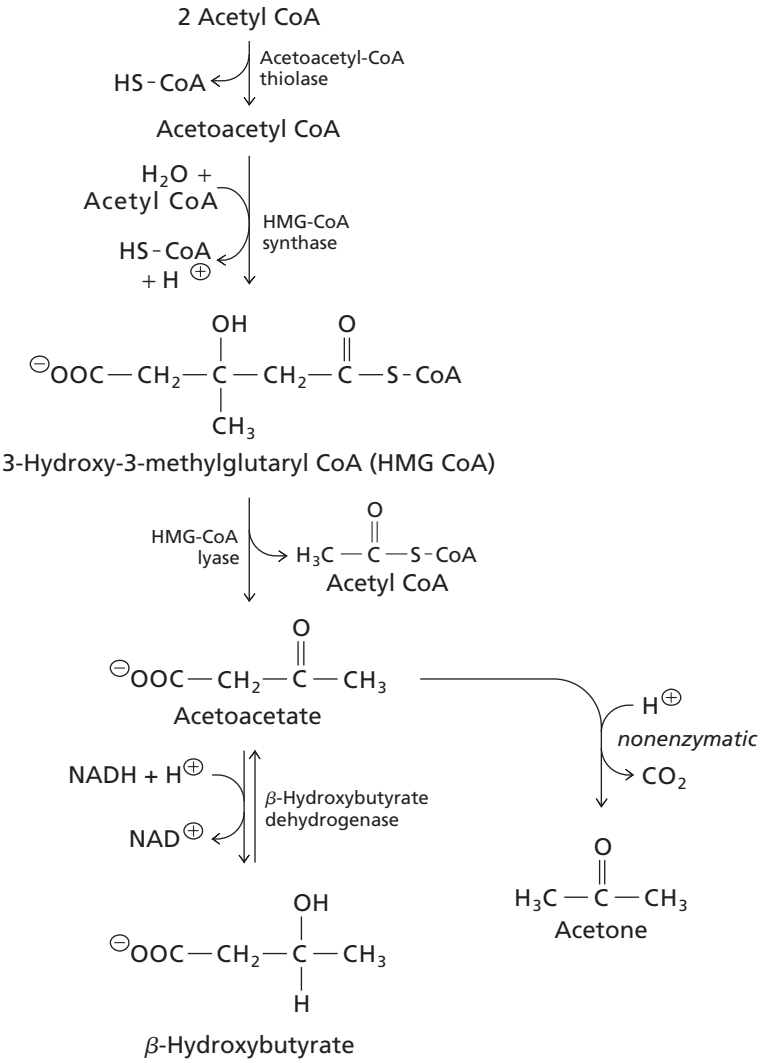
β-Hydroxybutyrate and acetoacetate are fuel molecules. They have less potential metabolic energy than the fatty acids from which they are derived but they make up for this deficiency by serving as "water-soluble lipids" that can be more readily transported in the blood plasma. During starvation, ketone bodies are produced in large amounts becoming substitutes for glucose as the principal fuel for brain cells. Ketone bodies are also metabolized in skeletal muscle and in the intestine during starvation.



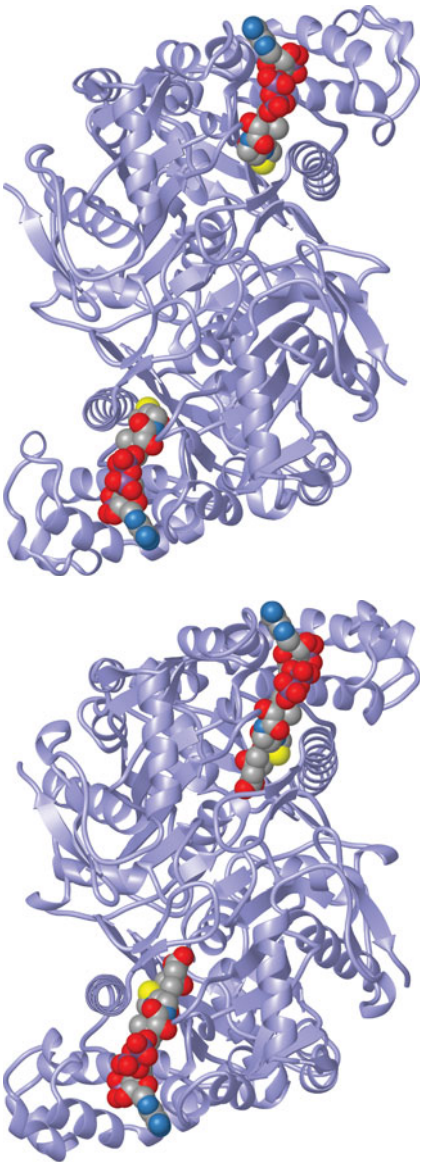
▲ **Figure 16.35**
Ketone bodies.

A. Ketone Bodies Are Synthesized in the Liver

In mammals, ketone bodies are synthesized in the liver and exported for use by other tissues. The pathway for ketone body synthesis is shown in Figure 16.36. First, two molecules of acetyl CoA condense to form acetoacetyl CoA and HS-CoA in a reaction catalyzed by acetoacetyl-CoA thiolase. Subsequently, a third molecule of acetyl CoA is added to acetoacetyl CoA to form 3-hydroxy-3-methylglutaryl CoA (HMG CoA) in a reaction catalyzed by HMG-CoA synthase. These steps are identical to the first two steps in the isopentenyl diphosphate biosynthesis pathway (Figure 16.17). The synthesis of



◀ **Figure 16.36**
Biosynthesis of β-hydroxybutyrate, acetoacetate, and acetone.



▲ **HMG-CoA synthase.** The human (*Homo sapiens*) isozymes are shown with bound HMG CoA. The cytosolic enzyme (top: PDB 2P8U) and the mitochondrial version (bottom: PDB 2WYA) are very similar.

Changes in carbohydrate metabolism during starvation are described in Section 13.10.

ketone bodies takes place in mitochondria but the synthesis of isopentenyl diphosphate (and cholesterol) takes place in the cytosol. Mammals have distinct isozymes of acetoacetyl-CoA thiolase and HMG-CoA synthase in the mitochondria and the cytosol. HMG-CoA synthase is only present in the mitochondria of liver cells and not in the mitochondria of any other cell types.

In the next step, HMG-CoA lyase catalyzes the cleavage of HMG CoA producing acetoacetate and acetyl CoA. HMG-CoA lyase is not present in the cytosol, which is why cytosolic HMG CoA is used exclusively in isopentenyl diphosphate synthesis and no ketone bodies are produced in the cytosol. NADH-dependent reduction of acetoacetate produces β-hydroxybutyrate in a reaction catalyzed by β-hydroxybutyrate dehydrogenase. Both acetoacetate and β-hydroxybutyrate can be transported across the inner mitochondrial membrane and the plasma membrane of liver cells. They enter the blood to be used as fuel by other cells of the body. Small amounts of acetoacetate are nonenzymatically decarboxylated to acetone in the bloodstream.

The main control point for ketogenesis is the mitochondrial isozyme of HMG-CoA synthase provided that fatty acyl CoA and acetyl CoA are available in the mitochondria. Succinyl CoA specifically inhibits this enzyme by covalent modification through succinylation. This is a short-term inactivation since reactivation occurs frequently by spontaneous desuccinylation. Glucagon lowers the amount of succinyl CoA in mitochondria, stimulating ketogenesis. Long-term regulation occurs by modification of gene expression. Starvation increases the level of HMG-CoA synthase (and its mRNA); refeeding or insulin produces a decrease in both activity and mRNA.

B. Ketone Bodies Are Oxidized in Mitochondria

In cells that use them as an energy source, β-hydroxybutyrate and acetoacetate enter mitochondria where they are converted to acetyl CoA that is oxidized by the citric acid cycle. β-Hydroxybutyrate is converted to acetoacetate in a reaction catalyzed by an isozyme of β-hydroxybutyrate dehydrogenase that is distinct from the liver enzyme. Acetoacetate reacts with succinyl CoA to form acetoacetyl CoA in a reaction catalyzed by succinyl-CoA transferase (also called succinyl-CoA:3-ketoacid-CoA transferase; Figure 16.37). Ketone bodies are broken down only in nonhepatic tissues because this transferase is present in all tissues except liver. The succinyl-CoA transferase reaction siphons some succinyl CoA from the citric acid cycle. Energy that would normally be captured as GTP in the substrate-level phosphorylation catalyzed by succinyl-CoA synthetase (Section 13.3#5) is used instead to activate acetoacetate to its CoA ester. Thiolase then catalyzes the conversion of acetoacetyl CoA to two molecules of acetyl CoA that can be oxidized by the citric acid cycle.

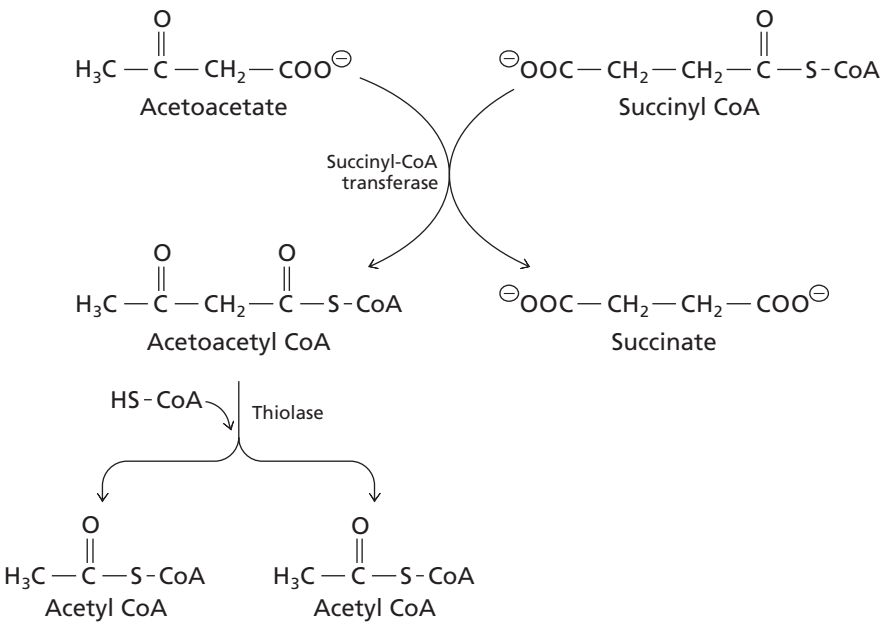


Figure 16.37 ►
Conversion of acetoacetate to acetyl CoA.

BOX 16.8 LIPID METABOLISM IN DIABETES

The breakdown of fats occurs because lipolysis is not inhibited by insulin, and other hormones trigger the release of fatty acids from adipocytes. The large amounts of fatty acids available to the liver lead to excess acetyl CoA that is diverted to form ketone bodies. In Type 2 diabetes (Section 12.7), the accumulation of glucose in the blood is caused mainly by poor uptake of glucose by peripheral tissues. Because obesity strongly predisposes a person to developing Type 2 diabetes, much research is focusing on the role of lipids in decreased insulin sensitivity. It appears that elevated free fatty acids in the blood may interfere with insulin signaling for glucose uptake into tissues.

Individuals suffering from untreated Type 1 diabetes produce large amounts of ketone bodies—more than the peripheral tissues can use. The smell of acetone can be discerned on the breath of diabetics. In fact, the levels of acetoacetic acid and β-hydroxybutyric acid in the blood can be so high that the pH of the serum can be lowered—a life-threatening condition called diabetic ketoacidosis. Type 1 diabetes must be treated with repeated injections of insulin and restricted glucose intake.

Although acute complications are rare in Type 2 diabetes, hyperglycemia can lead to tissue damage, particularly in the eye and the cardiovascular and renal systems. Dietary modifications are often sufficient to control Type 2 diabetes. In addition, oral drugs can increase insulin secretion and potentiate its action at peripheral tissues.

A novel approach for the treatment of Type 2 diabetes may be inhibition of the tyrosine phosphatase PTP-1B. PTP-1B inactivates the insulin receptor by catalyzing the removal of phosphate added to the receptor when insulin binds to it. After insulin injection, mice lacking PTP-1B have increased phosphorylation of insulin receptors in liver and muscle and enhanced sensitivity to insulin. These mice also maintain normal levels of blood glucose after a meal. A surprising observation was that mice lacking PTP-1B could eat a high fat diet yet be resistant to weight gain. PTP-1B may therefore also be a target for the treatment of obesity.

Summary

1. The pathway for fatty acid synthesis begins with synthesis of malonyl CoA in a reaction catalyzed by acetyl CoA-carboxylase. Malonyl CoA is converted to malonyl ACP and one molecule of malonyl ACP condenses with acetyl CoA (or acetyl ACP) to form acetoacetyl ACP.
2. The formation of long-chain fatty acids from a 3-ketoacyl ACP precursor occurs in four stages: reduction, dehydration, further reduction, and condensation . These four stages repeat to form a long-chain fatty acid. Fatty acids with more than 18 carbons and unsaturated fatty acids are produced by additional reactions.
3. Triacylglycerols and glycerophospholipids are derived from phosphatidate. The synthesis of triacylglycerols and neutral phospholipids proceeds via a 1,2-diacylglycerol intermediate. Acidic phospholipids are synthesized via a CDP–diacylglycerol intermediate.
4. Many eicosanoids are derived from arachidonate. The cyclooxygenase pathway leads to prostacyclin, prostaglandins, and thromboxane A_2. The products of the lipoxygenase pathway include leukotrienes.
5. Sphingolipids are synthesized from serine and palmitoyl CoA. Reduction, acylation, and oxidation produce ceramide, which can be modified by adding a polar head group and sugar residues.
6. Cholesterol is synthesized from acetyl CoA in a pathway leading to mevalonate and isopentenyl diphosphate. Both cholesterol and isopentenyl diphosphate are precursors of many other compounds.
7. Fatty acids are degraded to acetyl CoA by β-oxidation, the sequential removal of two-carbon fragments. Fatty acids are first activated by esterification to CoA and fatty acyl CoA is oxidized by a repeated series of four enzyme-catalyzed steps: oxidation, hydration, further oxidation, and thiolysis. Fatty acids yield more ATP per gram than glucose.
8. β-Oxidation of odd-chain fatty acids produces acetyl CoA and one molecule of propionyl CoA. The oxidation of most unsaturated fatty acids requires two enzymes, an isomerase and a reductase, in addition to those required for the oxidation of saturated fatty acids.
9. Fatty acid oxidation in animals is regulated by hormones according to the energy needs of the organism.
10. Dietary fat is hydrolyzed in the intestine to fatty acids and monoacylglycerols, which are absorbed. Lipoproteins transport lipids in the blood. In adipocytes, fatty acids are esterified for storage as triacylglycerols. Fatty acids are mobilized by the action of hormone-sensitive lipase.
11. The ketone bodies β-hydroxybutyrate and acetoacetate are water-soluble fuel molecules produced in the liver by the condensation of acetyl-CoA molecules.

Problems

1. (a) Familial hypercholesterolemia is a human genetic disorder in which LDL receptors are defective, leading to very high blood cholesterol levels and severe atherosclerosis at an early age. Explain why this disease results in high blood cholesterol levels.
   (b) Do high blood cholesterol levels affect cellular cholesterol synthesis in individuals with this disease?
   (c) Individuals with Tangier's disease lack the cellular protein ABCl, which is required for cholesterol uptake by HDL. How will this disease affect cholesterol transport?

2. Individuals with abnormally low levels of carnitine in their muscles suffer from muscular weakness during moderate exercise. In addition, their muscles have significantly increased levels of triacylglycerols.

(a) Explain these two effects.

(b) Can these individuals metabolize muscle glycogen aerobically?

3. How many ATP equivalents are generated by the complete oxidation of (a) laurate (dodecanoate) and (b) palmitoleate (*cis*-Δ^9-hexadecenoate)? Assume that the citric acid cycle is functioning.

4. Tetrahydrolipstatin (Orlistat) is a drug treatment for obesity. It is an inhibitor of pancreatic lipase. Suggest a rationale for use of tetrahydrolipstatin to treat obesity.

5. In addition to the enzymes of β-oxidation, what enzymes are necessary to degrade the following fatty acids to acetyl CoA or acetyl CoA and succinyl CoA?

(a) oleate (*cis* $CH_3(CH_2)_7CH{=}CH(CH_2)_7COO^{\ominus}$)

(b) arachidonate (all *cis* $CH_3(CH_2)_4(CH{=}CHCH_2)_4(CH_2)_2COO^{\ominus}$)

(c) *cis* $CH_3(CH_2)_9CH{=}CH(CH_2)_4COO^{\ominus}$)

6. Animals cannot carry out a net conversion of even chain fatty acid carbons to glucose. On the other hand, some of the carbons in odd-chain fatty acids can be gluconeogenic precursors to glucose. Explain.

7. Where is the labeled carbon found when the following molecules are added to a liver homogenate carrying out palmitate synthesis?

(a) $H^{14}CO_3^{\ominus}$

(b) $H_3{}^{14}C-\overset{\overset{O}{\|}}{C}-S-CoA$

8. Triclosan (2,4,4-trichloro-2-hydroxydiphenyl ether) is an effective antimicrobial agent that is used in a wide range of consumer products including soaps, toothpaste, toys, and cutting boards. Triclosan is effective against a broad spectrum of bacteria and mycobacteria and is an inhibitor of type II FAS enoyl acyl carrier protein reductase.

(a) What reaction is catalyzed by enoyl acyl carrier protein reductase?

(b) Why is enoyl acyl carrier protein reductase an appropriate target for antimicrobials?

(c) Suggest a reason why a compound may selectively inhibit fatty acid synthesis in bacteria and not in humans.

9. It has been proposed that malonyl CoA may be one of the signals sent to the brain to decrease the appetite response. When mice are given a derivative of cerulenin (a fungal epoxide) named C75, their appetite is suppressed and they rapidly lose weight. Cerulenin and its derivatives have been shown to be potent inhibitors of fatty acid synthase (FAS). Suggest how C75 might act as a potential weight reduction drug.

10. (a) Draw a general pathway for converting carbohydrates to fatty acids in a liver cell, and indicate which processes occur in the cytosol and which occur in motochondria.

(b) About half the reducing equivalents necessary for fatty acid synthesis are generated by glycolysis. Explain how these reducing equivalents can be used for fatty acid synthesis.

11. (a) Acetyl CoA carboxylase (ACC), a key regulator for fatty acid synthesis, exists in two different interconvertible forms: (1) an active filamentous polymer (dephosphorylated), and (2) an inactive protomeric form (phosphorylated). Citrate and palmitoyl CoA can regulate fatty acid synthesis by preferentially binding tightly to and stabilizing different forms of ACC. Explain how each of these regulator functions by interacting with ACC.

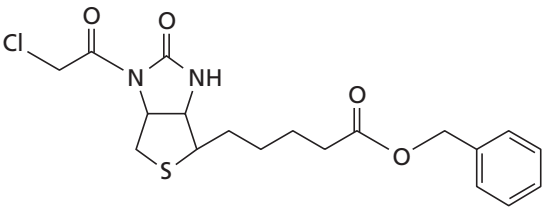
Filamentous polymer (active) $\rightleftharpoons$ Protomer (inactive)

(b) What role do glucagon and epinephrine play in regulating fatty acid synthesis?

12. Obesity is a serious health problem worldwide due in part to increased food intake and reduced physical activity. Obesity is associated with a variety of human disease including Type 2 diabetes and cardiovascular diseases. Selective and specific inhibitors of acetyl-CoA carboxylase have been proposed as potential anti-obesity drugs.

(a) What effect would an inhibitor of acetyl-CoA carboxylase have on fatty acid synthesis and fatty acid oxidation?

(b) One such inhibitor of acetyl-CoA carboxylase is CABI (structure below). What structural feature of CABI makes it a potential acetyl-CoA carboxylase inhibitor? (Levert, K. L., Waldrop, G. L., Stephens, J. M. (2002). *J. Biol. Chem.* A biotin analog inhibits acetyl CoA carboxylase activity and adipogenesis. 277:16347–16350.)



CABI

13. Write the equation for the conversion of eight acetyl CoA molecules to palmitate.

14. (a) In response to tissue damage in such injuries as heart attacks and rheumatoid arthritis, inflammatory cells (e.g., monocytes and neutrophils) invade the injured tissue and promote the synthesis of arachadonic acid. Explain the reason for this response.

(b) The biosynthesis of eicosanoids is affected by nonsteroidal drugs such as aspirin and ibuprofen and by steroidal drugs such as hydrocortisone and prednisone (which inhibit a specific phospholipase). Why do steroidal drugs inhibit the biosynthesis of both prostaglandins and leukotrienes, whereas aspirin-like drugs inhibit the biosynthesis of only prostaglandins?

15. Draw the correct structures of the following complex lipids.

(a) Phosphatidyl glycerol.

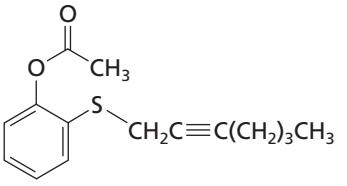
(b) Ethanolamine plasmalogen (1-alkyl-2-glycero-3-phosphoethanolamine).

(c) Glucocerebroside (1-β-D-glucoceramide).

16. Excess dietary fat can be converted to cholesterol in the liver. When palmitate labeled with ^{14}C at every odd-numbered carbon is added to a liver homogenate, where does the label appear in mevalonate?

17. The therapeutic anti-inflammatory effects of aspirin arise from its inhibition of the enzyme cyclooxygenase-2 (COX-2)—involved in the synthesis of prostaglandins, mediators of inflammation, pain, and fever. Aspirin irreversibly inhibits COX-2 by covalently

transferring an acetyl group to a serine residue at the enzyme active site. However, the undesirable side effect of stomach irritation arises from the irreversible inhibition of the related intestinal enzyme cyclooxygenase-1 (COX-1) by aspirin. COX-1 is involved in the synthesis of prostaglandins that regulate secretion of gastric mucin, which protects the stomach from acid. The aspirin analog APHS was synthesized and shown to be 60 times more selective as an inhibitor of COX-2 than of COX-1, suggesting that it could be an anti-inflammatory drug with far less gastrointestinal side effects. Draw the structure of the inactivated COX-2 enzyme-inhibitor complex with APHS. Since aspirin and structural analogs act at the active site of COX enzymes, will they exhibit competitive inhibition patterns?



APHS

Selected Readings

General

Nicholson, D. E. (2001). IUBMB–Nicholson metabolic pathways charts. *Biochem. Mol. Bio. Educ.* 29:42–44.

Vance, J. E., and Vance, D. E., eds. (2008). *Biochemistry of Lipids, Lipoproteins, and Membranes* (Amsterdam: Elsevier Science).

Lipid Synthesis

Athenstaedt, K., and Daum, G. (1999). Phosphatidic acid, a key intermediate in lipid metabolism. *Eur. J. Biochem.* 266:1–16.

Frye, L. L., and Leonard, D. A. (1999). Lanosterol analogs: dual-action inhibitors of cholesterol biosynthesis. *Crit. Rev. Biochem. Mol. Biol.* 34:123–124.

Kent, C. (1995). Eukaryotic phospholipid synthesis. *Annu. Rev. Biochem.* 64:315–343.

Leibundgut, M., Maier, T., Jenni, S., and Ban, N. (2008). The multienzyme architecture of eukaryotic fatty acid synthases. *Curr. Opin. Struct. Biol.* 18:714–726.

Simmons, D. L., Botting, R. M., and Hla, T. (2004). Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol. Rev.* 56:387–437.

Sommerville, C., and Browse, J. (1996). Dissecting desaturation: plants prove advantageous. *Trends in Cell Biol.* 6:148–153.

Wallis, J. G., Watts, J. L., and Browse, J. (2002). Polyunsaturated fatty acid synthesis: what will they think of next? *Trends Biochem. Sci.* 27:467–473.

White, S. W. Zheng, J., Zhang, Y-M., and Rock, C. O. (2005). The structural biology of type II fatty acid biosynthesis. *Ann. Rev. Biochem.* 74:791–831.

Lipid Catabolism

Bartlett, K., and Eaton, S. (2004). Mitochondrial β-oxidation. *Eur. J. Biochem.* 271:462–469.

Candlish, J. (1981). Metabolic water and the camel's hump—a textbook survey. *Biochem Ed.* 9:96–97.

Ishikawa, M., Tsuchiya, D., Oyama, T., Tsunaka, D., and Morikawa, K. (2004). Structural basis for channelling mechanism of a fatty acid β-oxidation multienzyme complex. *EMBO J.* 23:2745–2754.

Kim, J-J., and Battaile, K. P. (2002). Burning fat: the structural basis of fatty acid β-oxidation. *Curr. Opin. Struct. Biol.* 12:721–728.

Toogood. H. S., van Thiel, A., Basran, J., Sutcliffe, M. J., Scrutton, N. S., and Leys, D. (2004). Extensive domain motion and electron transfer in the human electron transferring flavoprotein medium chain acyl-CoA dehydrogenase complex. *J. Biol. Chem.* 279: 32904–32912.

Toogood, H. S., van Thiel, A., Scrutton, N. S., and Leys, D. (2005). Stabilization of non-productive conformations underpins rapid electron transfer to electron-transferring flavoprotein. *J. Biol. Chem.* 280:30361–30366.

Wanders, R. J. A., and Waterman, H. R. (2006). Biochemistry of mammalian peroxisomes revisited. *Ann. Rev. Biochem.* 75:295–332.

Lipoproteins

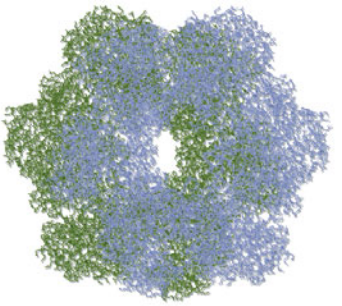
Bhattacharya, A. A., Grüne, T., and Curry, S. (2000). Crystallographic analysis reveals common modes of binding of medium- and long-chain fatty acids to human serum albumin. *J. Mol. Biol.* 303:721–732.

Fidge, N. H. (1999). High density lipoprotein receptors, binding proteins, and ligands. *J. Lipid Res.* 40:187–201.

Gagné, S. E., Larson, M. G., Pimstone, S. N., Schaefer, E. J., Kastelein, J. J. P., Wilson, P. W. F., Ordovas, J. M., and Hayden, M. R. (1999). A common truncation variant of lipoprotein lipase (S447X) confers protection against coronary heart disease: the Framingham Offspring Study. *Clin. Genet.* 55:450–454.

Kreiger, M. (1998). The "best" of cholesterols, the "worst" of cholesterols: a tale of two receptors. *Proc. Natl. Acad. Sci. USA* 95:4077–4080.





Amino Acid Metabolism

We live now in the "Age of Bacteria." Our planet has always been in the "Age of Bacteria," ever since the first fossils—bacteria, of course—were entombed in rocks more than 3 and a half billion years ago. On any possible, reasonable, or fair criterion, bacteria are—and always have been—the dominant forms of life on Earth.

Stephen Jay Gould, (1996), *Full House,* p. 176

Authors writing a chapter on amino acid metabolism have a nearly impossible task. Any description will be incomplete since there are 20 different amino acids and many intermediates in the biosynthesis and degradation pathways. Furthermore, alternate pathways are used by different tissues, organelles, and organisms. Fortunately, metabolic highlights can show the biological rationale of how amino acids are formed and degraded without getting into excessive detail. Here we describe a number of these highlights in order to illustrate the principles and concepts of amino acid metabolism.

The metabolism of amino acids includes hundreds of enzymatic interconversions of small molecules. Many of these reactions involve nitrogen atoms. Some of the intermediates appear in the metabolic pathways described in preceding chapters but many are described here for the first time. Although amino acids from the degradation of proteins can be a source of energy, we are more concerned with their biosynthesis. Life is compromised if all the amino acids are not available at the same time for protein synthesis. We can consider the metabolism of the 20 common amino acids from two points of view: the origins and fates of their nitrogen atoms and the origins and fates of their carbon skeletons.

The abilities of organisms to synthesize amino acids differ widely. A few organisms can assimilate N_2 and simple carbon compounds into amino acids—in other words, they are totally self-supporting for amino acid synthesis. Other species can synthesize the carbon chains of amino acids but require nitrogen in the form of ammonia. We begin this chapter with an overview of the principles of nitrogen metabolism.

Some species cannot synthesize the carbon skeletons of every amino acid. Mammals, for example, can make only about half of the amino acids they require; the rest—called **essential amino acids**—must be obtained from the diet. **Nonessential amino acids** are those that mammals can synthesize in sufficient quantity, provided they receive adequate total dietary protein.

The routes for disposal of the nitrogen-containing waste products of amino acid metabolism also vary among species. For example, excess nitrogen is excreted by aquatic animals as ammonia, by birds and most reptiles as uric acid, and by many other terrestrial

Top: Glutamine synthetase from the bacterium *Salmonella typhimurium*. Twelve identical subunits are arranged with hexagonal symmetry. [PDB 2GLS].

vertebrates as urea. We will end the chapter with a description of the urea cycle, a pathway for elimination of nitrogen in mammals.

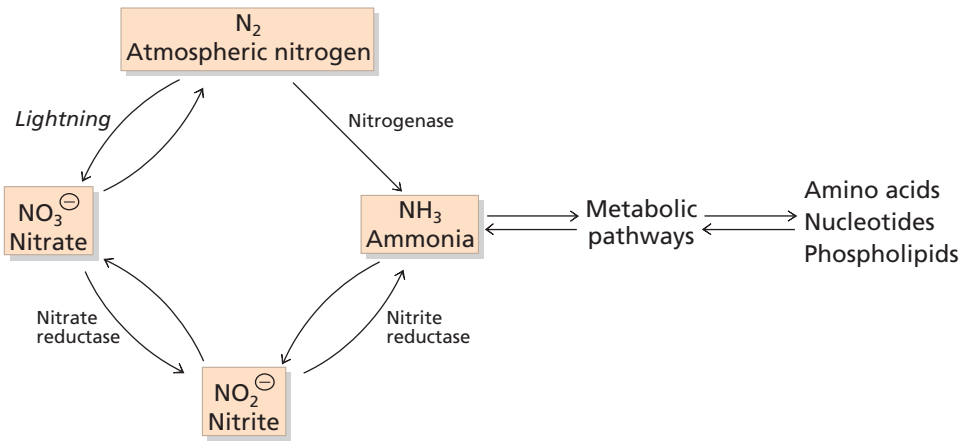
17.1 The Nitrogen Cycle and Nitrogen Fixation

The nitrogen needed for amino acids (and for the heterocyclic bases of nucleotides; Chapter 18) comes from two major sources—nitrogen gas in the atmosphere and nitrate ($NO_3^{\ominus}$) in soil and water. Atmospheric N_2, which constitutes about 80% of the atmosphere, is the ultimate source of biological nitrogen. This molecule can be metabolized, or fixed, by only a few species of bacteria. N_2 and $NO_3^{\ominus}$ must be reduced to ammonia in order to be used in metabolism. The ammonia produced is incorporated into amino acids via glutamate, glutamine, and carbamoyl phosphate.

N_2 is chemically unreactive because of the great strength of the $N \equiv N$ triple bond. Some bacteria have a very specific, sophisticated enzyme, called nitrogenase, that can catalyze the reduction of N_2 to ammonia in a process called **nitrogen fixation**. Ammonia is essential for life and bacteria are the only organisms capable of producing it from atmospheric nitrogen. Half of all biological nitrogen fixation is performed by various species of cyanobacteria in the ocean. The other half comes from soil bacteria.

There are two additional nitrogen-converting processes in addition to biological nitrogen fixation. During lightning storms, high-voltage discharges catalyze the oxidation of N_2 to nitrate and nitrite ($NO_2^{\ominus}$). Nitrogen is converted to ammonia for use in plant fertilizers by an energetically expensive industrial process that requires high temperature and pressure as well as special catalysts to drive the reduction of N_2 by H_2. The availability of biologically useful nitrogen is often a limiting factor for plant growth, and the application of nitrogenous fertilizers is important for obtaining high crop yields. Humans are now responsible for a substantial fraction of the total nitrogen fixation on the planet. Although only a small percentage of the nitrogen undergoing metabolism comes *directly* from nitrogen fixation, this process is the only way that organisms can use the huge pool of atmospheric N_2.

The overall scheme for the interconversion of the major nitrogen-containing compounds is shown in Figure 17.1. The flow of nitrogen from N_2 to nitrogen oxides, ammonia, and nitrogenous biomolecules and then back to N_2 is called the **nitrogen cycle**. Most of the nitrogen shuttles between ammonia and nitrate. Ammonia from decayed organisms is oxidized by soil bacteria to nitrate. This formation of nitrate is called nitrification. Some anaerobic bacteria can reduce nitrate or nitrite to N_2 (denitrification).



▲ **Figure 17.1**
Nitrogen cycle. A few free-living or symbiotic microorganisms can convert N_2 directly to ammonia. Ammonia is incorporated into biomolecules such as amino acids and proteins that subsequently are degraded, re-forming ammonia. Many soil bacteria and plants can carry out the reduction of nitrate to ammonia via nitrite. Several bacteria convert ammonia to nitrite. Others oxidize nitrite to nitrate and some can reduce nitrate to N_2.



▲ **Blooms of *Trichodesmium*.** *Trichodesmium* is one of the main nitrogen-fixing species of cyanobacteria. This large bloom of bacteria formed giant streaks in the ocean off the coast of Australia. The photograph was taken from the space shuttle. The average concentration of nitrogen-fixing bacteria in the ocean is about one million cells per liter.

KEY CONCEPT

Nitrogen is the most abundant gas in the atmosphere but only a few species of bacteria are capable of nitrogen fixation.

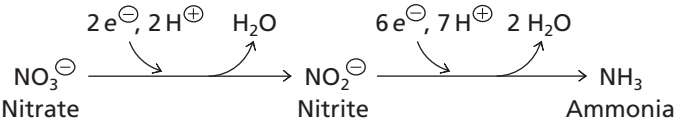


▲ **Lightning.** Lightning causes the conversion of nitrogen gas to nitrates. It is an important source of usable nitrogen for living organisms. This photograph was taken in 1908.



▲ **Figure 17.2**
Nodules on alfalfa roots. Symbiotic bacteria of the genus *Rhizobium* reside in these nodules where they reduce atmospheric nitrogen to ammonia.

Most green plants and some microorganisms contain nitrate reductase and nitrite reductase, enzymes that together catalyze the reduction of nitrogen oxides to ammonia.



(17.1)

This ammonia is used by plants, which supply amino acids to animals. Reduced ferredoxin (formed in the light reactions of photosynthesis, Section 15.2B) is the source of the reducing power in plants and photosynthetic bacteria.

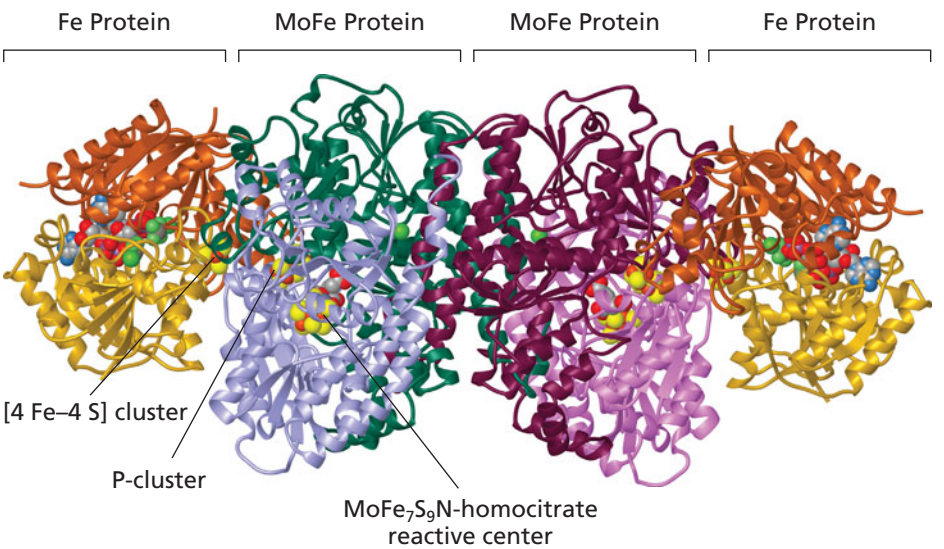
Let's examine the enzymatic reduction of N_2. Most nitrogen fixation in the biosphere is carried out by bacteria that synthesize the enzyme nitrogenase. This multisubunit protein catalyzes the conversion of each molecule of N_2 to two molecules of NH_3. Nitrogenase is present in various species of *Rhizobium* and *Bradyrhizobium* that live symbiotically in root nodules of many leguminous plants, including soybeans, peas, alfalfa, and clover (Figure 17.2). N_2 is also fixed by free-living soil bacteria such as *Agrobacteria, Azotobacter, Klebsiella*, and *Clostridium* and by cyanobacteria (mostly *Trichodesmuim spp.*) found in the ocean. Most plants require a supply of fixed nitrogen from sources such as decayed animal and plant tissue, nitrogen compounds excreted by bacteria, and fertilizers. Vertebrates obtain fixed nitrogen by ingesting plant and animal matter.

Nitrogenase is a protein complex that consists of two different polypeptide subunits forming an $\alpha_2\beta_2$ dimer of dimers (Figure 17.3). The two halves of the complex contain an [8 Fe–7 S] iron–sulfur cluster called the P-cluster. It is near the outer surface of the protein. The reactive center is a complex cluster of molybdenum, iron, and homocitrate [$MoFe_7S_9$-homocitrate]. A single $\alpha\beta$ dimer is called the iron–molybdenum (MoFe) protein.

Electrons are donated to the P-custer by a mobile iron (Fe) protein containing a [4 Fe–4 S] cluster. Fe protein, a homodimer, binds to the ends of MoFe protein near the P-cluster and a single electron is passed from Fe protein to MoFe protein. The reduction of iron in Fe protein is coupled to oxidation of ferredoxin or flavodoxin and each of these reduction reactions requires hydrolysis of two bound ATP molecules. Electrons are passed from Fe protein to the P-cluster to the FeMo-cluster. A total of six electrons are required for conversion of N_2 to $2NH_3$ and these must be passed one at a time from Fe protein as it binds and then dissociates from MoFe protein. An obligatory reduction of 2 $H^{\oplus}$ to H_2 accompanies the reduction of N_2. The overall stoichiometry is

$$N_2 + 8\,H^{\oplus} + 8\,e^{\ominus} + 16\,ATP \rightarrow 2\,NH_3 + H_2 + 16\,ADP + 16\,P_i \qquad (17.2)$$

Figure 17.3 ►
Structure of *Azotobacter vinelandii nitrogenase.* The Fe protein subunits are colored red and orange and the α and β subunits of each half of the MoFe protein are colored blue/green and purple/pink. This structure with bound Fe protein is stabilized by bound transition-state ATP analogs ADP-AlF_4 at the ATP binding sites. [PDB 1N2C]



This is a very expensive reaction in terms of ATP equivalents. It is also a very slow reaction in biochemical terms with a turnover number of only five ammonia atoms produced per second. The slowness of the reaction is due to the fact that eight reduced Fe proteins have to bind and dissociate from the MoFe protein during the conversion of nitrogen to ammonia.

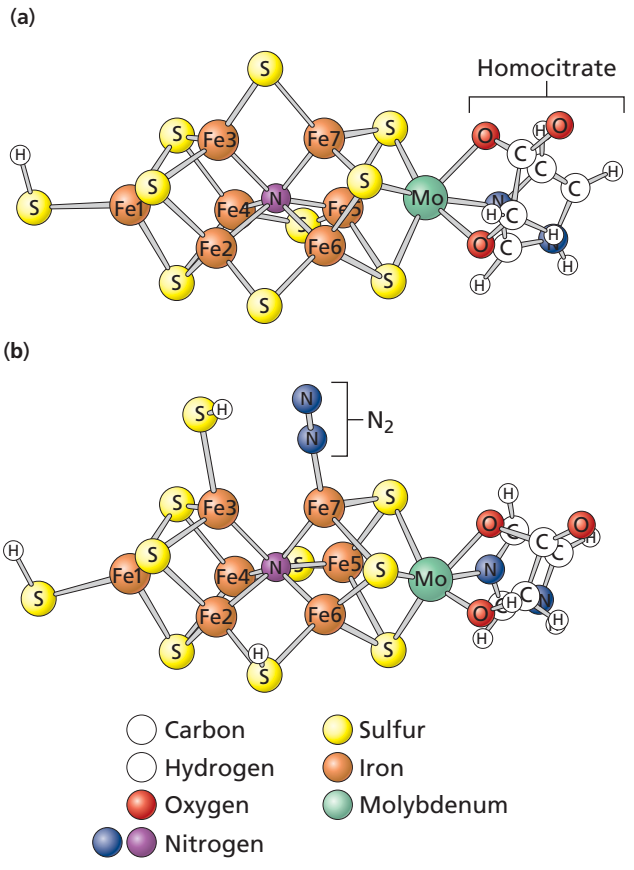
Nitrogenases must be protected from oxygen because the various oxidation–reduction centers are highly susceptible to inactivation by O_2. Strict anaerobes carry out nitrogen fixation in the absence of O_2. Within the root nodules of leguminous plants, the protein leghemoglobin (a homolog of vertebrate myoglobin; Section 4.12) binds O_2 and thereby keeps its concentration sufficiently low in the immediate environment of the nitrogen-fixing enzymes of *Rhizobium*. Nitrogen fixation in cyanobacteria is carried out in specialized cells (heterocysts) whose thick membranes inhibit entry of O_2 (Figure 10.8). In order to obtain the reducing power and ATP required for this process, symbiotic nitrogen-fixing microorganisms rely on nutrients obtained through photosynthesis carried out by the plants with which they are associated.

The actual reduction of nitrogen takes place at the iron–molybdenum–homocitrate cluster in the MoFe protein. This cluster is remarkably complex. It consists of a cage of Fe and S atoms surrounding a central N atom. A single Mo atom is bound to one edge of the Fe—S cage. It is chelated to a single molecule of homocitrate to form a $MoFe_7S_9N$·homocitrate cluster (Figure 17.4).

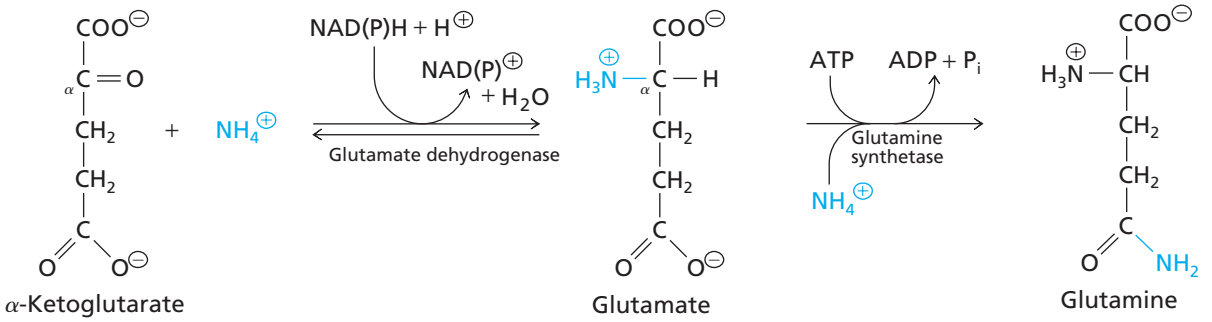
The detailed reaction mechanism of nitrogenase is unknown in spite of many years of intense study. It is likely that each of the three N≡N bonds is broken sequentially, giving rise to the intermediates diimine and hydrazine.

$$\mathrm{N{\equiv}N} \xrightarrow{2e^{\ominus},\ 2H^{\oplus}} \underset{\text{Diimine}}{\mathrm{H{-}N{=}N{-}H}} \xrightarrow{2e^{\ominus},\ 2H^{\oplus}} \underset{\text{Hydrazine}}{\mathrm{H_2N{-}NH_2}} \xrightarrow{2e^{\ominus},\ 2H^{\oplus}} \mathrm{NH_3 + NH_3} \qquad \textbf{(17.3)}$$

The reduction of 2 $H^{\oplus}$ to H_2, an essential coupled reaction, consumes the extra pair of electrons from ferredoxin as shown in Reaction 17.2.



◀ **Figure 17.4**
Structure of the $MoFe_7S_9N$ · homocitrate reactive center in *Azotobacter vinelandii*. **(a)** Resting state. **(b)** One possible structure with bound N_2. [PDB 2MIN]



▲ **Figure 17.5**
Incorporation of ammonia into glutamate and glutamine.

Synthetases are members of the Ligase class of enzymes. They require ATP as a cosubstrate. Synthases are members of the Transferase or Lyase class of enzymes. They do not use ATP as a cofactor. (Section 5.1, Section 13.3#1).

17.2 Assimilation of Ammonia

Ammonia is assimilated into a large number of low molecular weight metabolites, often via the amino acids glutamate and glutamine. At physiological pH the main ionic form of ammonia is the ammonium ion, $NH_4^{\oplus}$ ($pK_a = 9.2$). However, unprotonated ammonia (NH_3) is the reactive species in the catalytic centers of many enzymes.

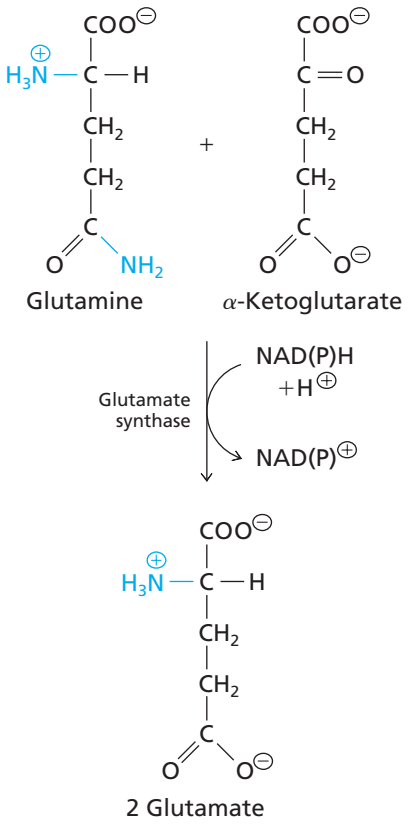
A. Ammonia Is Incorporated into Glutamate and Glutamine

The reductive amination of α-ketoglutarate to glutamate by glutamate dehydrogenase is one highly efficient route for the incorporation of ammonia into the central pathways of amino acid metabolism (Figure 17.5). The glutamate dehydrogenases of some species or tissues are specific for NADH while others are specific for NADPH. Still others can use either cofactor.

The glutamate dehydrogenase reaction can play different physiological roles depending on substrate and coenzyme availability and enzyme specificity. In *Escherichia coli*, for example, the enzyme generates glutamate when $NH_4^{\oplus}$ is present at high concentrations. In the mold *Neurospora crassa* an NADPH-dependent enzyme is used for the reductive amination of α-ketoglutarate to glutamate and the reverse reaction is catalyzed by an $NAD^{\oplus}$-dependent enzyme. Glutamate dehydrogenase is located in mitochondria in mammals and plants and it catalyzes a near equilibrium reaction with net flux usually from glutamate to α-ketoglutarate. The primary role of glutamate dehydrogenase in mammals is the degradation of amino acids and the release of $NH_4^{\oplus}$. Mammals probably assimilate very little nitrogen as free ammonia because they get most of their nitrogen from amino acids and nucleotides in the diet.

Another reaction critical to the assimilation of ammonia in many organisms is the formation of glutamine from glutamate and ammonia. This reaction is catalyzed by glutamine synthetase (Figure 17.5). Glutamine is a nitrogen donor in many biosynthetic reactions; for example, the amide nitrogen of glutamine is the direct precursor of several of the nitrogen atoms of the purine and pyrimidine ring systems of nucleotides (Sections 18.1 and 18.3). In mammals, glutamine carries nitrogen and carbon between tissues in order to avoid high levels of toxic $NH_4^{\oplus}$ in the bloodstream.

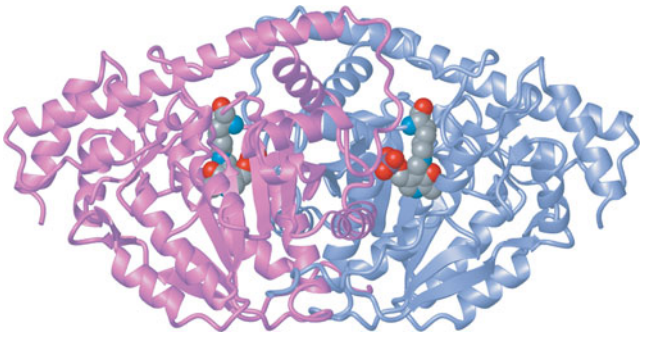
The amide nitrogen of glutamine can be transferred to α-ketoglutarate to produce two molecules of glutamate in a reductive amination reaction catalyzed by glutamate synthase (Figure 17.6). Like glutamate dehydrogenase, glutamate synthase requires a reduced pyridine nucleotide to reductively aminate α-ketoglutarate. Unlike the dehydrogenase, the synthase uses glutamine as the source of nitrogen. Animals do not have glutamate synthase.



▲ **Figure 17.6**
Glutamate synthase catalyzes the reductive amination of α-ketoglutarate.

B. Transamination Reactions

The amino group of glutamate can be transferred to many α-keto acids in reactions catalyzed by enzymes known as transaminases or aminotransferases. The general transamination reaction is shown in Figure 17.7.



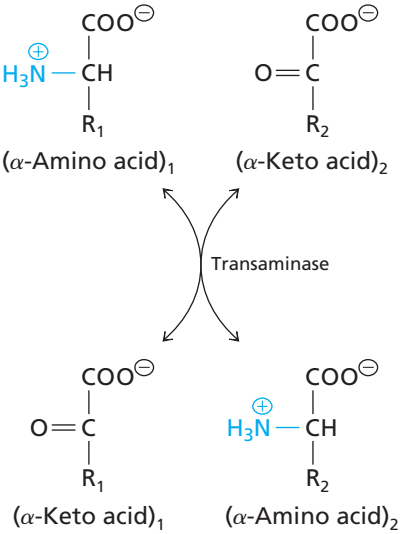
▲ **Pig (*Sus scrofa*) cytosolic aspartate transaminase.** The enzyme is a dimer of identical subunits (individual monomers are shown in purple and blue). A molecule of the coenzyme pyridoxal phosphate is shown (space-filling model) in each active site. [PDB 1AJR]

The amino group of glutamate is transferred to various α-keto acids generating the corresponding α-amino acids during amino acid synthesis. Most of the common amino acids can be formed by transamination. In amino acid catabolism, amino groups are transferred from various amino acids to α-ketoglutarate or oxaloacetate generating glutamate or aspartate.

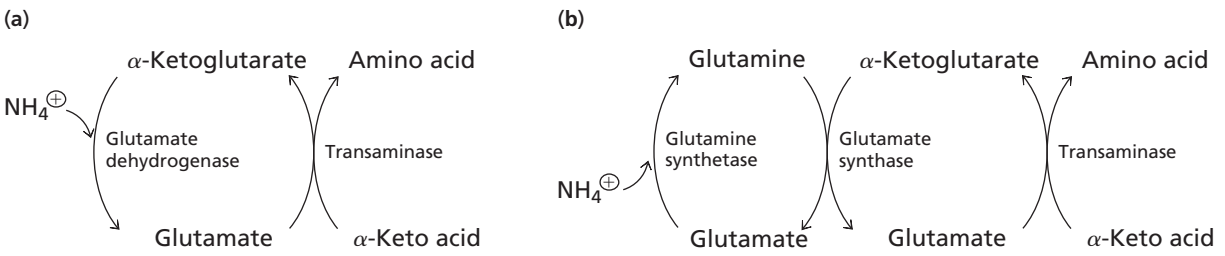
All known transaminases require the coenzyme pyridoxal phosphate (Section 7.8). The chemical mechanism of the initial half-reaction of transamination was shown in Figure 7.18. The complete transamination requires two coupled half-reactions, with enzyme-bound pyridoxamine phosphate (PMP) transiently carrying the amino group being transferred.

The transaminases catalyze near-equilibrium reactions. The direction in which the reactions proceed *in vivo* (flux) depends on the supply of substrates and the removal of products. For example, in cells with an excess of α-amino nitrogen groups the amino groups can be transferred via one or a series of transamination reactions to α-ketoglutarate to yield glutamate that can undergo oxidative deamination catalyzed by glutamate dehydrogenase. Transamination occurs in the opposite direction when amino acids are being actively formed and the amino groups are donated by glutamate.

An important alternative to the glutamate dehydrogenase reaction in bacteria uses coupled reactions catalyzed by glutamine synthetase and glutamate synthase for the assimilation of ammonia into glutamate, especially when the concentration of ammonia is low. Figure 17.8 shows how the combined actions of glutamine synthetase and glutamate synthase can lead to the incorporation of ammonia into a variety of amino acids. After formation, glutamate undergoes transamination with α-keto acids to form the corresponding amino acids. The conversion of α-ketoglutarate to glutamate can occur via the glutamine synthetase–glutamate synthase pathway at the low concentrations of $NH_4^{\oplus}$ present in most bacterial cells because the K_m of glutamine synthetase for NH_3 is much lower than the K_m of glutamate dehydrogenase for $NH_4^{\oplus}$.



▲ **Figure 17.7**
Transfer of an amino group from an α-amino acid to an α-keto acid, catalyzed by a transaminase. In biosynthetic reactions (α-amino acid)$_1$ is often glutamate, with its carbon skeleton producing α-ketoglutarate [= (α-keto acid)$_1$]. (α-keto acid)$_2$ represents the precursor of a newly formed acid, (α-amino acid)$_2$.



▲ **Figure 17.8**
Assimilation of ammonia into amino acids. (a) The glutamate dehydrogenase pathway. **(b)** Combined action of glutamine synthetase and glutamate synthase under conditions of low $NH_4^{\oplus}$ concentration.

17.3 Synthesis of Amino Acids

We now turn our attention to the origins of the carbon skeletons of amino acids. Figure 17.9 shows how the biosynthesis pathways leading to the 20 common amino acids are related to other metabolic pathways. Note that 11 of the 20 common amino acids are synthesized from intermediates in the citric acid cycle. The others require simple precursors that we have encountered in previous chapters.

A. Aspartate and Asparagine

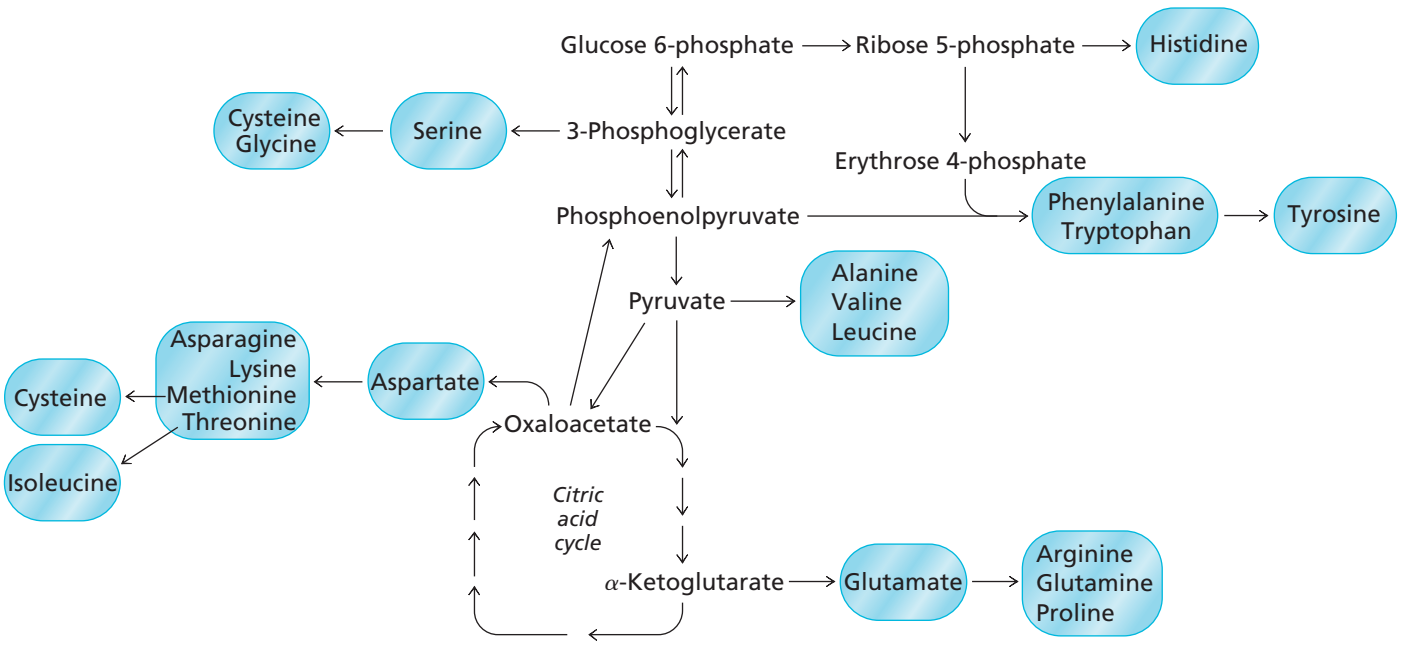
Oxaloacetate is the amino group acceptor in a transamination reaction that produces aspartate (Figure 17.10). The enzyme that catalyzes this reaction is aspartate transaminase (L-aspartate:2-oxoglutarate aminotransferase). Asparagine is synthesized in most species by an ATP-dependent transfer of the amide nitrogen of glutamine to aspartate in a reaction catalyzed by asparagine synthetase. In some bacteria, asparagine synthetase catalyzes the formation of asparagine from aspartate using ammonia instead of glutamine as the source of the amide group. This reaction is similar to the reaction catalyzed by glutamine synthetase.

Some asparagine synthetases can use either ammonia or glutamine as the substrate. These enzymes use $NH_4^{\oplus}$ at the primary reaction site but they have a second site that catalyzes hydrolysis of glutamine and release of $NH_4^{\oplus}$. The $NH_4^{\oplus}$ intermediate diffuses through a tunnel in the protein that connects the two active sites. This example of molecular channeling ensures that the hydrolysis of glutamine is tightly coupled to the formation of asparagine and it prevents the accumulation of $NH_4^{\oplus}$ in the cell. There are many examples of molecular tunnels that facilitate the channeling of $NH_4^{\oplus}$ (see Box 18.2).

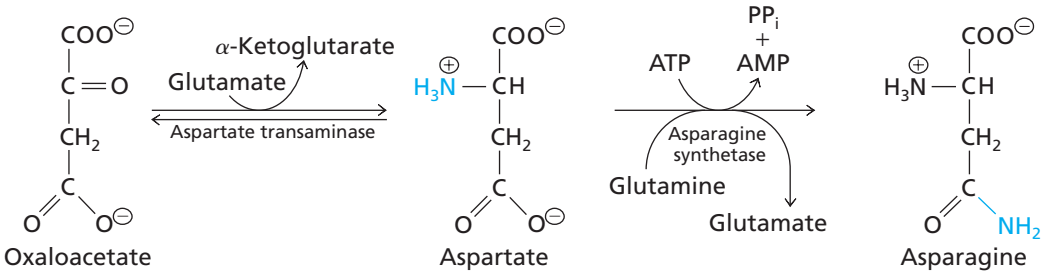
B. Lysine, Methionine, and Threonine

Aspartate is the precursor for synthesis of lysine, methionine, and threonine (Figure 17.11). The first step in the pathway is the phosphorylation of aspartate in a reaction catalyzed by aspartate kinase. In the second step, aspartyl phosphate is converted to aspartate β-semialdehyde. This second reaction is catalyzed by aspartate semialdehyde dehydrogenase. These two enzymes are present in bacteria, protists, fungi, and plants but they are missing in animals. Consequently, animals are not able to synthesize lysine, methionine, and threonine (see Box 17.3).

The first two reactions leading to aspartate β-semialdehyde are common to the formation of all three amino acids. In the branch leading to lysine, pyruvate is the source of carbon atoms added to the skeleton of aspartate β-semialdehyde and glutamate is the



▼ **Figure 17.9**
Biosynthesis of amino acids, showing the connections to glycolysis/gluconeogenesis and the citric acid cycle.



◀ **Figure 17.10**
Synthesis of aspartate and asparagine.

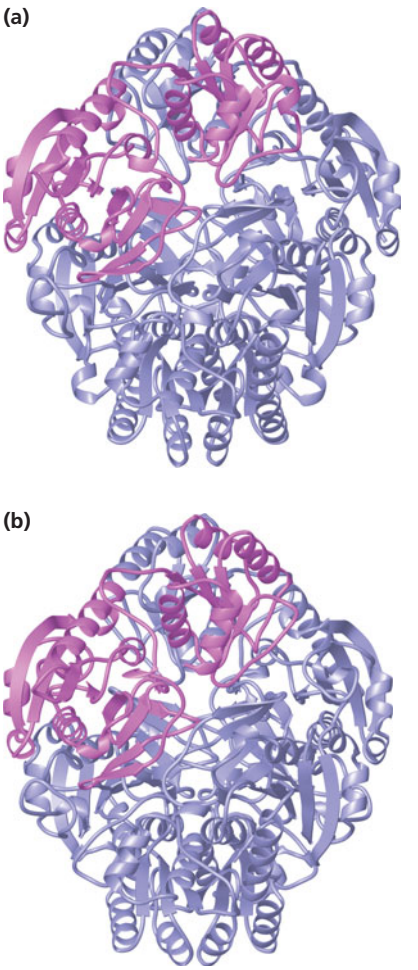
source of the ε-amino group. Lysine is produced by an entirely different route in yeast and some algae.

Homoserine is formed from aspartate β-semialdehyde. It is a branch point for the formation of threonine and methionine. Threonine is derived from homoserine in two steps, one of which requires PLP. In the methionine pathway homoserine is converted to homocysteine in three steps. The sulfur atom of homocysteine then accepts a methyl group derived from 5-methyltetrahydrofolate forming methionine. The enzyme that catalyzes this reaction is homocysteine methyltransferase, one of the few enzymes that requires cobalamin (Section 7.12). Homocysteine methyltransferase is found in mammals but its activity is low and the supply of homocysteine is limited. Therefore, methionine remains an essential amino acid in mammals due primarily to the absence of the first two enzymes in the pathway.

C. Alanine, Valine, Leucine, and Isoleucine

Pyruvate is the amino group acceptor in the synthesis of alanine by a transamination reaction (Figure 17.12). Pyruvate is also a precursor in the synthesis of the branched chain amino acids valine, leucine, and isoleucine. The first step in the branched chain pathway is the synthesis of α-ketobutyrate from threonine.

Pyruvate combines with α-ketobutyrate in a series of three reactions leading to the branched chain intermediate α-keto-β-methylvalerate. This intermediate is converted to isoleucine in a transamination reaction. Note the similarity between the structures of α-ketobutyrate and pyruvate. The same enzymes that catalyze the synthesis of α-keto-β-methylvalerate also catalyze the synthesis of α-ketoisovalerate by combining two molecules of pyruvate instead of one molecule of pyruvate and one molecule of α-ketobutyrate. α-Ketoisovalerate is converted directly to valine by valine transaminase—the same enzyme catalyzes the synthesis of isoleucine from α-keto-β-methylvalerate (Figure 17.13). These pathways illustrate an important point, namely that some enzymes recognize several different but similar substrates. At some point in the future the eukaryotic genes for these enzymes might be duplicated and each of the two copies would evolve to become specific for either the isoleucine or valine pathways. If this happened, it would be an example of pathway evolution by gene duplication and divergence (Section 10.2D). We see



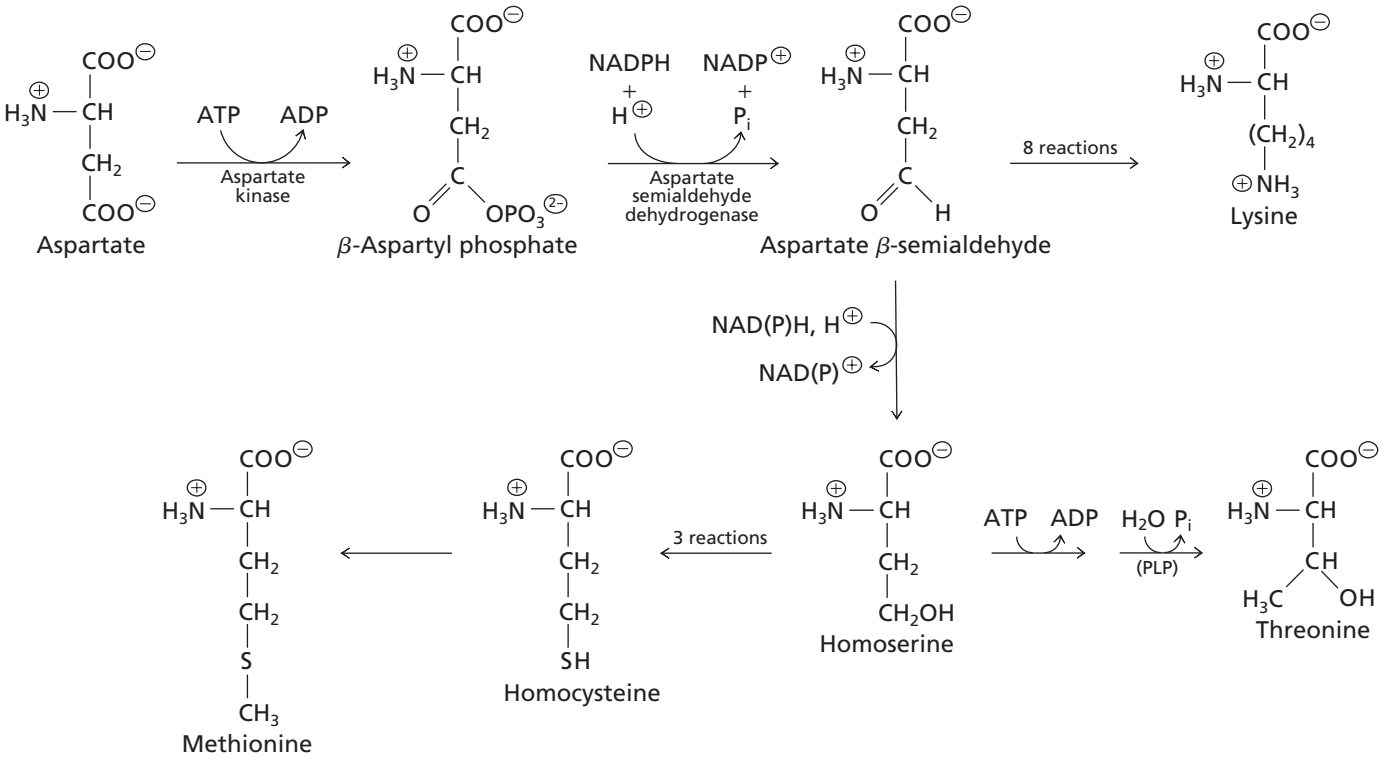
▲ **Asparaginases. (a)** *Escherichia coli* [PDB 1NNS] **(b)** *Erwinia chrysanthemi* [PDB 1O7J]

BOX 17.1 CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA CAN BE TREATED WITH ASPARAGINASE

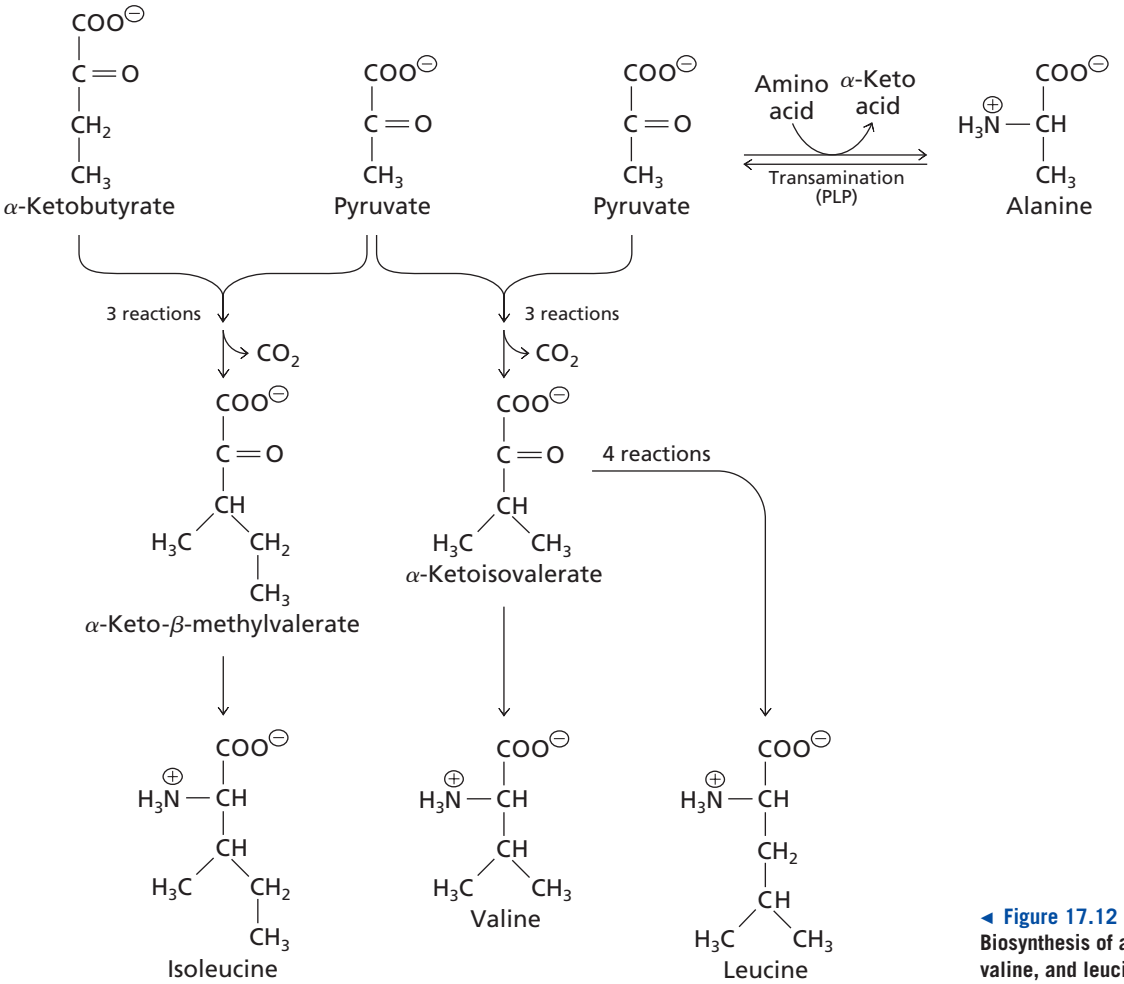
Acute lymphoblastic leukemia (ALL) is caused by the proliferation of malignant T-cell lymphoblasts due, in most cases, to a mutation caused by mistakes in genetic recombination during the activation of T-cell receptor genes. Malignant lymphoblasts have reduced levels of asparagine synthetase and are unable to synthesize enough asparagine to support their rapid growth and proliferation. Unlike normal cells, they must obtain asparagine from the blood plasma.

This cancer can be successfully treated with injections of asparaginase, an *E. coli* enzyme that breaks down asparagine in the plasma (Section 17.6A). The malignant cells die in the absence of an available source of asparagine. Treatment with asparaginase alone causes remission in 50% of all cases of childhood acute lymphoblastic leukemia and the success rate is even higher when the enzyme treatment is combined with other chemotherapy. The primary cause of resistance to the treatment is due to increased expression of asparagine synthetase in the cancer cells.

Patients often develop antibodies to the *E. coli* enzyme during treatment. Switching to the homologous enzyme from *Erwinia chrysanthemi* is often effective because the amino acid side chains on the surface of the two proteins are different. Antibodies directed against one enzyme usually don't recognize the other.



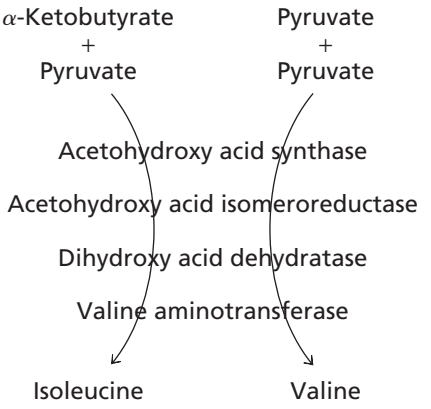
▲ **Figure 17.11**
Biosynthesis of lysine, threonine, and methionine from aspartate.



◀ **Figure 17.12**
Biosynthesis of alanine, isoleucine, valine, and leucine.

many examples of pathway evolution by gene duplication involving enzymes of amino acid metabolism (see below). The basic requirement is that in the early stages the same enzyme can catalyze two similar reactions and that is what we see in the isoleucine and valine synthesis pathways.

The carbon skeleton of α-ketoisovalerate is lengthened by one methylene group to form leucine in a pathway that branches from the valine biosynthetic pathway. Two of the enzymes in this pathway are homologous to aconitase and isocitrate dehydrogenase in the citric acid cycle lending support to the idea that citric acid cycle enzymes evolved from preexisting enzymes required for amino acid biosynthesis (Section 13.8).



▲ **Figure 17.13**
The isoleucine and valine synthesis pathways share four enzymes.

D. Glutamate, Glutamine, Arginine, and Proline

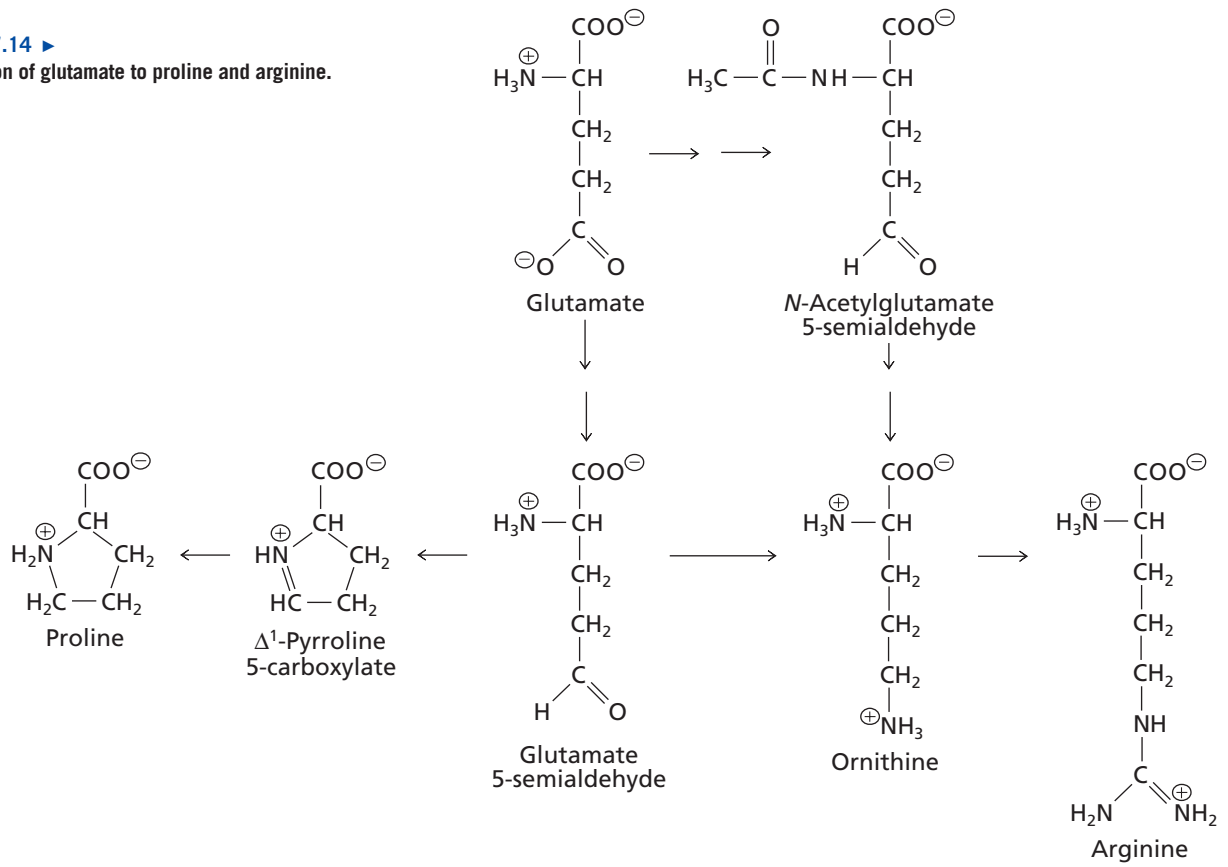
We have seen how glutamate and glutamine are formed from the citric acid cycle intermediate α-ketoglutarate (Section 17.2B). The carbon atoms of proline and arginine also come from α-ketoglutarate, via glutamate. Proline is synthesized from glutamate by a four-step pathway in which the 5-carboxylate group of glutamate is reduced to an aldehyde. The glutamate 5-semialdehyde intermediate undergoes nonenzymatic cyclization to a Schiff base, 5-carboxylate, that is reduced by a pyridine nucleotide coenzyme to produce proline (Figure 17.14).

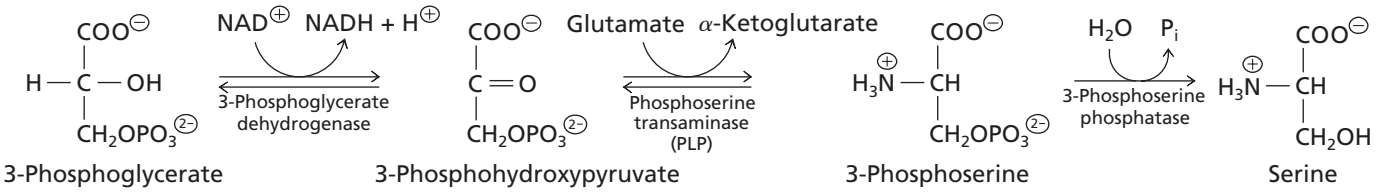
The pathway to arginine is similar in most species except that the α-amino group of glutamate is acetylated before the aldehyde is formed. This step prevents the cyclization that occurs in the synthesis of proline. The *N*-acetylglutamate 5-semialdehyde intermediate is then converted to *N*-acetyl ornithine and ornithine. In mammals, glutamate 5-semialdehyde is transaminated to ornithine and ornithine is converted to arginine by the reactions of the urea cycle (Section 17.7).

E. Serine, Glycine, and Cysteine

Three amino acids—serine, glycine, and cysteine—are derived from the glycolytic/gluconeogenic intermediate 3-phosphoglycerate. Serine is synthesized from 3-phosphoglycerate in three steps (Figure 17.15). First, the secondary hydroxyl substituent of

Figure 17.14 ►
Conversion of glutamate to proline and arginine.





▲ **Figure 17.15**
Biosynthesis of serine.

3-phosphoglycerate is oxidized to a keto group, forming 3-phosphohydroxypyruvate. This compound undergoes transamination with glutamate to form 3-phosphoserine and α-ketoglutarate. Finally, 3-phosphoserine is hydrolyzed to give serine and P_i.

Serine is a major source of glycine via a reversible reaction catalyzed by serine hydroxymethyltransferase (Figure 17.16). In plant mitochondria and bacteria, the flux through this reaction is toward serine providing a route to serine that differs from that in Figure 17.15. The serine hydroxymethyltransferase reaction requires two cofactors: the prosthetic group PLP and the cosubstrate tetrahydrofolate.

The biosynthesis of cysteine from serine occurs in two steps (Figure 17.17). First, an acetyl group from acetyl CoA is transferred to the β-hydroxyl substituent of serine, forming *O*-acetylserine. Next, sulfide ($S^{2\ominus}$) displaces the acetate group, and cysteine is formed.

Animals do not have the normal cysteine biosynthesis pathway shown in Figure 17.17. Nevertheless, cysteine can still be synthesized in animals as a by-product of methionine degradation (Section 17.6F). Serine condenses with homocysteine, an intermediate in the degradation of methionine. The product of the condensation reaction, crystathionine, is cleaved to α-ketobutyrate and cysteine (Figure 17.18).

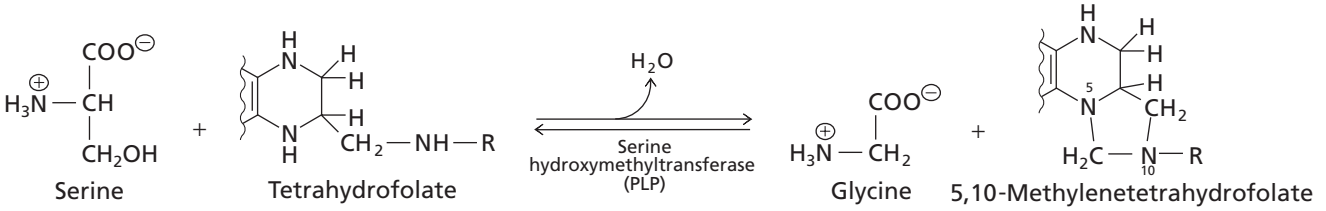
F. Phenylalanine, Tyrosine, and Tryptophan

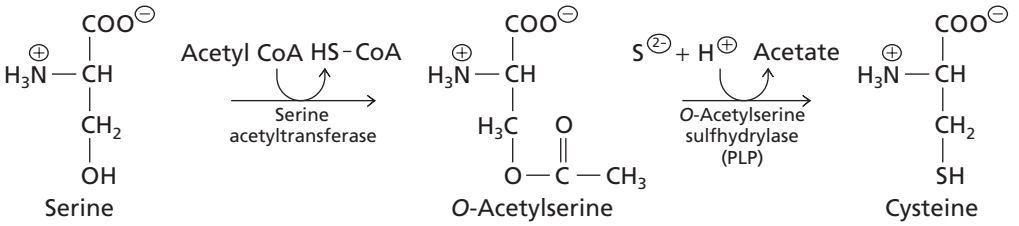
The key to elucidation of the pathway for aromatic amino acid synthesis was the observation that some bacteria with single-gene mutations require as many as five compounds for growth: phenylalanine, tyrosine, tryptophan, *p*-hydroxybenzoate, and *p*-aminobenzoate. These compounds all contain an aromatic ring. The inability of these mutants to grow without these compounds is reversed when shikimate is provided indicating that shikimate is an intermediate in the biosynthesis of all these aromatic compounds.

Chorismate, a derivative of shikimate, is a key branch-point intermediate in aromatic amino acid synthesis. The pathway to shikimate and chorismate (Figure 17.19) begins with condensation of phosphoenolpyruvate and erythrose 4-phosphate to form a seven-carbon sugar derivative and P_i. Three additional steps, including cyclization, are required to produce shikimate. The pathway from shikimate to chorismate involves phosphorylation of shikimate, addition of a three-carbon group from phosphoenolpyruvate, and dephosphorylation. Pathways from chorismate lead to phenylalanine, tyrosine, and tryptophan. Animals do not have the enzymes of the chorismate pathway. They cannot synthesize chorismate and, consequently, cannot synthesize any of the aromatic amino acids.

A branched pathway leads from chorismate to phenylalanine or tyrosine (Figure 17.20). In phenylalanine synthesis in *E. coli*, a bifunctional chorismate mutase–prephenate dehydratase catalyzes the rearrangement of chorismate to produce prephenate, a highly reactive compound. Next, the enzyme catalyzes the elimination of a hydroxide ion and CO_2

▼ **Figure 17.16**
Biosynthesis of glycine.





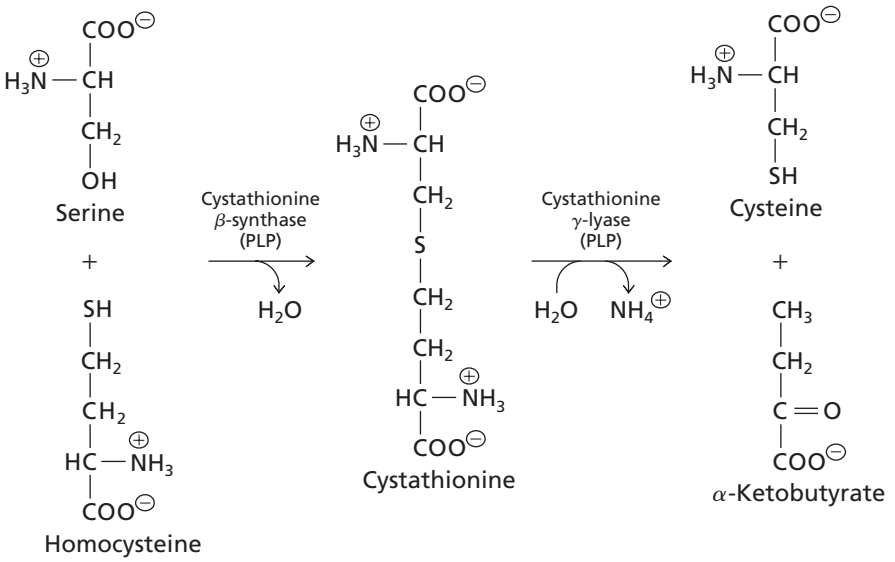
◀ **Figure 17.17**
Biosynthesis of cysteine from serine in many bacteria and plants.

from prephenate to form the fully aromatic product phenylpyruvate that is then transaminated to phenylalanine.

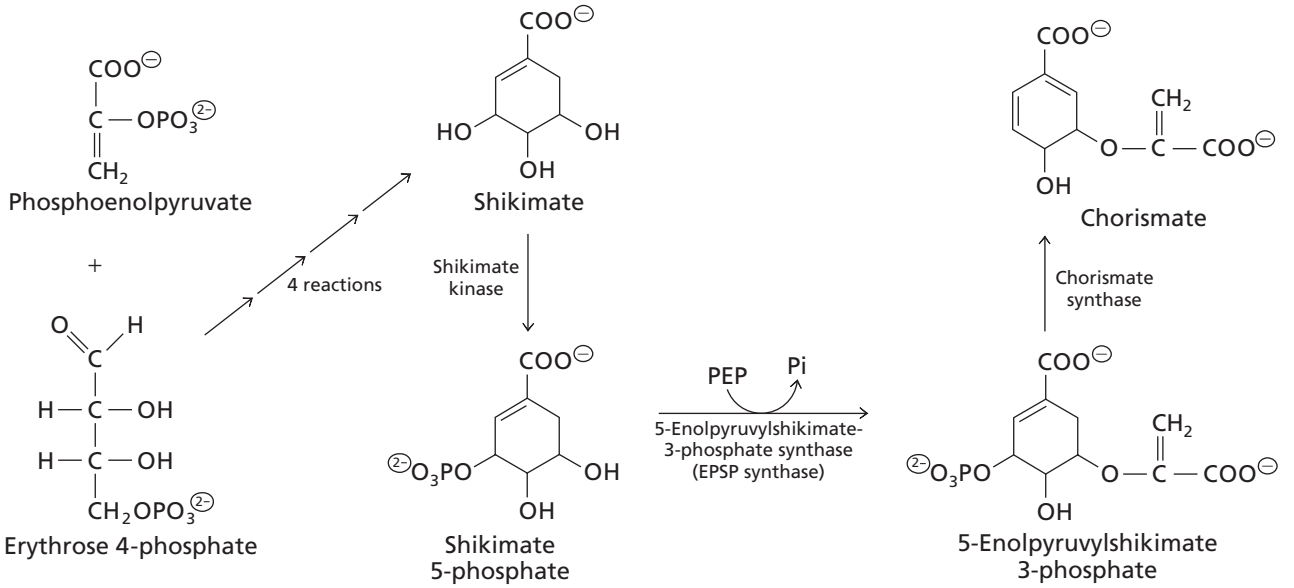
A similar bifunctional chorismate mutase–prephenate dehydrogenase catalyzes the formation of prephenate and then 4-hydroxyphenylpyruvate in the tyrosine branch. The intermediate undergoes transamination to form tyrosine. Several bacteria and some plants follow the same pathways from chorismate to phenylalanine and tyrosine as *E. coli* although their chorismate mutase and prephenate dehydratase or prephenate dehydrogenase activities are on separate polypeptide chains. Some other bacteria use alternate pathways in which prephenate is first transaminated and then decarboxylated.

The biosynthesis of tryptophan from chorismate requires five enzymes. In the first step, the amide nitrogen of glutamine is transferred to chorismate. Subsequent elimination of the hydroxyl group and the adjacent pyruvate moiety of chorismate produces the aromatic compound anthranilate (Figure 17.21). Anthranilate accepts a phosphoribosyl moiety from PRPP. Rearrangement of the ribose, decarboxylation, and ring closure generate indole glycerol phosphate.

The final two reactions of tryptophan biosynthesis are catalyzed by tryptophan synthase (Figure 17.22). In some organisms, the two independent catalytic domains of tryptophan synthase are contained on a single polypeptide chain but in some species the enzyme contains two types of subunits in an $\alpha_2\beta_2$ tetramer. The α subunit, or domain, catalyzes the cleavage of indole glycerol phosphate to glyceraldehyde 3-phosphate and indole. The β subunit, or domain, catalyzes the condensation of indole and serine in a reaction that requires PLP as a cofactor. The indole produced in the reaction catalyzed by the α subunit of $\alpha_2\beta_2$ tetramers is channeled (i.e., transferred directly) to the active site of the β subunit. When the three-dimensional structure of tryptophan synthase from *Salmonella typhimurium* (an organism whose tryptophan synthase has the $\alpha_2\beta_2$ oligomeric structure) was determined by X-ray crystallography, a tunnel joining the α and β active sites was discovered. The diameter of the tunnel matches the molecular dimensions of indole, so passage of indole through the tunnel would explain why

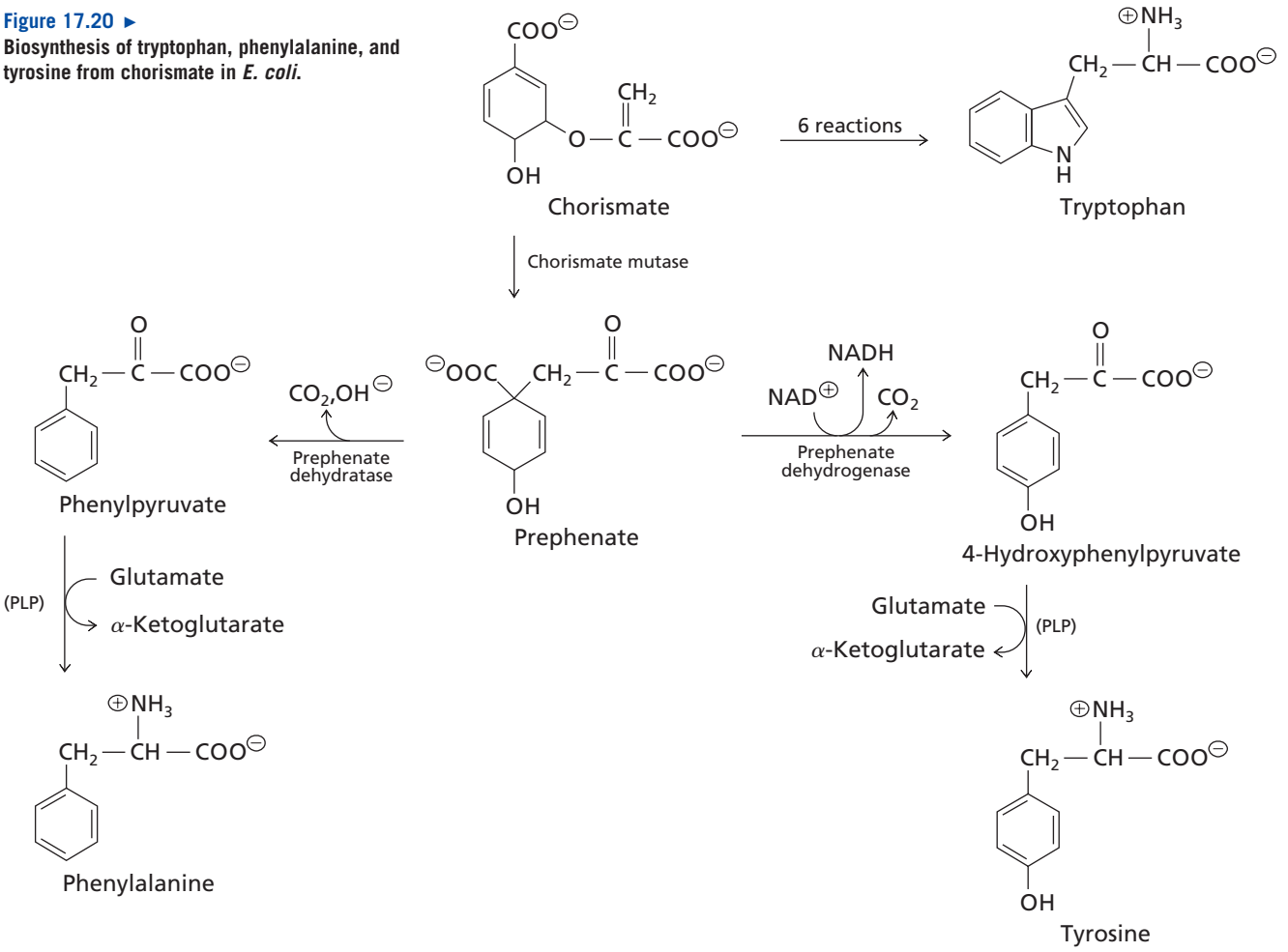


◀ **Figure 17.18**
Biosynthesis of cysteine in mammals.

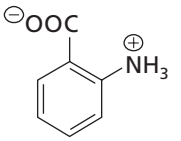


▲ **Figure 17.19**
Synthesis of shikimate and chorismate.

Figure 17.20 ►
Biosynthesis of tryptophan, phenylalanine, and tyrosine from chorismate in *E. coli*.



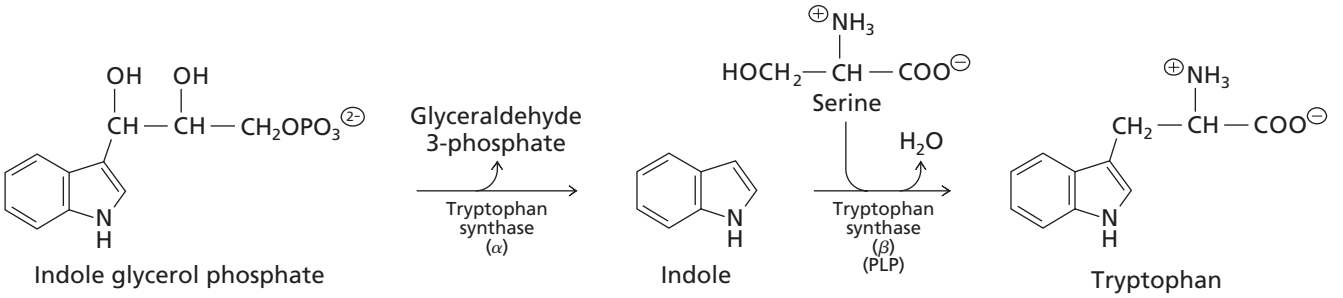
indole does not diffuse away. This was one of the earliest examples of metabolic channeling (Section 5.10). Up until quite recently there were only a few other examples and the phenomenon was thought to be rare. The huge increase in structural and genomic studies has revealed many more examples—including half a dozen in this chapter alone.



▲ **Figure 17.21**
Anthranilate.

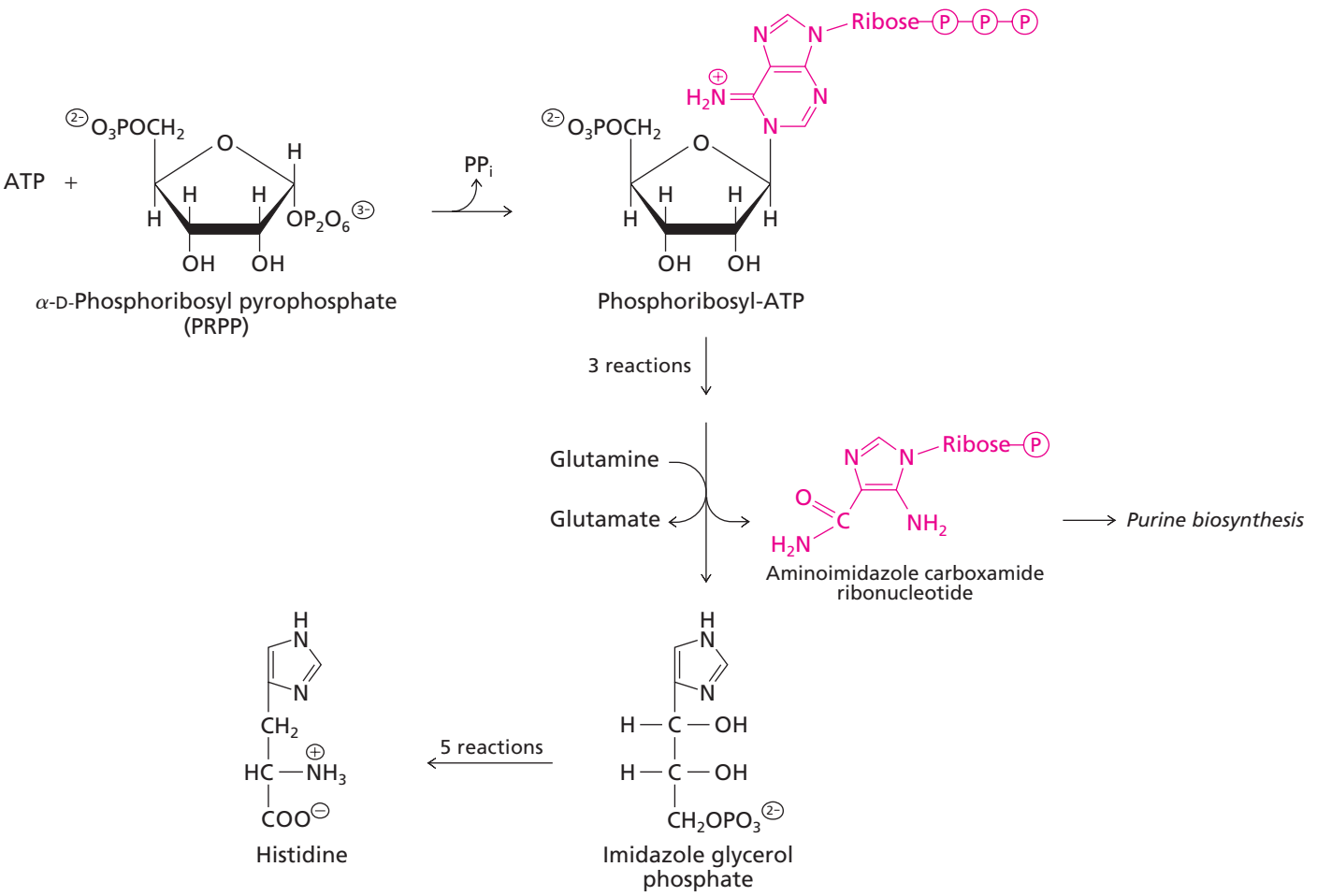
G. Histidine

The ten-step pathway for the biosynthesis of histidine in bacteria begins with a condensation between the six-membered ring of ATP and a ribose derivative, phosphoribosyl pyrophosphate (PRPP) (Figure 17.23). The six-membered ring of the adenine moiety is then cleaved and glutamine donates a nitrogen atom that is incorporated via cyclization



▲ **Figure 17.22**
Reactions catalyzed by tryptophan synthase.

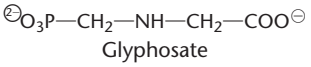
▼ **Figure 17.23**
Synthesis of histidine from phosphoribosyl pyrophosphate (PRPP) and ATP.
Histidine is derived from PRPP (5 C atoms), the purine ring of ATP (1 N and 1 C), glutamine (1 N), and glutamate (1 N).



BOX 17.2 GENETICALLY MODIFIED FOOD

The chorismate pathway is an effective target for herbicides since compounds that specifically block this pathway in plants will have no effect on animals. One of the most effective general herbicides is glyphosate. Glyphosate inhibits the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSP synthase) by acting as a competitive inhibitor of PEP binding (Section 5.7A).

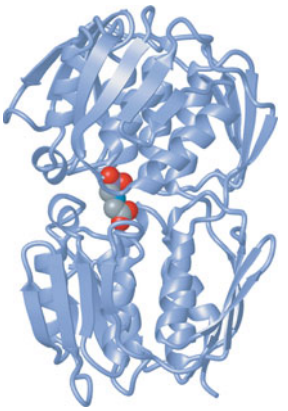
Glyphosate is the active ingredient in Roundup®, a herbicide that kills all plants. It is used to remove weeds from driveways and stone pathways. Although it is cheap and effective as a weed killer, glyphosate cannot be used to spray actively growing food crops since it indiscriminately kills all plants, including the crop!



Glyphosate
(*N*-(phosphonomethyl) glycine)

Resistant versions of EPSP synthase have been identified in many species of bacteria. The enzyme from strain CP4 of *Agrobacterium sp.* has been genetically modified to remain fully active in the presence of high concentrations of glyphosate. The gene for this bacterial CP4-EPSP synthase was patented and then introduced into soybeans creating a genetically modified plant that is resistant to glyphosate. The new strain of soybeans is marketed by Monsanto as Roundup Ready® soybeans. Farmers who grow crops of Roundup Ready® soybeans are able to spray them with Roundup® (also sold by Monsanto) to kill weeds. The economic advantages to farmers are significant. Most of the soybeans currently grown in North America are genetically modified.

Other Roundup Ready® crop plants are now available. Versions of corn, cotton, and canola are widely used.

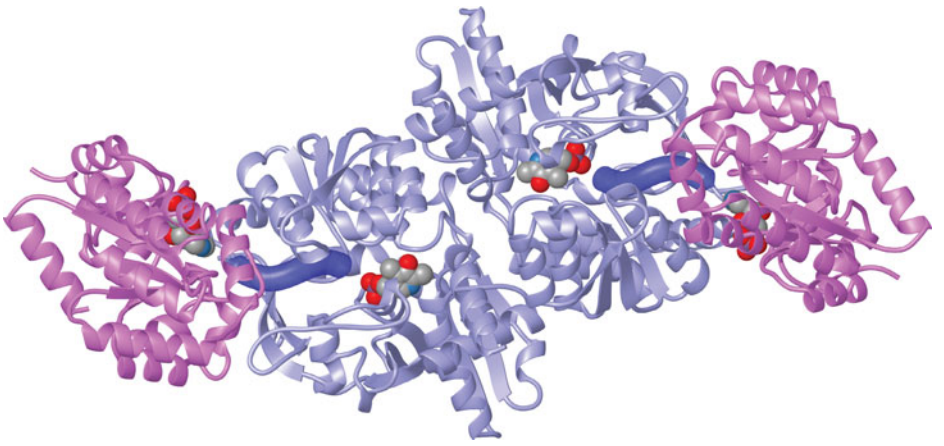


▲ ***E. coli*** **5-enolpyruvylshikimate-3-phosphate synthase with a molecule of glyphosate bound to the active site. [PDB 2AAY]**

KEY CONCEPT

Metabolic channeling evolves to improve kinetic efficiency.

into the imidazole ring of the product, imidazole glycerol phosphate. Most of the carbon and nitrogen atoms of ATP are released as aminoimidazole carboxamide ribonucleotide, an intermediate in purine biosynthesis (Section 18.1). This metabolite then can be recycled into ATP. Imidazole glycerol phosphate undergoes dehydration, transamination by glutamate, hydrolytic removal of its phosphate, and oxidation from the level of a primary alcohol to that of a carboxylic acid in two sequential $NAD^{\oplus}$-dependent steps forming histidine.



▲ **Tryptophan synthase from *Salmonella typhimurium*.** The substrate indole glycerol phosphate is shown as a space-filling molecule bound to the α subunits. The cofactor PLP is bound to the β subunits. The enzyme contains a channel leading from the indole glycerol phosphate binding site to the PLP reaction site. [PDB 1QOQ]

BOX 17.3 ESSENTIAL AND NONESSENTIAL AMINO ACIDS IN ANIMALS

Humans and other animals do not possess the enzymes required for the synthesis of all amino acids. Those that cannot be synthesized are, therefore, essential components of the human diet. As a general rule, the pathways that have been lost are the ones with the most steps. A crude measure of the complexity of a pathway is the number of moles of ATP (or its equivalent) required in a pathway.

The table shows the correlation between the expense of a particular pathway and whether an amino acid is essential. The amino acids are grouped according to their common precursors as described in the previous sections. Note that lysine, methionine, and threonine are derived from a common precursor (Section 17.3B). All three amino acids are essential because animals cannot synthesize the precursor. Valine, leucine, and isoleucine are essential because animals lack the key enzymes that all three biosynthesis pathways share (Section 17.3C).

Energy requirements for biosynthesis of amino acids

Amino acids	Moles of ATP required per mole of amino acid produced[a]	
	Nonessential	Essential
Aspartate	21	
Asparagine	22–24	
Lysine		50 or 51
Methionine		44
Threonine		31
Alanine	20	
Valine		39
Leucine		47
Isoleucine		55
Glutamate	30	
Glutamine	31	
Arginine	44[b]	
Proline	39	
Serine	18	
Glycine	12	
Cysteine	19[c]	
Phenylalanine		65
Tyrosine	62[d]	
Tryptophan		78
Histidine		42

[a]Moles of ATP required includes ATP used for synthesis of precursors and conversion of precursors to products.
[b]Essential in some mammals.
[c]Cysteine can be synthesized from homocysteine and homocysteine is a degradation product of methionine. The biosynthesis of cysteine depends on an adequate supply of methionine in the diet.
[d]Tyrosine can be synthesized from the essential amino acid phenylalanine.

17.4 Amino Acids as Metabolic Precursors

The primary role of amino acids is to serve as substrates for protein synthesis. In this role, newly synthesized amino acids are activated by covalent attachment to tRNA and the pool of aminoacyl–tRNAs is used as the substrate for polypeptide synthesis by the protein synthesis machinery. We devote an entire chapter to this fundamentally important biosynthesis pathway (Chapter 22).

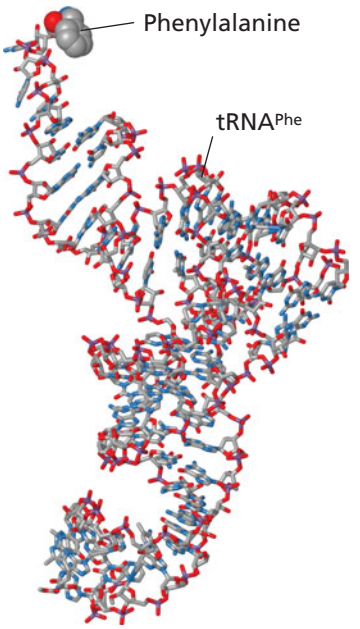
Some amino acids are essential precursors in other biosynthesis pathways. The list is long and it's impossible to mention every pathway. Some important regulatory amines were described in Section 3.3 (histamine, GABA, epinephrine, thyroxine). The important role of methionine in the synthesis of *S*-adenosylmethionine will be described in Section 17.6F.

A. Products Derived from Glutamate, Glutamine, and Aspartate

We've already seen that glutamate and glutamine are important players in nitrogen assimilation. In addition, glutamate and aspartate are amino group donors in many transamination reactions. We will see that glutamate and aspartate are required in the urea cycle. Glutamine and aspartate are also required as precursors in both purine biosynthesis (Section 18.1) and pyrimidine biosynthesis (Section 18.3). Recall that synthesis of biologically active tertahydrofolate involves addition of up to six glutamate residues to the tetrahydrofolate moiety (Section 7.10).

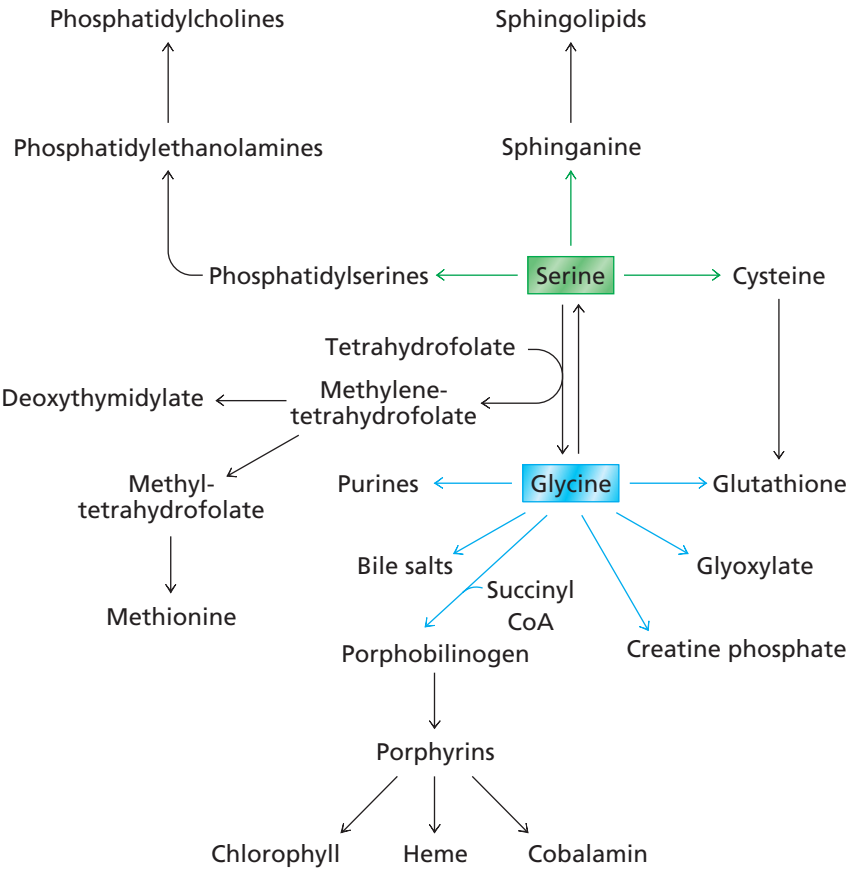
B. Products Derived from Serine and Glycine

Serine and glycine are metabolic precursors of many other compounds (Figure 17.24). The role of serine in lipid biosynthesis has already been described in the previous chapter.



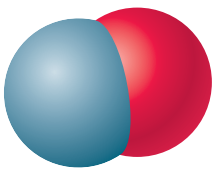
▲ **Phenylanyl–tRNA^{Phe}.** Most newly synthesized amino acids are rapidly attached to their corresponding tRNAs and used in protein synthesis. [PDB 1TTT]

Figure 17.24 ►
Compounds formed from serine and glycine.



Glycine and succinyl CoA are the main precursors in the porphyrin pathway leading to heme and chlorophyll. Glycine is also required in purine biosynthesis (Section 18.1).

The conversion of serine to glycine is coupled to the synthesis of methylene tetrahydrofolate. Tetrahydrofolate derivatives are important in many reactions that catalyze transfer of one-carbon units (Section 17.10). One of the most important of these reactions is the synthesis of deoxythymidylate (Figure 18.15).



▲ Nitric oxide

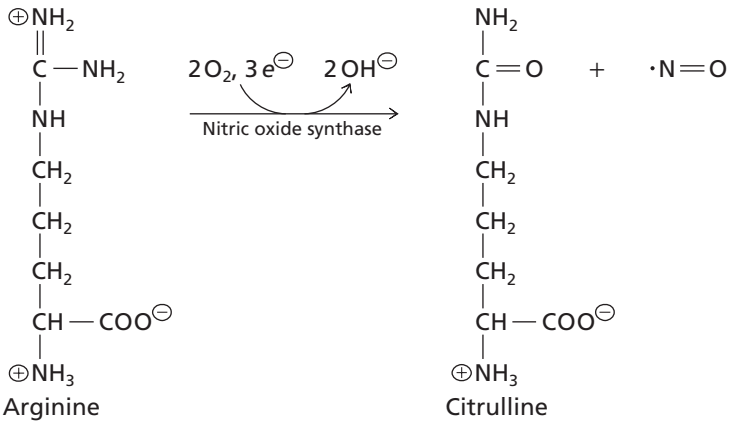
C. Synthesis of Nitric Oxide from Arginine

One of the more interesting examples of amino acids as metabolic precursors is the role of arginine as substrate for synthesis of nitric oxide, an unstable gaseous derivative of nitrogen with an odd number of electrons (•N═O). Although it is a reactive free radical and potentially toxic, nitric oxide is physiologically important—so important, in fact, that it was named the 1992 "Molecule of the Year" by the journal *Science*. As a gas, NO can diffuse rapidly into cells. It exists *in vivo* for only a few seconds because nitric oxide in aqueous solution reacts rapidly with oxygen and water to form nitrates and nitrites.

An enzyme found in mammals, nitric oxide synthase, catalyzes the formation of nitric oxide and citrulline from arginine (Figure 17.25). The reaction requires the cofactors NADPH, FMN, FAD, a cytochrome P450, and tetrahydrobiopterin (Section 7.10). The mechanism of action of tetrahydrobiopterin in this reaction has not yet been elucidated but it appears to be a reducing agent needed for the hydroxylation of arginine. Nitric oxide synthase is present in two forms, a constitutive (i.e., constantly synthesized) calcium-dependent form in brain and endothelial cells and an inducible (i.e., variably synthesized) calcium-independent form in macrophages (a type of white blood cell).

In 1998 Robert F. Furchgott, Louis J. Ignarro, and Ferid Murad were awarded the Nobel Prize in Physiology or Medicine "for their discoveries concerning nitric oxide as a signaling molecule in the cardiovascular system."

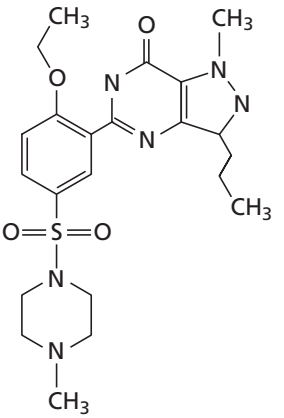
Nitric oxide is a messenger molecule that binds to a soluble guanylyl cyclase and stimulates the formation of cyclic GMP (Section 9.12B). It has several functions; for example, when macrophages are stimulated, they synthesize nitric oxide. The short-lived nitric oxide free radical is one of the weapons used by macrophages to kill bacteria and tumor cells. Nitric oxide may interact with superoxide anions ($\bullet O_2^{\ominus}$) to form more toxic reactants that account for the cell-killing activity.



◀ **Figure 17.25**
Conversion of arginine to nitric oxide and citrulline. NADPH is the source of the three electrons.

Nitric oxide synthase is also present in the cells that line blood vessels. Under certain conditions, nitric oxide is produced and diffuses to the smooth muscle cells of the vessels, causing them to relax and lower blood pressure. Hypertension and heart failure involve impaired relaxation of blood vessels. Nitroglycerin, used to dilate coronary arteries in the treatment of angina pectoris, exerts its effect by virtue of its metabolic conversion to nitric oxide.

Nitric oxide also functions as a neurotransmitter in brain tissue. Abnormally high amounts of nitric oxide formed during a stroke appear to kill some neurons in the same way macrophages kill bacteria. Administering an inhibitor of nitric oxide synthase to an animal produces some protection from stroke damage. One role of nitric oxide as a neurotransmitter is to stimulate erection of the penis. Sildenafil, the active ingredient in Viagra, is a drug used to alleviate erectile dysfunction. Sildenafil is a phosphodiesterase inhibitor that blocks the hydrolysis of cyclic GMP and therefore prolongs the stimulatory effect of nitric oxide. Tadalafil (Cialis) and vardenafil (Levitra) inhibit the same enzyme.



▲ **Sildenafil.** Sildenafil is the active ingredient in Viagra®.

D. Synthesis of Lignin from Phenylalanine

Lignin (Figure 17.26) is a series of complex polymers synthesized from phenylalanine. It is a major component of wood in flowering plants and may be the second most abundant biopolymer on the planet (after cellulose). Lignin cannot be broken down during digestion so in spite of the fact that animals ingest huge amounts of lignin it is metabolically inert. The only species that can break it down are various fungi that degrade fallen trees in the forest.

E. Melanin Is Made from Tyrosine

Melanin is a dark pigment found in bacteria, fungi, and animals. In humans it is responsible for skin color and hair color. Melanin is also the main component of the ink released by a frightened octopus.

The structure of melanin (eumelanin) is complex but the precursors are well known and the enzymes required in the pathway have been identified in a number of species. The first steps involve the conversion of L-tyrosine to L-DOPA and L-dopaquinone (Figure 17.27).

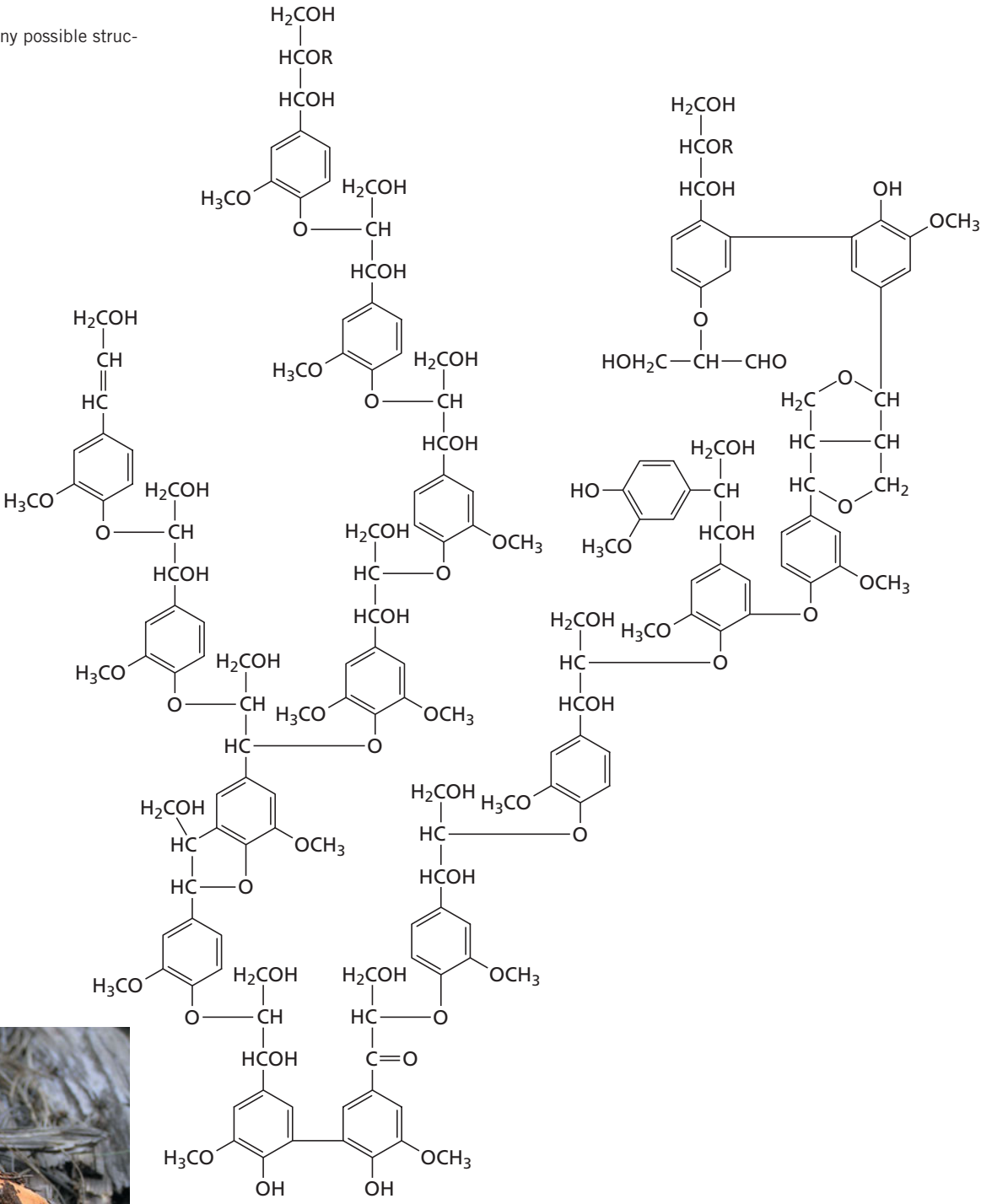


▲ **Octopus ink is mostly melanin.**

17.5 Protein Turnover

One might assume that only growing or reproducing cells would require new protein molecules (and therefore a supply of amino acids) but this is not the case. Proteins are continually synthesized and degraded in all cells, a process called turnover. Individual proteins turn over at different rates. Their half-lives can vary from a few minutes to several weeks but the half-life of a given protein in different organs and species is generally similar. Rapid protein turnover ensures that some regulatory proteins are degraded so that the cell can respond to constantly changing conditions. Such proteins have evolved to be relatively unstable.

Figure 17.26 ►
Lignin. This is one of many possible structures of plant lignin.



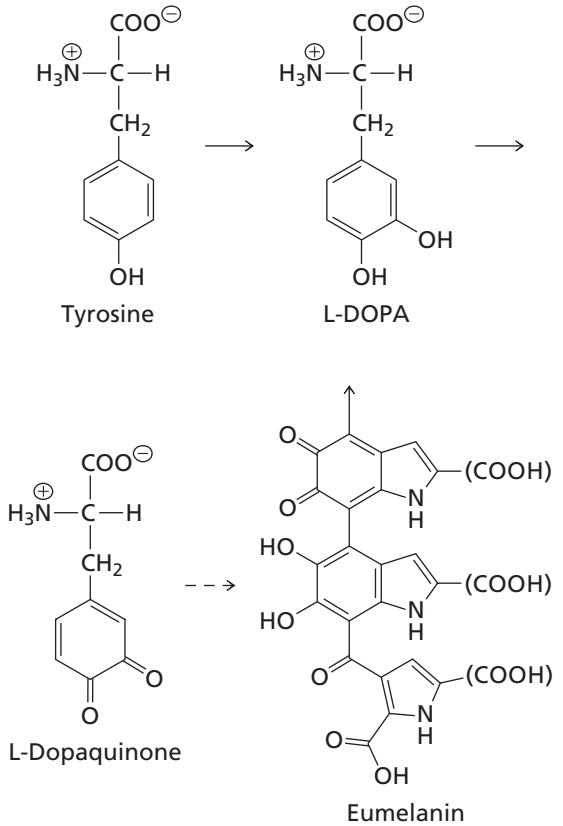


▲ **Rotting wood.** This mushroom is growing on rotting wood in a deciduous forest. Fungi are the only organisms that produce enzymes for breaking down lignin.

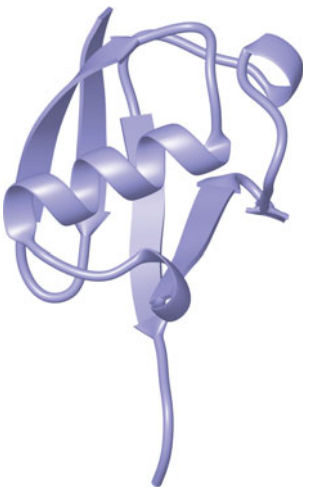
The rate of hydrolysis of a protein can be inversely related to the stability of its tertiary structure. Misfolded and unfolded proteins are quickly degraded (Section 4.10).

Some proteins are degraded to amino acids through lysosomal hydrolysis (in eukaryotic cells). Vesicles containing material to be destroyed fuse with lysosomes, and various lysosomal proteases hydrolyze the engulfed proteins. The lysosomal enzymes have broad substrate specificities so all the trapped proteins are extensively degraded.

Some proteins have very short half-lives because they are specifically targeted for degradation. Abnormal (mutated) proteins are also selectively hydrolyzed. The pathway for the selective hydrolysis of these proteins in eukaryotic cells requires the protein ubiquitin. Side-chain amino groups of lysine residues in the target protein are covalently linked to the C-terminus of ubiquitin in a complex pathway that involves



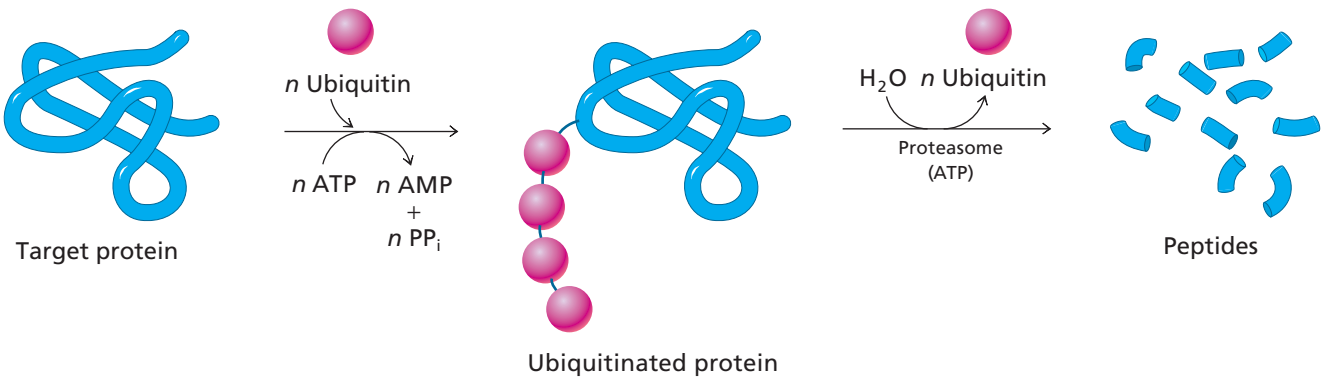
◀ **Figure 17.27**
Synthesis of eumelanin from tyrosine and L-DOPA.



▲ **Ubiquitin (*Homo sapiens*).** Ubiquitin is a small, highly conserved, eukaryotic protein used as a marker that targets proteins for degradation. [PDB 1UBI]

ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin-protein ligase (E3). This pathway is coupled to ATP hydrolysis—one ATP molecule is hydroylzed for every ubiquitin molecule attached to the target protein. The ubiquitinated protein is hydrolyzed to peptides by the action of a large multiprotein complex called the proteasome (or proteosome) (Figure 17.28). This process occurs in both the cytosol and the nucleus. Other proteases catalyze hydrolysis of the resulting peptides. ATP is required to assemble the proteasome and to hydrolyze the ubiquitinated protein. Before this pathway was discovered there was no explanation for the surprising observation that the degradation of many proteins requires ATP. (Recall from Section 2.6 that hydrolysis of a peptide bond is a thermodynamically favorable reaction.)

Aaron Ciechanover (1947–), Avram Hershko (1937–), and Irwin Rose (1926–) won the 2004 Nobel Prize in Chemistry "for the discovery of ubiquitin-mediated protein degradation."



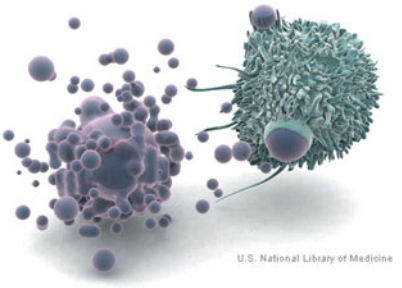
▲ **Figure 17.28**
Ubiquitination and hydrolysis of a protein. Ubiquitination enzymes catalyze the attachment of numerous molecules of ubiquitin to the protein targeted for degradation. The proteasome catalyzes ATP-dependent hydrolysis of the substituted protein, releasing peptides and ubiquitin.

BOX 17.4 APOPTOSIS—PROGRAMMED CELL DEATH

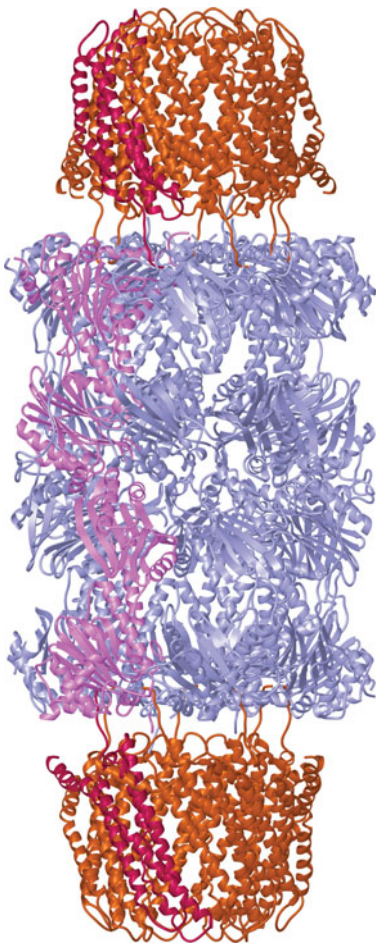
Apoptosis (often pronounced with the second *p* silent) is a series of morphological changes in a cell that leads to its death. The changes include a decrease in cell volume, damage to the plasma membrane, swelling of mitochondria, and fragmentation of chromatin. Surplus and harmful cells are removed principally by the action of proteases.

Some cells die normally during development or in the regulation of antibody production. Others die as a result of diseases or from faulty apoptosis (as in some neurodegenerative diseases). As a result of apoptosis, vesicles containing cellular contents form and are engulfed by neighboring cells. Some of the protein contents of the vesicles can be saved and reused by the other cells.

All eukaryotes have a similar set of endogenous enzymes responsible for cell death. These enzymes (first described as being involved in apoptosis in 1993) include about a dozen proteases called caspases—meaning cysteine-containing hydrolases that act on the carboxyl side of aspartate residues.



▲ **Apoptosis.** The drawing depicts vesicles from a dead apoptotic cell (purple) being taken up by a white blood cell (green). [Courtesy of the United States National Library of Medicine.]



▲ **Proteasome from yeast (*Saccharomyces cerevisiae*). (a)** Side view. The complete proteasome consists of two seven-member rings of β subunits (blue) with their active protease sites on the interior of the cylinder. The outer two rings have seven α subunits (purple). **(b)** Top view. Ubiquinated proteins enter the cylinder through a pore at the top or bottom of the structure. [PDB 1FNT]

17.6 Amino Acid Catabolism

Amino acids obtained from the degradation of endogenous proteins or from the diet can be used for the biosynthesis of new proteins. Amino acids not needed for the synthesis of proteins are catabolized in order to make use of their nitrogen and their carbon skeletons. The first step in amino acid degradation is often removal of the α-amino group. Next, the carbon chains are altered in specific ways for entry into the central pathways of carbon metabolism. We first consider the metabolic fates of the various carbon skeletons. In the next section we examine the metabolism of the ammonia arising from amino acid degradation. These catabolic pathways are present in all species but they are especially important in animals since amino acids are a significant part of fuel metabolism.

Removal of the α-amino group of an amino acid occurs in several ways. The amino acid usually undergoes transamination with α-ketoglutarate to form an α-keto acid and glutamate. The glutamate is oxidized to α-ketoglutarate and ammonia by the action of mitochondrial glutamate dehydrogenase. The net effect of these two reactions is the release of α-amino groups as ammonia and the formation of NADH and α-keto acids. This is the reverse of the pathway shown in Figure 17.8A.

$$\text{Amino acid} + \alpha\text{-Ketoglutarate} \rightleftharpoons \alpha\text{-Keto acid} + \text{Glutamate}$$

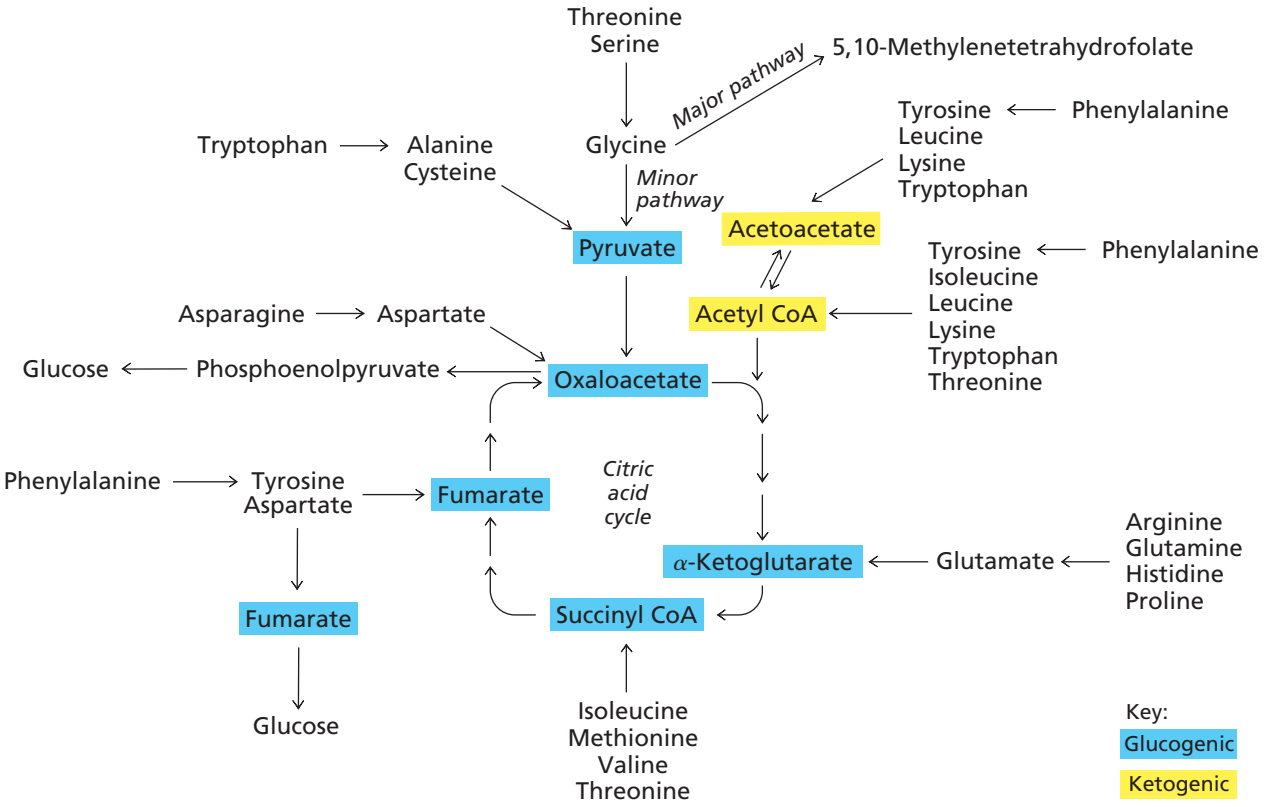
$$\text{Glutamate} + \text{NAD}^{\oplus} + \text{H}_2\text{O} \rightleftharpoons \alpha\text{-Ketoglutarate} + \text{NADH} + \text{H}^{\oplus} + \text{NH}_4^{\oplus}$$

$$\text{Sum: Amino acid} + \text{NAD}^{\oplus} + \text{H}_2\text{O} \rightleftharpoons \alpha\text{-Keto acid} + \text{NADH} + \text{H}^{\oplus} + \text{NH}_4^{\oplus} \quad \textbf{(17.4)}$$

The amide groups of glutamine and asparagine are hydrolyzed by specific enzymes—glutaminase and asparaginase, respectively—to produce ammonia and the corresponding dicarboxylic amino acids glutamate and aspartate. Ammonia from amides and amino groups that is not used in biosynthesis reactions is excreted.

Once the amino groups have been removed, the carbon chains of the 20 amino acids can be degraded. Some are degraded to one of four citric acid cycle intermediates while others are degraded to pyruvate, and still others to acetyl CoA or acetoacetate (Figure 17.29). Each amino acid follows its own route to one or more of these seven compounds.

While all these products can be oxidized to CO_2 and H_2O they can also have other metabolic fates. Amino acids that are degraded to pyruvate or citric acid cycle intermediates are called glucogenic because they can directly supply the pathway of gluconeogenesis. Those that form acetyl CoA or acetoacetate can contribute to the formation of fatty acids or ketone bodies and are called ketogenic. Some amino acids are both glucogenic



▲ **Figure 17.29**
Degradation of amino acids. The carbon skeletons of amino acids are converted to pyruvate, acetoacetate, acetyl CoA, or citric acid cycle intermediates.

and ketogenic because different parts of their carbon chains form different products. The distinction between glucogenic and ketogenic products is important in animals since amino acids are significant fuel metabolites in the diet. Animals do not possess a direct pathway leading from acetyl CoA to glucose and the production of excess acetyl CoA stimulates formation of ketone bodies (Section 16.11). The distinction between glucogenic and ketogenic products is less important in bacteria, protists, fungi, and plants since they can convert acetyl CoA to oxaloacetate via the glyoxylate pathway (Section 13.7). In these organisms, acetyl CoA is glucogenic.

In this section, we examine the pathways of amino acid degradation beginning with the simplest routes. Our aim is to show how the carbon atoms of each amino acid reach "glucogenic" metabolites (pyruvate and citric acid cycle intermediates) or "ketogenic" metabolites (acetyl CoA and acetoacetate). The ultimate fates of these metabolites depend on the species and are covered in earlier chapters.

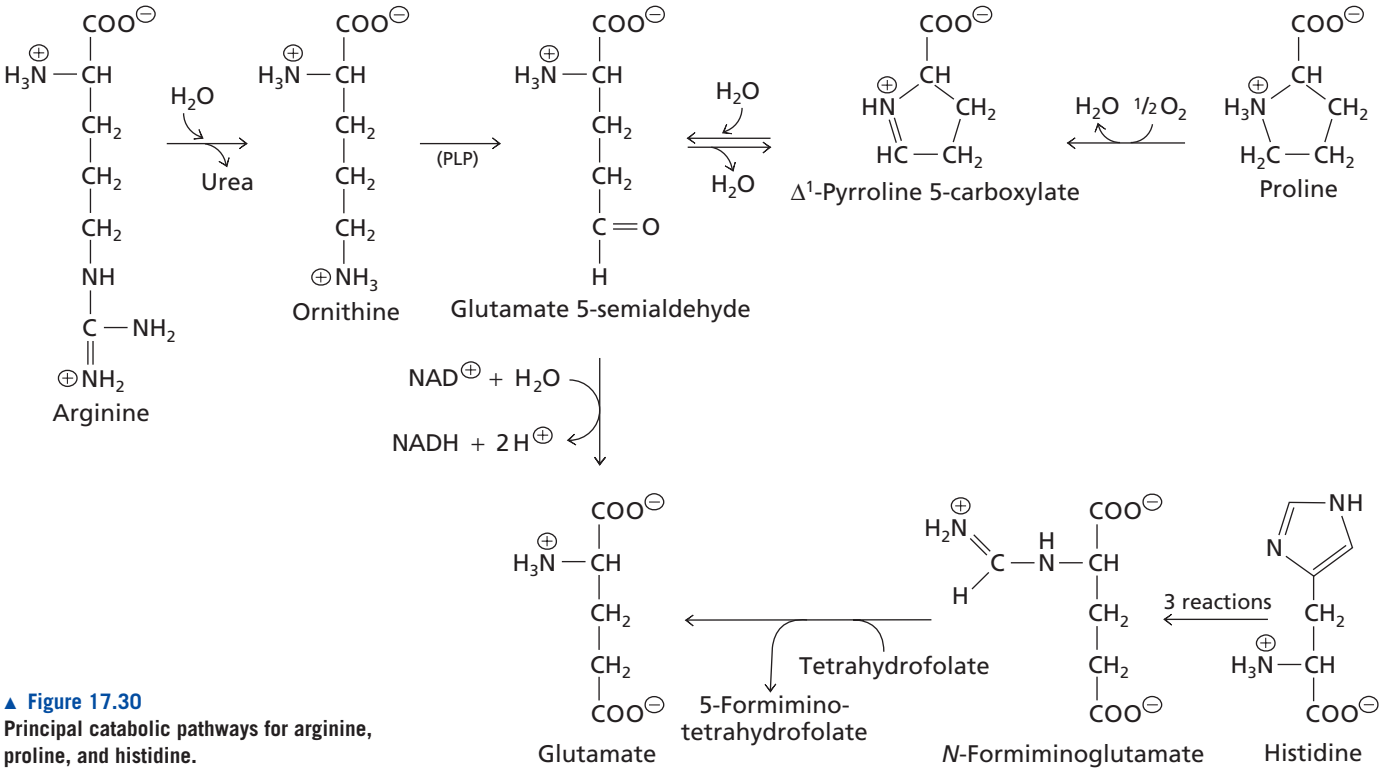
A. Alanine, Asparagine, Aspartate, Glutamate, and Glutamine

Alanine, aspartate, and glutamate are synthesized by reversible transamination reactions (Sections 17.3A,C,D). The breakdown of these three amino acids involves their re-entry into the pathways from which their carbon skeletons arose. Alanine gives rise to pyruvate, aspartate to oxaloacetate, and glutamate to α-ketoglutarate by reversal of the original transamination reactions. All three amino acids are glucogenic since aspartate and glutamate are converted to citric acid cycle intermediates and alanine is converted to pyruvate.

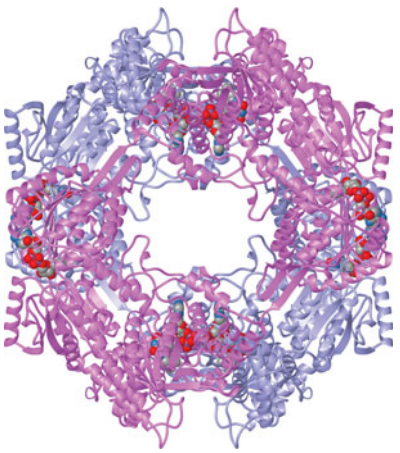
The degradation of both glutamine and asparagine begins with their hydrolysis to glutamate and aspartate, respectively. Thus, glutamine and asparagine are both glucogenic. The hydrolysis reactions are catalyzed by specific enzymes—asparaginase (Box 17.1) and glutaminase.

B. Arginine, Histidine, and Proline

The pathways for the degradation of arginine, histidine, and proline converge on glutamate (Figure 17.30). In the case of arginine and proline, the degradation pathways resemble the biosynthesis pathways. Arginine degradation commences with the reaction catalyzed by arginase. The ornithine produced is transaminated to glutamate 5-semialdehyde, which is oxidized to form glutamate.



▲ **Figure 17.30**
Principal catabolic pathways for arginine, proline, and histidine.



▲ **Proline utilization A flavoprotein.** This enzyme from *Bradyrhizobium japonicum* combines the first two enzymes in the proline degradation pathway into a large complex consisting of six subunits of bifunctional proteins. The two identical subunits of one core dimer are colored blue and purple and the entire structure consists of three such dimers arranged in a circle. The bound FAD and $NAD^{\oplus}$ coenzymes are shown as space-filling models. This enzyme presumably confers a selective advantage over species containing two separate enzymes so why hasn't it evolved in eukaryotes? [PDB 3HAZ]

Proline is converted to glutamate in three steps. The first step is an oxidation reaction catalyzed by the FAD-containing enzyme proline dehydrogenase. The electron acceptor is sometimes molecular oxygen although other acceptors can be used. The product of the first reaction is Δ^1-pyrroline 5-carboxylate (P5C) that exists in equilibrium with the open-chain form, glutamate 5-semialdehyde. Glutamate 5-semialdehyde is converted to glutamate by the action of $NAD^{\oplus}$-dependant P5C dehydrogenase. Note that the conversion of Δ^1-pyrroline 5-carboxylate to glutamate 5-semialdehyde is spontaneous as in the proline synthesis pathway (Section 17.3D).

The first two enzymes in this pathway are separate enzymes in all eukaryotes and most bacteria but in some species of bacteria the two genes for these enzymes have fused to create a bifunctional hexameric protein that catalyzes both reactions. This is kinetically advantageous since the intermediates (Δ^1-pyrroline 5-carboxylate and glutamate 5-semialdehyde) do not dissociate from the complex before being converted to glutamate.

The major pathway for histidine degradation also produces glutamate. Histidine undergoes nonoxidative deamination, hydration, and ring opening to form *N*-formiminoglutamate. The formimino moiety ($-CH=NH_2^{\oplus}$) is then transferred to tetrahydrofolate, forming 5-formiminotetrahydrofolate and glutamate. 5-Formiminotetrahydrofolate is then enzymatically deaminated to form 5,10-methenyltetrahydrofolate. The one-carbon (methenyl) group of this tetrahydrofolate derivative can be used in pathways such as pyrimidine synthesis (Section 18.6).

C. Glycine and Serine

There are two pathways for the breakdown of serine (Figure 17.31). A small amount of serine is converted directly to pyruvate by the action of serine dehydratase, a PLP-dependent enzyme. Most serine, however, is converted to glycine by the action of serine hydroxymethyltransferase. This is the same reaction that results in synthesis of glycine in the biosynthesis pathway (Figure 17.16) and it is a reaction that produces 5,10-methylene tetrahydrofolate (5,10-methylene THF).

Some glycine can be converted to serine by the reverse reaction of serine hydroxylmethyltransferase and the glycine carbon atoms can end up in pyruvate when the serine molecules are deaminated. However, the major pathway for degradation of glycine in all species is conversion to $NH_4^{\oplus}$ and $HCO_3^{\ominus}$ by the glycine cleavage system.

Catalysis by the glycine cleavage system requires an enzyme complex containing four nonidentical subunits. PLP, lipoamide, and FAD are prosthetic groups, and $NAD^{\oplus}$ and tetrahydrofolate (THF) are cosubstrates. Initially, glycine is decarboxylated and the $—CH_2—NH_3^{\oplus}$ group is transferred to lipoamide. Then, $NH_4^{\oplus}$ is released, and the remaining one-carbon group is transferred to tetrahydrofolate to form 5,10-methylenetetrahydrofolate (5,10-methylene THF). Reduced lipoamide is oxidized by FAD and $FADH_2$ reduces the mobile carrier $NAD^{\oplus}$.

As shown in Figure 17.32 the glycine cleavage system is another example of a lipoamide swinging arm mechanism similar in principle to that of pyruvate dehydrogenase (Section 13.1). Although glycine breakdown is reversible *in vitro*, the glycine cleavage system catalyzes an irreversible reaction in cells. The irreversibility of the reaction sequence is due in part to the K_m values for the products ammonia and methylenetetrahydrofolate that are far greater than the concentrations of these compounds *in vivo*.

D. Threonine

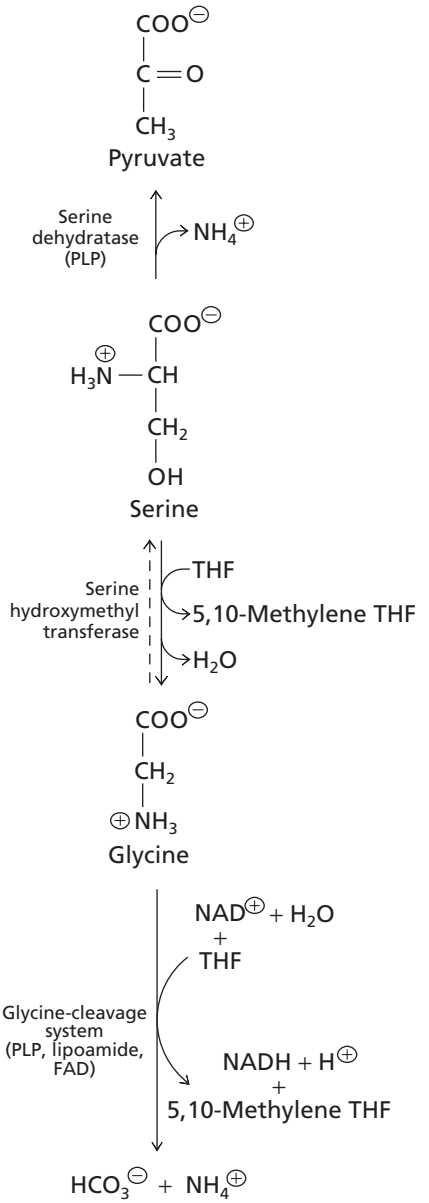
There are several routes for the degradation of threonine. In the major pathway, threonine is oxidized to 2-amino-3-ketobutyrate in a reaction catalyzed by threonine dehydrogenase (Figure 17.33). 2-Amino-3-ketobutyrate can undergo thiolysis to form acetyl CoA and glycine. Another route for threonine catabolism is cleavage to acetaldehyde and glycine by the action of threonine aldolase. Threonine aldolase is actually a minor activity of serine hydroxymethyltransferase in many tissues and organisms. Acetaldehyde can be oxidized to acetate by the action of acetaldehyde dehydrogenase and acetate can be converted to acetyl CoA by acetyl-CoA synthetase.

A third route for threonine catabolism in mammals is deamination to α-ketobutyrate. This reaction is catalyzed by serine dehydratase, the same enzyme that catalyzes the conversion of serine to pyruvate. This reaction produces α-ketobutyrate for synthesis of isoleucine in most species (Section 17.3C). α-Ketobutyrate can be converted to propionyl CoA in the degradative pathway and propionyl CoA is a precursor of the citric acid cycle intermediate succinyl CoA (Section 16.7 F). Threonine can thus produce either succinyl CoA or glycine + acetyl CoA depending on the pathway by which it is degraded.

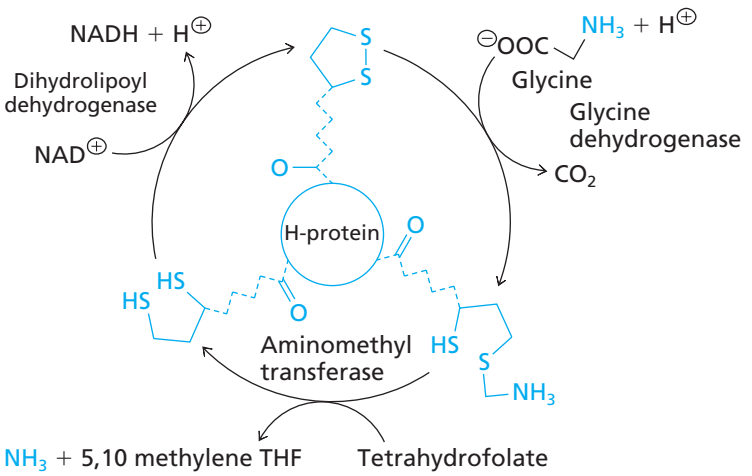
E. The Branched Chain Amino Acids

Leucine, valine, and isoleucine are degraded by related pathways (Figure 17.34). The same three enzymes catalyze the first three steps in all pathways. The first step, transamination, is catalyzed by branched chain amino acid transaminase.

The second step in the catabolism of branched chain amino acids is catalyzed by branched chain α-keto acid dehydrogenase. In this reaction, the branched chain α-keto acids undergo oxidative decarboxylation to form branched chain acyl CoA molecules one carbon atom shorter than the precursor α-keto acids. Branched chain α-keto acid dehydrogenase is a multienzyme complex containing lipoamide and thiamine pyrophosphate (TPP) and requires $NAD^{\oplus}$ and coenzyme A. Its catalytic mechanism is similar to

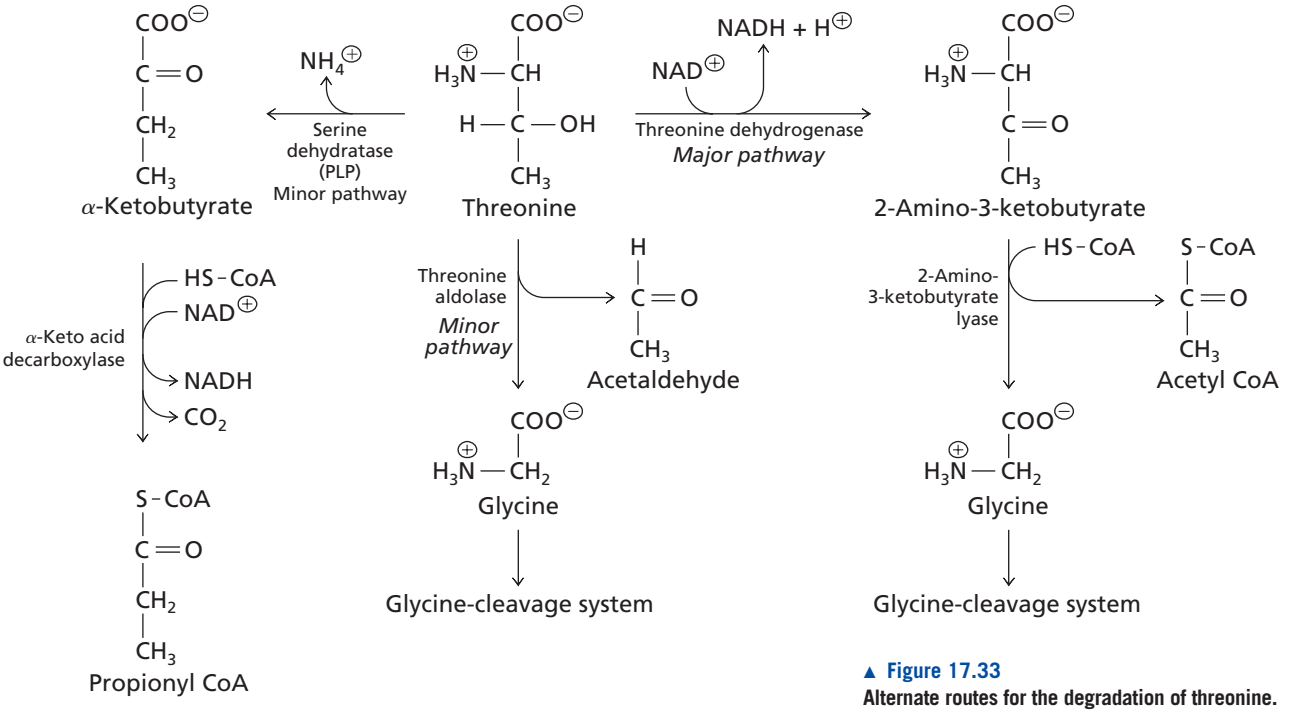


▲ **Figure 17.31**
Catabolism of serine and glycine.

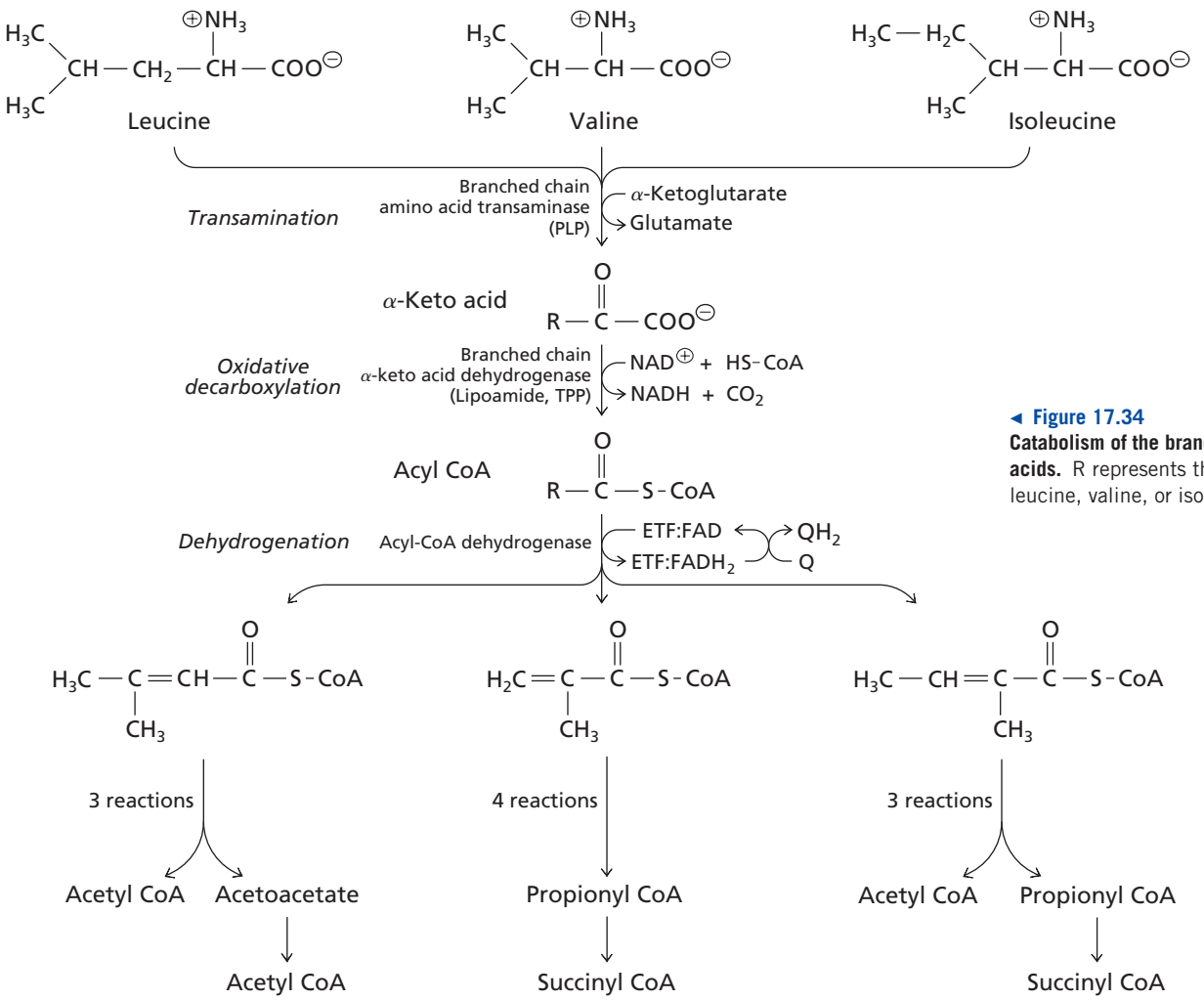


The pathway from propionyl CoA to succinyl CoA is shown in detail in Figure 16.22.

◀ **Figure 17.32**
Glycine cleavage system. A lipoamide swinging arm is attached to the core structural component (H-protein). The swinging arm visits the active sites of the three enzymes of the pathway.



▲ **Figure 17.33**
Alternate routes for the degradation of threonine.



◀ **Figure 17.34**
Catabolism of the branched chain amino acids. R represents the side chain of leucine, valine, or isoleucine.

that of the pyruvate dehydrogenase complex (Section 13.1) and the α-ketoglutarate dehydrogenase complex (Section 13.3#4), and it contains the same dihydrolipoamide dehydrogenase (E_3) subunits as those found in the other two dehydrogenase complexes.

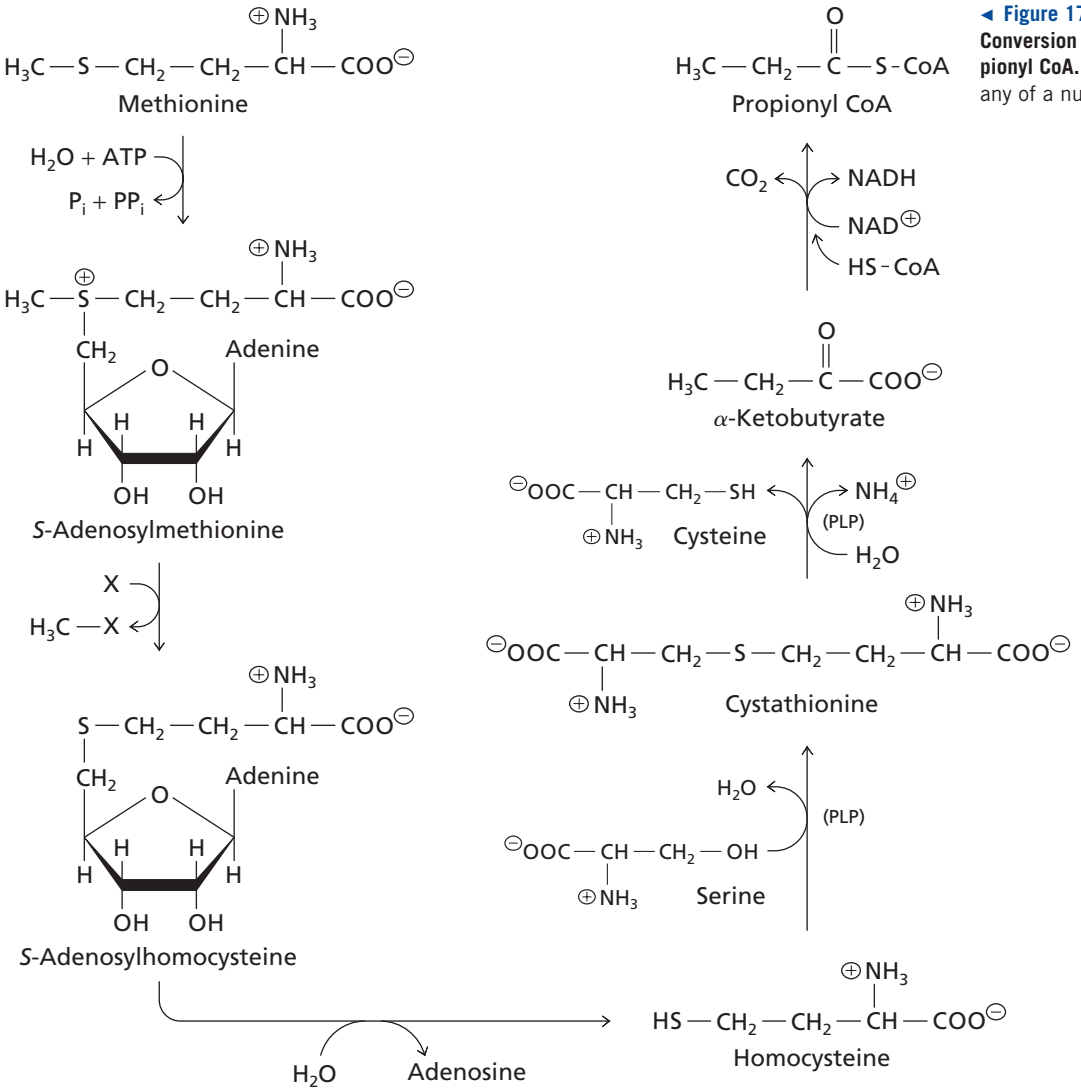
Branched chain acyl CoA molecules are oxidized by an FAD-containing acyl-CoA dehydrogenase in a reaction analogous to the first step in fatty acyl CoA oxidation (Figure 16.19). The electrons removed in this oxidation step are transferred via the electron transferring flavoprotein (ETF) to ubiquinone (Q).

At this point, the steps in the catabolism of branched chain amino acids diverge. All the carbons of leucine are ultimately converted to acetyl CoA, so leucine is purely ketogenic. Valine is ultimately converted to propionyl CoA. As in the degradation of threonine, propionyl CoA is converted to succinyl CoA that enters the citric acid cycle. Valine is glucogenic. The isoleucine degradation pathway leads to both propionyl CoA and acetyl CoA. Isoleucine is therefore both glucogenic (via succinyl CoA formed from propionyl CoA) and ketogenic (via acetyl CoA). Thus, although the initial steps in the degradation of the three branched chain amino acids are similar, their carbon skeletons have different fates—at least in animals.

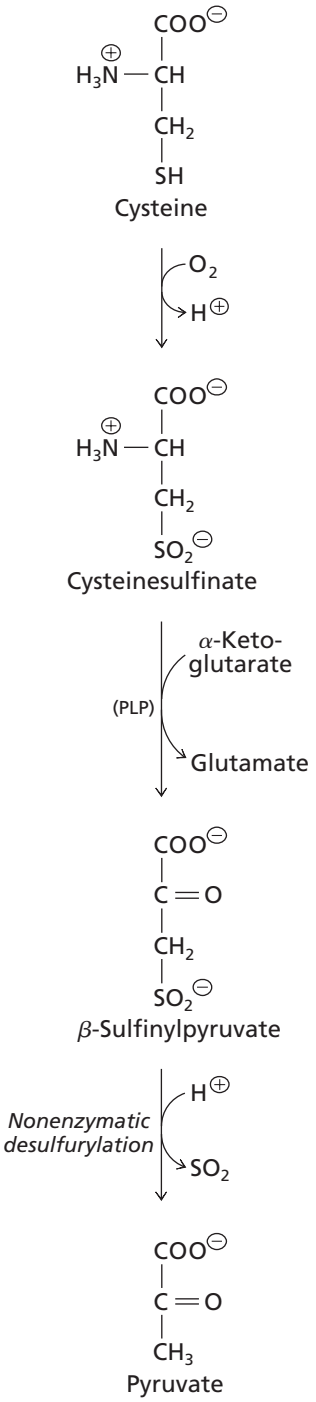
Remember that the distinction between ketogenic and glucogenic pathways is only relevant in animals because all other species can convert acetyl CoA to glucose.

F. Methionine

One major role of methionine is conversion to the activated methyl donor *S*-adenosylmethionine (Section 7.3). Transfer of the methyl group from *S*-adenosylmethionine to a methyl acceptor leaves *S*-adenosylhomocysteine that is degraded by hydrolysis to homocysteine and adenosine (Figure 17.35). Homocysteine can either be methylated by



◀ **Figure 17.35**
Conversion of methionine to cysteine and propionyl CoA. X in the second step represents any of a number of methyl-group acceptors.



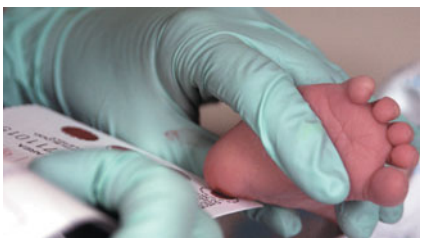
▲ **Figure 17.36**
Conversion of cysteine to pyruvate.

BOX 17.5 PHENYLKETONURIA IS A DEFECT IN TYROSINE FORMATION

One of the most common disorders of amino acid metabolism is phenylketonuria (PKU). The disease is caused by a mutation in the gene that encodes phenylalanine hydroxylase (*PAH* gene on chromosome 12q: OMIN MIN=261600). Affected individuals are unable to convert dietary phenylalanine to tyrosine so the blood of children with this disease contains very high levels of phenylalanine and low levels of tyrosine. Instead of being converted to tyrosine, phenylalanine is metabolized to phenylpyruvate in the reverse of the transamination reaction shown in Figure 17.20. (Transamination of phenylalanine does not occur in unaffected individuals because the K_m of the transaminase for phenylalanine is much higher than the normal concentration of phenylalanine.) Elevated levels of phenylpyruvate and its derivatives inhibit brain development.

Newborns are routinely screened for PKU by testing for elevated levels of phenylpyruvate in the urine or of phenylalanine in the blood during the first days after birth. Phenylalanine hydroxylase-deficient individuals often develop normally if the dietary intake of phenylalanine is strictly limited during the first decade of life. Some women with PKU must restrict their dietary intake of phenylalanine during pregnancy to ensure proper fetal development. Elevated levels of phenylalanine are also observed in individuals with deficiencies in dihydropteridine reductase or 4a-carbinolamine dehydratase or defects in the biosynthesis of tetrahydrobiopterin because each of these disorders results in impairment of the hydroxylation of phenylalanine.

Control of diet can successfully treat PKU but the restrictions exclude many natural, protein-rich foods such as meat, fish, milk, bread, and cake. The food of this strict diet is not appetizing. Tests have been performed by feeding PKU victims an enzyme that catalyzes degradation of phenylalanine to ammonia and a nontoxic carbon product. This enzyme does not fully replace dietary restriction of phenylalanine but it may increase a patient's tolerance for protein-containing foods.

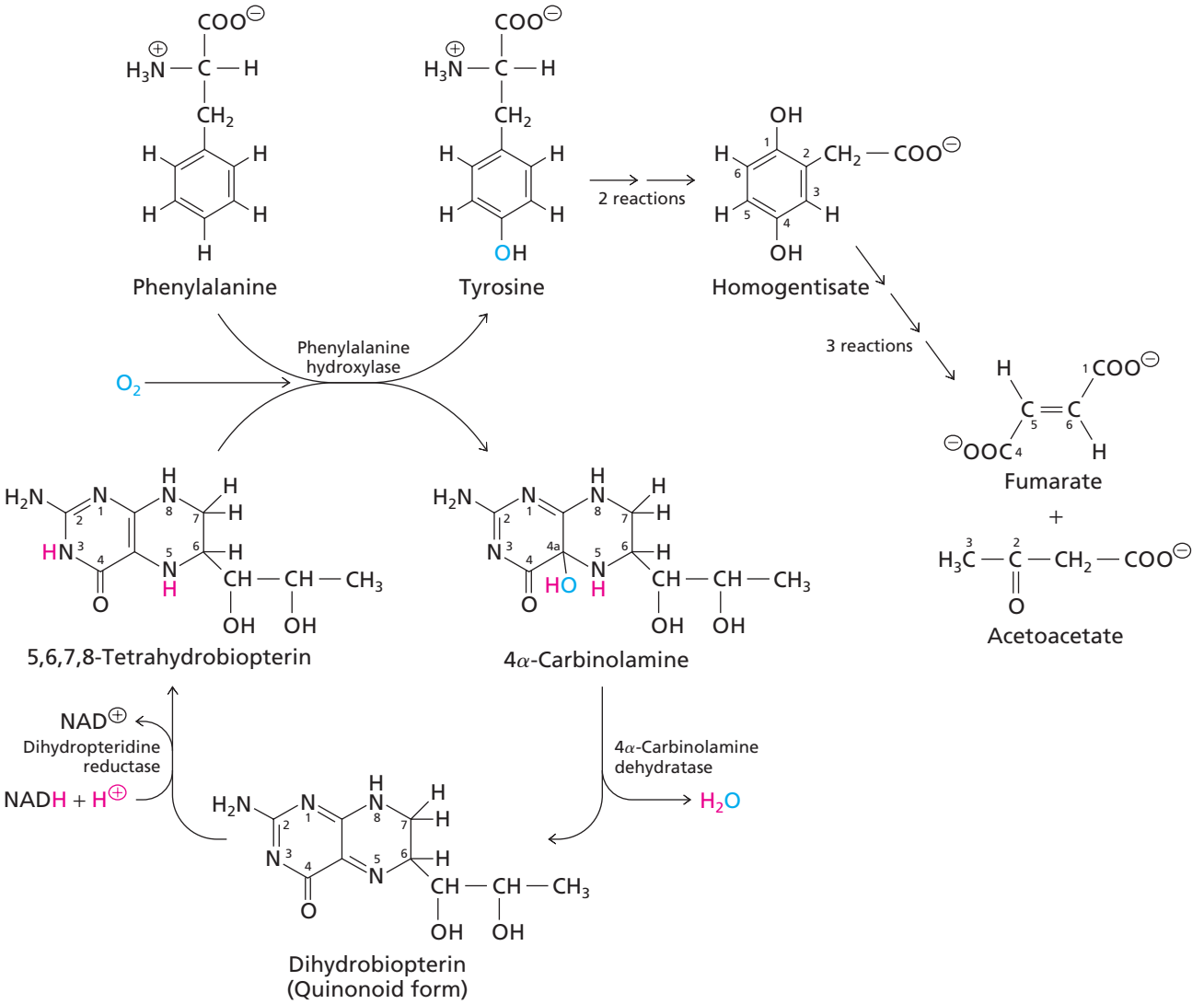


▲ Newborn infants are tested for phenylketonuria by analyzing blood drawn from the heel of the foot.

5-methyltetrahydrofolate to form methionine or it can react with serine to form cystathionine that can be cleaved to cysteine and α-ketobutyrate. We encountered this series of reactions earlier as part of a pathway for the formation of cysteine (Figure 17.18). By this pathway, mammals can form cysteine using a sulfur atom from the essential amino acid methionine. α-Ketobutyrate is converted to propionyl CoA by the action of an α-keto acid dehydrogenase. Propionyl CoA can be further metabolized to succinyl CoA, so methionine is glucogenic.

G. Cysteine

The major route of cysteine catabolism is a three-step pathway leading to pyruvate (Figure 17.36). Therefore, cysteine is glucogenic. Cysteine is first oxidized to cysteinesulfinate that loses its amino group by transamination to form β-sulfinylpyruvate. Nonenzymatic desulfurylation produces pyruvate.



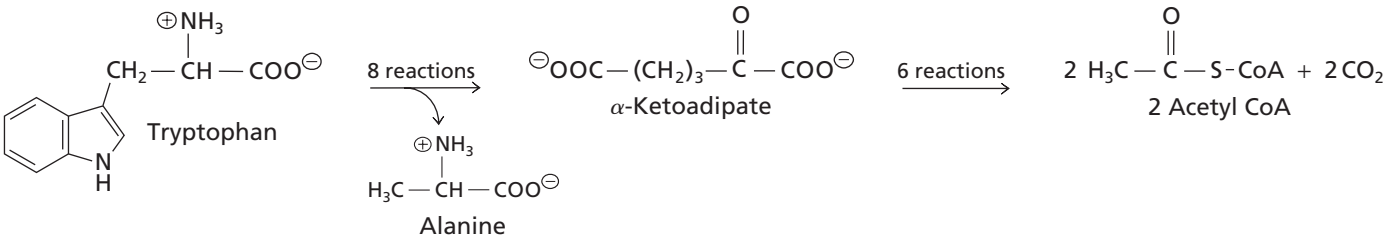
▲ **Figure 17.37**
Conversion of phenylalanine and tyrosine to fumarate and acetoacetate. The tetrahydrobiopterin cofactor is regenerated via dehydration and NADH-dependent reduction.

H. Phenylalanine, Tryptophan, and Tyrosine

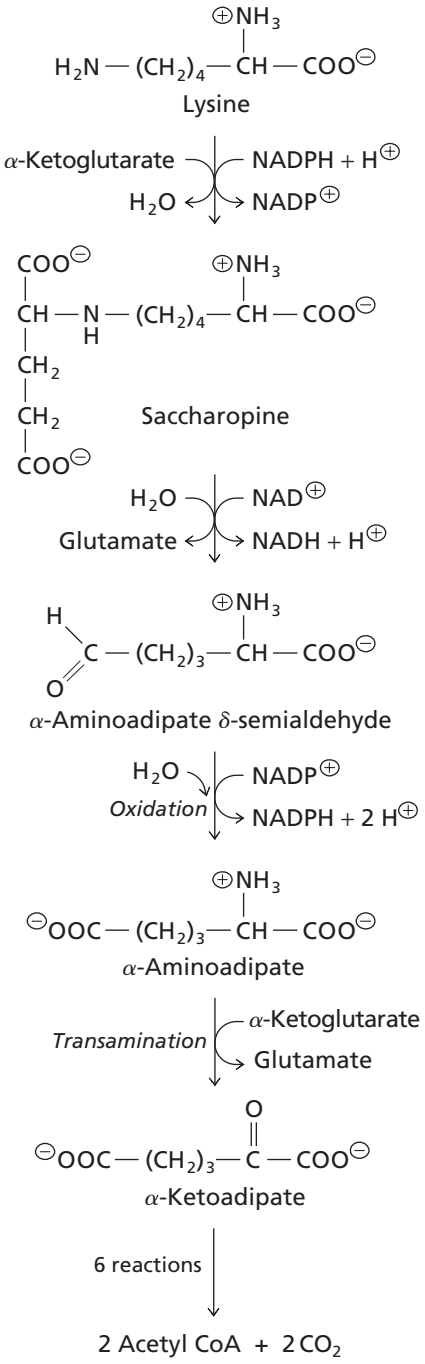
The aromatic amino acids share a common pattern of catabolism. In general, the pathways begin with oxidation, followed by removal of nitrogen by transamination or hydrolysis and then ring opening coupled with oxidation.

The conversion of phenylalanine to tyrosine, catalyzed by phenylalanine hydroxylase, is an important step in the catabolism of phenylalanine (Figure 17.37). It also serves as a source of tyrosine in animals since they lack the normal chorismate pathway for tyrosine synthesis. The phenylalanine hydroxylase reaction requires molecular oxygen and the reducing agent tetrahydrobiopterin. One oxygen atom from O_2 is incorporated into tyrosine and the other is converted to water.

Tetrahydrobiopterin is regenerated in two steps. 4α-Carbinolamine dehydratase catalyzes the dehydration of the first oxidized product and prevents its isomerization to an inactive form in which the side chain is on C-7, not C-6. Dihydropteridine reductase catalyzes the reduction of the resulting quinonoid dihydrobiopterin to 5,6,7,8-tetrahydrobiopterin in a reaction that requires NADH. Tetrahydrobiopterin is also a reducing agent in the biosynthesis of nitric oxide from arginine (Section 17.4C).



▲ **Figure 17.38**
Conversion of tryptophan to alanine and acetyl CoA.



▲ **Figure 17.39**
Conversion of lysine to acetyl CoA.

Further degradation of uric acid is described in Section 18.8.

The catabolism of tyrosine begins with the removal of its α-amino group in a transamination reaction with α-ketoglutarate. Subsequent oxidation steps lead to ring opening and eventually to the final products, fumarate and acetoacetate. This fumarate is cytosolic and is converted to glucose. Acetoacetate is a ketone body. Thus, tyrosine is both glucogenic and ketogenic.

The indole ring system of tryptophan has a more complex degradation pathway that includes two ring-opening reactions. The major route of tryptophan catabolism in the liver and many microorganisms leads to α-ketoadipate and ultimately to acetyl CoA (Figure 17.38). Alanine, produced early in tryptophan catabolism, is transaminated to pyruvate. Thus, the catabolism of tryptophan is both ketogenic and glucogenic.

I. Lysine

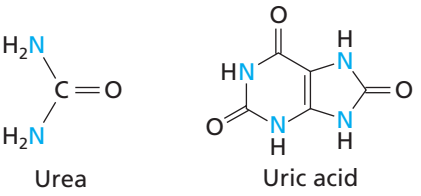
The main pathway for the degradation of lysine generates the intermediate saccharopine, the product of the condensation of α-ketoglutarate with lysine (Figure 17.39). Sequential oxidation reactions produce α-aminoadipate that loses its amino group by transamination with α-ketoglutarate to become α-ketoadipate. α-Ketoadipate is subsequently converted to acetyl CoA by the same steps that occur in the degradation of tryptophan. Like leucine, lysine is ketogenic (these two are the only common amino acids that are purely ketogenic).

17.7 The Urea Cycle Converts Ammonia into Urea

High concentrations of ammonia are toxic to cells. Different organisms have evolved different strategies for eliminating waste ammonia. The nature of the excretory product depends on the availability of water. In many aquatic organisms, ammonia diffuses directly across the cell membranes and is diluted by the surrounding water. This route is inefficient in large terrestrial multicellular organisms and the buildup of ammonia inside internal cells must be avoided.

Most terrestrial vertebrates convert waste ammonia to urea, a less toxic product (Figure 17.40). Urea is an uncharged and highly water-soluble compound produced in the liver and carried in the blood to the kidneys where it is excreted as the major solute of urine. (Urea was first described around 1720 as the essential salt of urine. The name "urea" is derived from "urine.") Birds and many terrestrial reptiles convert surplus ammonia to uric acid, a relatively insoluble compound that precipitates from aqueous solution to form a semisolid slurry. Uric acid is also a product of the degradation of purine nucleotides by birds, some reptiles, and primates.

The synthesis of urea occurs almost exclusively in the liver. Urea is the product of a set of reactions called the urea cycle—a pathway discovered by Hans Krebs and Kurt Henseleit in 1932 several years before Krebs discovered the citric acid cycle. Several observations led to the identification of the urea cycle; for example, slices of rat liver can bring about the net conversion of ammonia to urea. Synthesis of urea by these preparations is markedly stimulated when the amino acid ornithine is added and the amount of urea synthesized greatly exceeds the amount of ornithine that is added, suggesting that ornithine acts catalytically. Finally, it was known that high levels of the enzyme arginase occur in the livers of all organisms that synthesize urea.



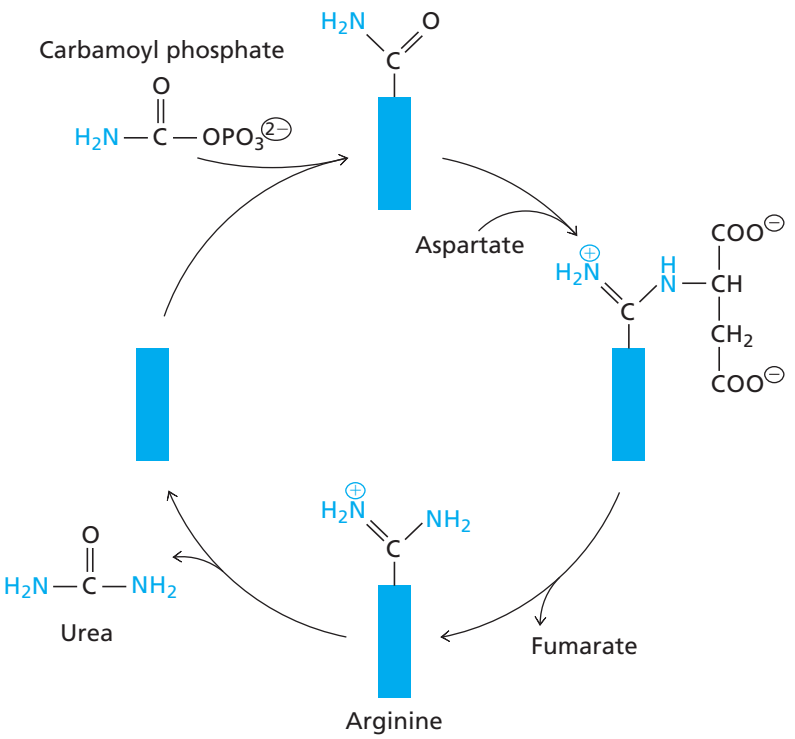
◀ **Figure 17.40**
Urea and uric acid.

A. Synthesis of Carbamoyl Phosphate

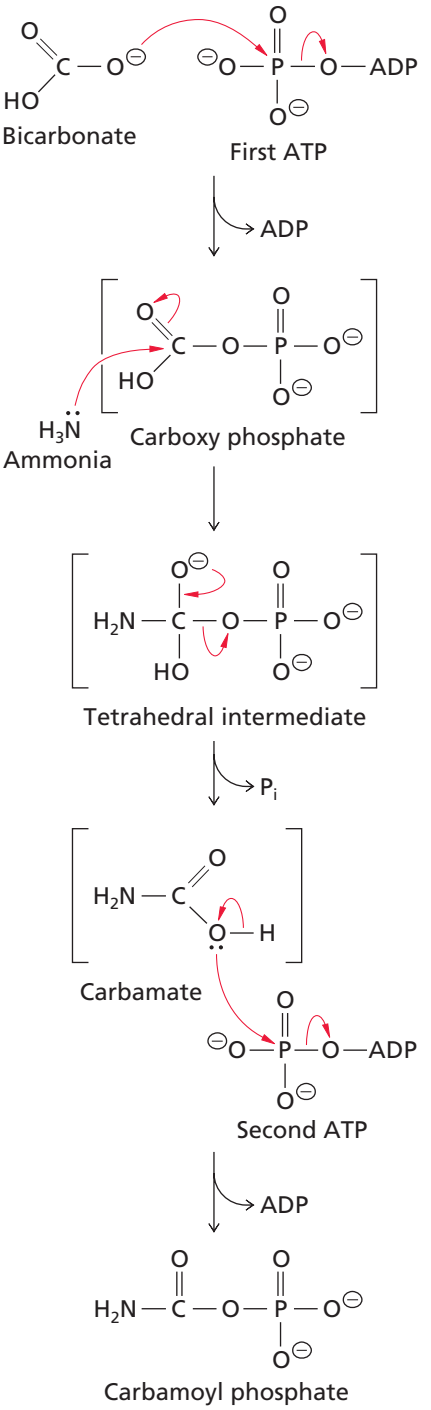
The ammonia released by oxidative deamination of glutamate reacts with bicarbonate to form carbamoyl phosphate. This reaction requires two molecules of ATP and is catalyzed by carbamoyl phosphate synthetase (Figure 17.41). This enzyme is present in all species since carbamoyl phosphate is an essential precursor in pyrimidine biosynthesis and it's also required in the synthesis of arginine in species that don't have a urea cycle. Mammals have two versions of this enzyme. The cytosolic version is called carbamoyl phosphate synthetase II and it uses glutamine rather than ammonia as the nitrogen donor. This is the enzyme used in pyrimidine synthesis (Section 18.3). The bacterial enzymes also use glutamine. The second mammalian version, carbamoyl phosphate I, is the one involved in the urea cycle. It is one of the most abundant enzymes in liver mitochondria accounting for as much as 20% of the protein of the mitochondrial matrix. The nitrogen atom of carbamoyl phosphate is incorporated into urea via the urea cycle.

B. The Reactions of the Urea Cycle

The first nitrogen atom of urea is contributed by carbamoyl phosphate and the second is derived from aspartate. The synthesis of urea takes place while the intermediates are covalently bound to an ornithine skeleton. Ornithine is regenerated when urea is released and it re-enters the urea cycle. Thus, ornithine acts catalytically in the synthesis of urea (Figure 17.42). The carbon, nitrogen, and oxygen atoms of ornithine are not



▲ **Figure 17.42**
The urea cycle. The blue rectangular box represents ornithine.



▲ **Figure 17.41**
Synthesis of carbamoyl phosphate catalyzed by carbamoyl phosphate synthetase I. The reaction involves two phosphoryl-group transfers. First, nucleophilic attack by bicarbonate on ATP produces carboxy phosphate and ADP. Next, ammonia reacts with carboxy phosphate, forming a tetrahedral intermediate. Elimination of a phosphate group produces carbamate. A second phosphoryl-group transfer from another ATP forms carbamoyl phosphate and ADP. Structures in brackets remain enzyme bound during the reaction.

BOX 17.6 DISEASES OF AMINO ACID METABOLISM

Hundreds of human metabolic diseases involving single-gene defects (often termed inborn errors of metabolism) have been discovered. Many are due to defects in the breakdown of amino acids. We have already discussed phenylketonuria, the defect in tyrosine formation from phenylalanine (Box 17.5). A few more examples are mentioned here. Defects in some pathways are severe and even life-threatening; defects in other pathways can result in less severe symptoms. The results indicate that some amino acid degradation pathways are almost dispensable whereas others are essential for survival following birth.

Alkaptonuria

The first metabolic disease to be characterized as a genetic defect was alkaptonuria, a rare disease in which one of the intermediates in the catabolism of phenylalanine and tyrosine (homogentisate) accumulates (Figure 17.37). A deficiency of homogentisate dioxygenase, the enzyme that catalyzes oxidative cleavage of this intermediate, prevents further metabolism of this catabolite. The gene is *HGD* on chromosome 3 (OMIM MIM=203500). Solutions of homogentisate turn dark on standing because this compound is converted to a pigment. Alkaptonuria was recognized by observing the darkening of urine. Individuals with alkaptonuria are prone to develop arthritis, but it is not known how the metabolic defect produces this complication; possibly it is from the deposit of pigments in bones and connective tissues.

Cystinuria

If there is a defect in kidney transport of cysteine and the basic amino acids, then cysteine accumulates in blood and oxidizes to cystine producing a condition called cystinuria. Cystine has a low solubility and forms calculi. Patients suffering from cystinuria drink large amounts of water to dissolve these stones or are given compounds that react with cystine to form soluble derivatives. (See OMIM MIM=220100.)

Gyrate Atrophy

A defect in ornithine transaminase activity causes the metabolic disease gyrate atrophy of the choroid and retina of the eyes. The affected gene is *OAT* on chromosome 10 (OMIM MIM=258870). Gyrate atrophy leads to tunnel vision and later to blindness. The progress of this disorder can be slowed by restricting the dietary intake of arginine or by the administration of pyridoxine.

Maple Syrup Urine Disease

Patients suffering from maple syrup urine disease excrete urine that smells like maple syrup. The disease is caused by a genetic defect at the second step in catabolism of branched chain amino acids—the step catalyzed by the branched chain α-keto acid dehydrogenase complex. Those afflicted with this disease have short lives unless they follow a diet very low in branched chain amino acids. (OMIM MIM=248600)

Nonketotic Hyperglycinemia (Glycine Encephalopathy)

Defects in the enzyme complex that catalyzes glycine cleavage lead to the accumulation of large amounts of glycine in body fluids. This is the main biochemical symptom of a disease called nonketotic hyperglycinemia. Most individuals with this disorder have severe mental deficiencies and die in infancy. The severity of the disease indicates the crucial importance of the glycine cleavage system. (OMIM MIM=605899)

KEY CONCEPT

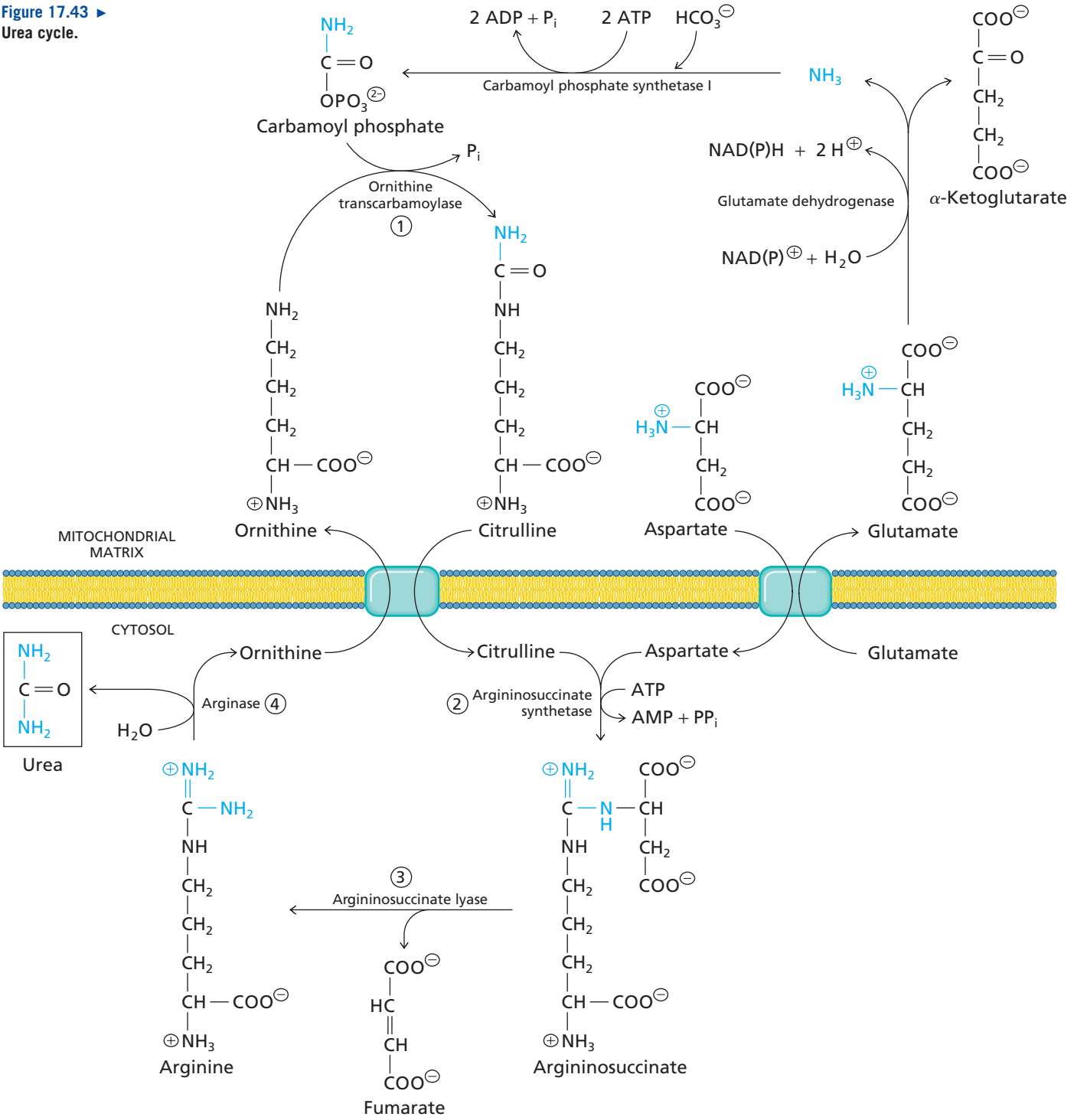
All species need to eliminate ammonia produced by degradation reactions. Some can excrete it directly while others have to convert it to less toxic compounds that are subsequently excreted.

exchanged in the urea cycle. Its role as a catalyst is more obvious than the role of oxaloacetate in the citric acid cycle (Section 13.3) but the principle is the same.

The actual urea cycle reactions are more complex than the simple scheme shown in Figure 17.42. This is because the first reaction occurs in the mitochondrial matrix and the other three occur in the cytosol (Figure 17.43). Two transport proteins connecting the mitochondrial matrix and the cytosol are required: the citrulline–ornithine exchanger and the glutamate-aspartate translocase.

1. The cycle begins when carbamoyl phosphate reacts in the mitochondrion with ornithine to form citrulline in a reaction catalyzed by ornithine transcarbamoylase. This step incorporates the nitrogen atom originating from ammonia into citrulline; citrulline thus contains half the nitrogen destined for urea. Citrulline is then transported out of the mitochondrion in exchange for cytosolic ornithine.
2. The second nitrogen atom destined for urea comes from aspartate and is incorporated when citrulline condenses with aspartate to form argininosuccinate in the cytosol. This ATP-dependent reaction is catalyzed by argininosuccinate synthetase. Most aspartate in cells originates in mitochondria although aspartate is sometimes generated in the cytosol. Mitochondrial aspartate enters the cytosol in exchange for cytosolic glutamate. (This translocase reaction is part of the malate–aspartate shuttle we described in Section 14.12.)
3. Argininosuccinate is cleaved nonhydrolytically to form arginine plus fumarate in an elimination reaction catalyzed by argininosuccinate lyase. Arginine is the immediate

Figure 17.43 ►
Urea cycle.



precursor of urea. (Together, the second and third steps of the urea cycle exemplify a strategy for donating the amino group of aspartate. We will encounter this strategy twice more in the next chapter as part of purine biosynthesis. The key processes are a nucleoside triphosphate–dependent condensation, followed by the elimination of fumarate.)

4. Finally, the guanidinium group of arginine is hydrolytically cleaved to form ornithine and urea in a reaction catalyzed by arginase. Arginase has a pair of Mn^{2+} ions in its active site and this binuclear manganese cluster binds a molecule of water forming a nucleophilic hydroxide ion that attacks the guanidinium carbon atom of arginine. The ornithine generated by the action of arginase is transported

into the mitochondrion where it reacts with carbamoyl phosphate to support continued operation of the urea cycle.

The overall reaction for urea synthesis is

$$NH_3 + HCO_3^{\ominus} + \text{Aspartate} + 3\ \text{ATP} \longrightarrow \text{Urea} + \text{Fumarate} + 2\ \text{ADP} + 2\ P_i + \text{AMP} + PP_i \qquad (17.5)$$

The two nitrogen atoms of urea are derived from ammonia and aspartate. The carbon atom of urea comes from bicarbonate. Four equivalents of ATP are consumed per molecule of urea synthesized. Three molecules of ATP are converted to two ADP and one AMP during the formation of one molecule of urea and the hydrolysis of inorganic pyrophosphate accounts for cleavage of the fourth phosphoanhydride bond.

The carbon skeleton of fumarate is converted to glucose and CO_2. Cytosolic fumarate does not enter the citric acid cycle (which occurs in mitochondria) but instead is hydrated to malate by the action of a cytosolic fumarase. Malate is oxidized to oxaloacetate by the action of malate dehydrogenase and oxaloacetate enters the pathway of gluconeogenesis. This fate is shared by the fumarate produced during tyrosine degradation (Section 17.6H).

C. Ancillary Reactions of the Urea Cycle

The reactions of the urea cycle convert equal amounts of nitrogen from ammonia and from aspartate into urea. Many amino acids can function as amino-group donors via transamination reactions with α-ketoglutarate to form glutamate. Glutamate can undergo either transamination with oxaloacetate to form aspartate or deamination to form ammonia. Both glutamate dehydrogenase and aspartate transaminase are abundant in liver mitochondria and catalyze near-equilibrium reactions. The concentrations of ammonia and aspartate must be approximately equal for efficient synthesis of urea and elimination of nitrogen.

Consider the theoretical case of a relative surplus of ammonia (Figure 17.44a). In this situation, the near-equilibrium reaction catalyzed by glutamate dehydrogenase would proceed in the direction of glutamate formation. Elevated concentrations of glutamate would then result in increased flux to aspartate through aspartate transaminase, also a near-equilibrium step. In contrast, when excess aspartate is present the net flux in the reactions catalyzed by glutamate dehydrogenase and aspartate transaminase would occur in the opposite direction to provide ammonia for urea formation (Figure 17.44b).

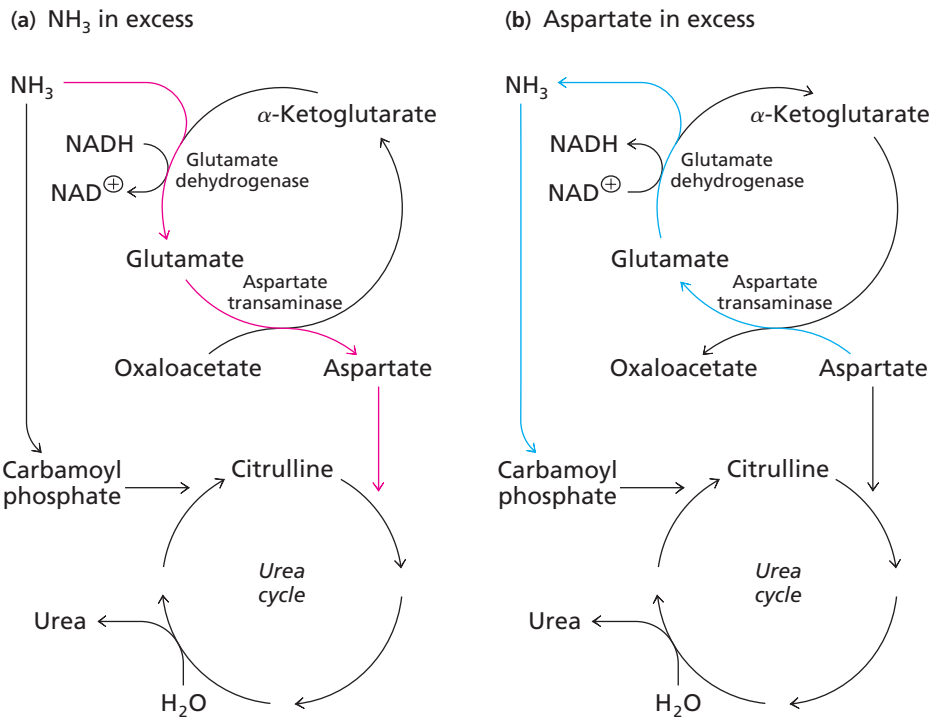
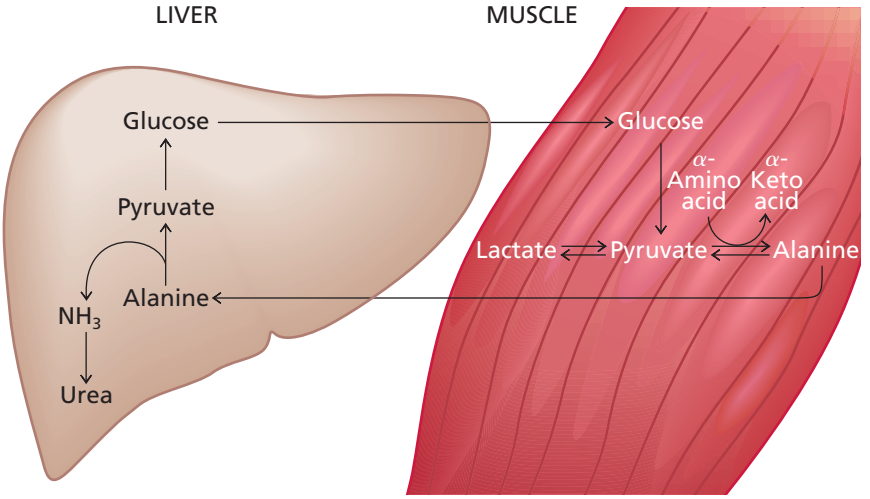


Figure 17.44 ▶
Balancing the supply of nitrogen for the urea cycle. Two theoretical situations are described: **(a)** NH_3 in extreme excess and **(b)** aspartate in extreme excess. Flux through the near-equilibrium reactions catalyzed by glutamate dehydrogenase and aspartate transaminase reverses, depending on the relative amounts of ammonia and amino acids.



◀ **Figure 17.45**
Glucose–alanine cycle.

Some amino acids are deaminated in muscle, not in the liver. Glycolysis—a major source of energy in muscle—produces pyruvate. The transfer of amino groups from α-amino acids to pyruvate generates large amounts of alanine. Alanine travels through the bloodstream to the liver where it is deaminated back to pyruvate. The amino group is used for urea synthesis and the pyruvate is converted to glucose by gluconeogenesis. Recall that neither of these pathways operates in muscle. Glucose can return to the muscle tissue. Alternatively, pyruvate can be converted to oxaloacetate that becomes the carbon chain of aspartate—the metabolite that donates one of the nitrogen atoms of urea. The exchange of glucose and alanine between muscle and liver is called the glucose–alanine cycle (Figure 17.45) and it provides an indirect means for muscle to eliminate nitrogen and replenish its energy supply.

17.8 Renal Glutamine Metabolism Produces Bicarbonate

The body often produces acids as metabolic end products. The resulting anions are eliminated in the urine and the protons remain in the body. One example is β-hydroxybutyric acid, a ketone body that is produced in massive amounts during uncontrolled diabetes mellitus. Another example is sulfuric acid produced during catabolism of the sulfur-containing amino acids cysteine and methionine. These acid metabolites dissociate to give protons and the corresponding anion, β-hydroxybutyrate or sulfate ($SO_4^{2\ominus}$). The blood has an effective buffer system for the protons—they react with bicarbonate to produce CO_2 that is eliminated by the lungs and H_2O (Figure 17.46). While this system effectively neutralizes the excess hydrogen ions it does so at the cost of depleting blood bicarbonate. Bicarbonate is replenished by glutamine catabolism in the kidneys.

In the kidneys, the two nitrogen atoms of glutamine are removed by the sequential action of glutaminase and of glutamate dehydrogenase to produce α-ketoglutarate$^{2\ominus}$ and 2 $NH_4^{\oplus}$.

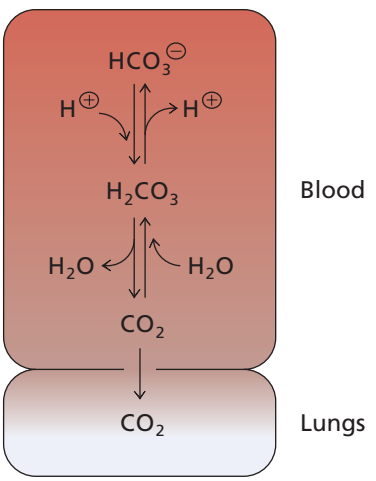
$$\text{Glutamine} \longrightarrow \longrightarrow \alpha\text{-Ketoglutarate}^{2\ominus} + 2\ NH_4^{\oplus} \qquad \textbf{(17.6)}$$

Two molecules of the divalent anion α-ketoglutarate can be converted to one molecule of neutral glucose and four molecules of bicarbonate. The α-ketoglutarate is converted to glucose by oxidation to oxaloacetate, leading to gluconeogenesis. The overall process (ignoring ATP involvement) is

$$\underset{\text{Glutamine}}{2\ C_5H_{10}N_2O_3} + 3\ O_2 + 6\ H_2O \longrightarrow$$

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + 4\ HCO_3^{\ominus} + 4\ NH_4^{\oplus} \qquad \textbf{(17.7)}$$

The $NH_4^{\oplus}$ is excreted in the urine and the $HCO_3^{\ominus}$ is added to the venous blood, replacing the bicarbonate lost in buffering metabolic acids. The excreted $NH_4^{\oplus}$ is accompanied in the urine by the anion (e.g., β-hydroxybutyrate or sulfate) of the original acid metabolite.



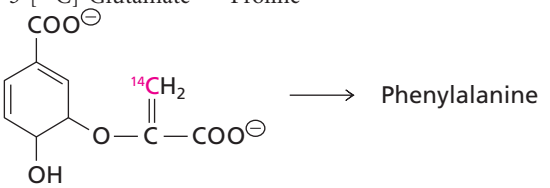
▲ **Figure 17.46**
$H^{\oplus}$ buffering in blood. The $H^{\oplus}$ buffer system leads to bicarbonate loss.

Summary

1. Nitrogen is fixed in only a few species of bacteria by the nitrogenase-catalyzed reduction of atmospheric N_2 to ammonia. Plants and microorganisms can reduce nitrate and nitrite to ammonia.
2. Ammonia is assimilated into metabolites by the reductive amination of α-ketoglutarate to glutamate, catalyzed by glutamate dehydrogenase. Glutamine, a nitrogen donor in many biosynthetic reactions, is formed from glutamate and ammonia by the action of glutamine synthetase.
3. The amino group of glutamate can be transferred to an α-keto acid in a reversible transamination reaction to form α-ketoglutarate and the corresponding α-amino acid.
4. Pathways for the biosynthesis of the carbon skeletons of amino acids begin with simple metabolic precursors such as pyruvate and citric acid cycle intermediates.
5. In addition to their role in protein synthesis, amino acids serve as precursors in a number of other metabolic pathways.
6. Protein molecules in all living cells are continually synthesized and degraded.
7. Amino acids obtained from protein degradation or directly from food can be catabolized. Catabolism often begins with deamination, followed by modification of the remaining carbon chains for entry into the central pathways of carbon metabolism.
8. The pathways for the degradation of amino acids lead to pyruvate, acetyl CoA, or intermediates of the citric acid cycle. Amino acids that are degraded to citric acid cycle intermediates are glucogenic. Those that form acetyl CoA are ketogenic.
9. Most nitrogen in mammals is excreted as urea that is formed by the urea cycle in the liver. The carbon atom of urea is derived from bicarbonate. One amino group is derived from ammonia and the other from aspartate.
10. The metabolism of glutamine in the kidneys produces the bicarbonate needed to neutralize acids produced in the body.

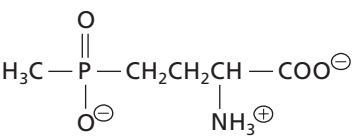
Problems

1. The heterocysts of cyanobacteria contain high concentrations of nitrogenase. These cells have retained photosystem I (PSI) but they do not contain photosystem II (PSII). Why?
2. Write the net equation for converting one molecule of α-ketoglutarate into one molecule of glutamine by assimilating two molecules of ammonia in the following coupled reactions: (a) glutamate dehydrogenase and glutamine synthetase and (b) glutamine synthetase and glutamate synthase. Compare the energy requirements of the two pathways and account for any difference.
3. When ^{15}N-labeled aspartate is fed to animals the ^{15}N label quickly appears in many amino acids. Explain this observation.
4. (a) Three of the 20 common amino acids are synthesized by simple transamination of carbohydrate metabolites. Write the equations for these three transamination reactions.
   (b) One amino acid can be synthesized by reductive amination of a carbohydrate metabolite. Write the equation for this reaction.
5. Animals rely on plants or microorganisms for the incorporation of sulfur into amino acids and their derivatives. However, methionine is an essential amino acid in animals while cysteine is not. If the donor of a sulfur atom in the conversion of homoserine to homocysteine by plants is cysteine, outline the overall path by which sulfur is incorporated into cysteine and methionine in plants and into cysteine in animals.
6. Serine is a source of one-carbon fragments for certain biosynthetic pathways.
   (a) Write the equations that show how two carbon atoms from serine are made available for biosynthesis.
   (b) Assuming that the precursor of serine is produced by glycolysis, which carbon atoms of glucose are the ultimate precursors of these one-carbon fragments?
7. Indicate where the label appears in the product for each of the following precursor–product pairs:
   (a) 3-[^{14}C]-Oxaloacetate ⟶ Threonine
   (b) 3-[^{14}C]-Phosphoglycerate ⟶ Tryptophan
   (c) 3-[^{14}C]-Glutamate ⟶ Proline
   (d)



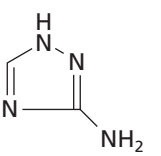
Chorismate

8. (a) PPT (phosphinothricin) is a herbicide that is relatively safe for animals because it is not transported from the blood into the brain and it is rapidly cleared by animal kidneys. PPT effectively inhibits an enzyme in plant amino acid metabolism because it is an analog of the amino acid substrate. What amino acid does PPT resemble?



PPT

   (b) The herbicide aminotriazole inhibits imidazole glycerol phosphate dehydrogenase. What amino acid pathway is inhibited in plants?



Aminotriazole

9. Children with phenylketonuria should not consume the artificial sweetener aspartame (Figure 3.10). Why?

10. (a) A deficiency of α-keto acid dehydrogenase is the most common enzyme abnormality in branched chain amino acid catabolism. Individuals with this disease excrete branched chain α-keto acids. Write the structures of the α-keto acids that would result during the catabolism of leucine, valine, and isoleucine when this enzyme is defective.

(b) A disorder of amino acid catabolism results in the accumulation and excretion of saccharopine. What amino acid pathway is involved and what enzyme is defective?

(c) Citrullinemia is characterized by accumulation of citrulline in the blood and its excretion in the urine. What metabolic pathway is involved and what enzyme is deficient for this disease?

11. Which amino acids yield the following α-keto acids by transamination?

(a) $^{\ominus}OOC-C(=O)-CH_3$

(b) $^{\ominus}OOC-C(=O)-CH_2-COO^{\ominus}$

(c) $^{\ominus}OOC-C(=O)-CH$

(d) $^{\ominus}OOC-C(=O)-CH_2-SH$

12. Animal muscles use two mechanisms to eliminate excess nitrogen generated during the deamination of amino acids. What are the two pathways and why are they necessary?

13. Thiocitrulline and *S*-methylthiocitrulline prevent experimentally induced blood vessel dilation, reduced blood pressure, and shock in animals. What enzyme that produces a gaseous blood vessel dilating messenger is being inhibited? Suggest why these two molecules might act as inhibitors of this enzyme.

14. Why are there so few metabolic diseases associated with defects in amino acid biosynthesis?

15. Pathways for the biosynthesis of the 21st, 22nd, and 23rd amino acids (Section 3.3) are not described in this chapter. Why not? What are the immediate precursors of three additional amino acids?

16. The cost of making amino acids, in ATP equivalents, can be calculated using values for the cost of making each of the precursors plus the cost of each reaction in the amino acid biosynthesis pathway. Assuming that the cost of making glyceraldehyde-3-phosphate is 24 ATP equivalents (Section 15.4C), calculate the cost of making serine (Figure 17.15) and alanine (Figure 17.12). How do your values compare to those in Box 17.3?

Selected Readings

Nitrogen Cycle

Dixon, R., and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* 2:621–631.

Moisander, P. H., Beinart, R. A., Hewson, I., White, A. E., Johnson, K. S., Carlson, C. A., Montoya, J. P., and Zehr, J. P. (2010). Unicellular cyanobacterial distributions broaden the oceanic N_2 fixation domain. *Science* 327:1512–1524.

Montoya, J. P., Holl, C. M., Zehr, J. P., Hansen, A., Villareal, T. A., and Capone, D. G. (2004). High rates of N_2 fixation in the oligotrophic Pacific ocean. *Nature* 430:1027–1031.

Schimpl, J., Petrilli, H. M., and Blöchl, P. E. (2003). Nitrogen binding to the FeMo-cofactor of nitrogenase. *J. Am. Chem. Soc.* 125:15772–15778.

Seefeldt, L. C., Hoffman, B. M., and Dean, D. R. (2009). Mechanism of Mo-dependent nitrogenase. *Annu. Rev. Biochem.* 78:701–722.

Amino Acid Metabolism

Fitzpatrick, P. F. (1999). Tetrahydropterin-dependent amino acid hydroxylases. *Annu. Rev. Biochem.* 68:355–381.

Häussinger, D. (1998). Hepatic glutamine transport and metabolism. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72:43–86.

Huang. X., Holden, H. M., and Raushel, F. M. (2001). Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.* 70:149–180.

Katagiri, M., and Nakamura, M. (2003). Reappraisal of the 20th-century version of amino acid metabolism. *Biochem. Biophys, Res, Comm.* 312:205–208.

Levy, H. L. (1999). Phenylketonuria: old disease, new approach to treatment. *Proc. Natl. Acad. Sci. USA* 96:1811–1813.

Perham, R. N. (2000). Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* 69:961–1004.

Purich, D. L. (1998). Advances in the enzymology of glutamine synthesis. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72:9–42.

Raushel, F. M., Thoden, J. B., and Holden, H. M. (2003). Enzymes with molecular tunnels. *Acc. Chem. Res.* 36:539–548.

Richards, N. G. and Kilberg, M. S. (2006). Asparagine synthetase chemotherapy. *Annu. Rev. Biochem.* 75:629–654.

Scapin, G., and Blanchard, J. S. (1998). Enzymology of bacterial lysine biosynthesis. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72:279–324.

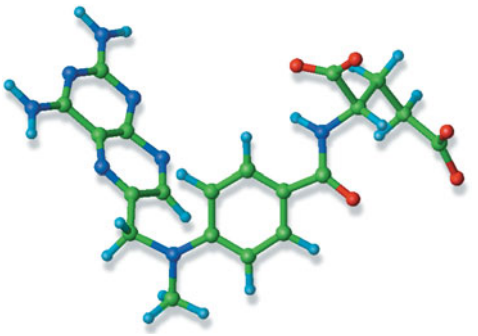
Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds. (1995). *The Metabolic Basis of Inherited Disease*, Vols. 1, 2, and 3. (New York: McGraw-Hill).

Srivastava, D., Schuermann, J. P., White, T. A., Krishnan, N., Sanyal, N., Hura, G. L., Tan, A., Henzl, M. T., Becker, D. F., and Tanner, J. J. (2010). Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*. *Proc. Natl. Acad. Sci. USA* 107:2878–2883.

Wu, G., and Morris, S. M., Jr. (1998). Arginine metabolism: nitric oxide and beyond. *Biochem. J.* 336:1–17.

Zalkin, H., and Smith, J. L. (1998). Enzymes utilizing glutamine as an amide donor. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72:87–144.





Nucleotide Metabolism

[Sven] Furberg, reasoning with marked brilliance and luck from data that were meagre but included his own x-ray studies, got right the absolute three-dimensional configuration of the individual nucleotide . . . "Furberg's nucleotide . . . was absolutely essential to us." Crick told me.

Horace Freeland Judson (1996), *The Eighth Day of Creation*, p. 94.

We have encountered nucleotides and their constituents throughout this book. Nucleotides are probably best known as the building blocks of DNA and RNA; however, as we have seen, they are involved in almost all the activities of the cell either alone or in combination with other molecules. Some nucleotides (such as ATP) function as cosubstrates, and others (such as cyclic AMP and GTP) are regulatory compounds.

One of the components of every nucleotide is a purine or pyrimidine base. The other components are a five-carbon sugar—ribose or deoxyribose—and one or more phosphoryl groups. The standard bases (adenine, guanine, cytosine, thymine, uracil) are almost always found as constituents of nucleotides and polynucleotides. All organisms and cells can synthesize purine and pyrimidine nucleotides because these molecules are essential for information flow. In non-dividing cells, nucleotide biosynthesis is almost exclusively devoted to the production of ribonucleotides for RNA synthesis and various nucleotide cofactors. Deoxyribonucleotides are required for DNA replication in dividing cells and consequently, deoxynucleotide synthesis is closely linked to cell division. Its study is particularly important in modern medicine since synthetic agents that inhibit deoxyribonucleotide synthesis are useful as therapeutic agents against cancer.

We begin this chapter with a description of the biosynthesis of purine and pyrimidine nucleotides. Next, we present the conversion of purine and pyrimidine ribonucleotides to their 2′-deoxy forms, the forms incorporated into DNA. We then discuss how purines and pyrimidines obtained from the breakdown of nucleic acids or extracellularly from food can be incorporated directly into nucleotides—a process called salvage. The salvage pathways conserve energy by recycling the products of nucleic acid turnover. Finally, we examine the biological degradation of nucleotides. The breakdown of purines leads to the formation of potentially toxic compounds that are excreted, whereas the breakdown of pyrimidines leads to readily metabolized products.

18.1 Synthesis of Purine Nucleotides

The identification of the enzymes and intermediates in the pathway for the synthesis of the two purine nucleotides, adenosine 5′-monophosphate (AMP) and guanosine

Top: Methotrexate, one of the most commonly used anticancer drugs. Methotrexate is an analog of folate that inhibits the reaction cycle generating deoxythymidylate for DNA synthesis.

5′-monophosphate (GMP), began with studies of nitrogen metabolism in birds. The major end product of nitrogen metabolism in birds and some reptiles is uric acid (Figure 18.1) rather than urea, as in mammals. Researchers in the 1950s discovered that uric acid and nucleic acid purines arise from the same precursors and reaction sequence. Homogenates of pigeon liver—a tissue in which purines are actively synthesized—were a convenient source of enzymes for studying the steps in purine biosynthesis. The pathway in avian liver has since been found in many other organisms.

When isotopically labeled compounds such $^{13}CO_2$, $H^{13}COO^-$(formate), and $^+H_3N\text{-}CH_2\text{—}^{13}COO^-$ (glycine) were administered to pigeons and rats the result was excretion of labeled uric acid. This uric acid was isolated and chemically degraded to identify the positions of the labeled carbon and nitrogen atoms. The carbon from carbon dioxide was incorporated into C-6, and the carbon from formate into C-2 and C-8 of purines. Ultimately, the sources of the ring atoms were shown to be N-1, aspartate; C-2 and C-8, formate via 10-formyltetrahydrofolate (Section 7.9); N-3 and N-9, amide groups from glutamine; C-4, C-5, and N-7, glycine; and C-6, carbon dioxide. These findings are summarized in Figure 18.2.

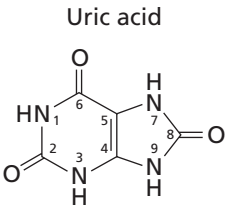
The purine ring structure is not synthesized as a free base but as a substituent of ribose 5-phosphate. The ribose 5-phosphate for purine biosynthesis comes from 5-phosphoribosyl 1-pyrophosphate (PRPP) also known as 5-phosphoribosyl 1-diphosphate. PRPP is synthesized from ribose 5-phosphate and ATP in a reaction catalyzed by ribose-phosphate diphosphokinase (Figure 18.3); PRPP then donates ribose 5-phosphate to serve as the foundation on which the purine structure is built. PRPP is also a precursor for the biosynthesis of pyrimidine nucleotides, although in that pathway it reacts with a preformed pyrimidine to form a nucleotide. PRPP is also used in the nucleotide salvage pathways and in the biosynthesis of histidine (Figure 17.23).

The initial product of the purine nucleotide biosynthetic pathway is inosine 5′-monophosphate (IMP, or inosinate) (Figure 18.4), a nucleotide containing hypoxanthine (6-oxopurine) as its base. The ten-step pathway for the *de novo* synthesis of IMP was discovered in the 1950s by the research groups of John M. Buchanan and G. Robert Greenberg. The painstaking isolation and structural characterization of the intermediates took about ten years.

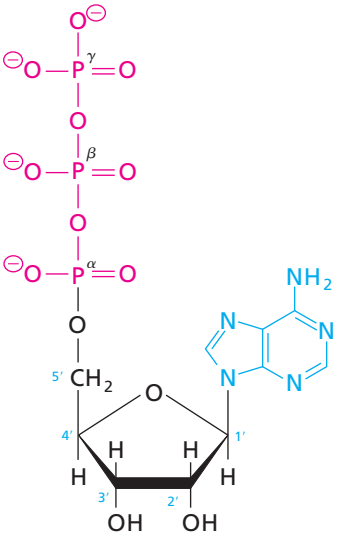
The pathway to IMP is shown in Figure 18.5. It begins with displacement of the pyrophosphoryl group of PRPP by the amide nitrogen of glutamine in a reaction catalyzed by glutamine–PRPP amidotransferase. Note that the configuration of the anomeric carbon is inverted from α to β in this nucleophilic displacement—the β configuration persists in completed purine nucleotides. The amino group of the product, phosphoribosylamine, is then acylated by glycine to form glycinamide ribonucleotide. The mechanism of this reaction, in which an enzyme-bound glycyl phosphate is formed, resembles that of glutamine synthetase that has γ-glutamyl phosphate as an intermediate (Reaction 10.17).

Step 3 consists of transfer of a formyl group from 10-formyltetrahydrofolate to the amino group destined to become N-7 of IMP. In step 4, an amide is converted to an amidine (RHN—C=NH) in an ATP-dependent reaction in which glutamine is the nitrogen donor. Step 5 is a ring-closure reaction that requires ATP and produces an imidazole derivative. CO_2 is incorporated in step 6 by attachment to the carbon that becomes C-5 of IMP. This carboxylation is unusual because it does not require biotin. Bicarbonate is first attached, in an ATP-dependent step, to the amino group that becomes N-3 of IMP. The carboxylated intermediate then undergoes a rearrangement in which the carboxylate group is transferred to the carbon atom that becomes C-5 of the purine ring (Figure 18.6). These steps are catalyzed by two separate proteins in *Escherichia coli* but in eukaryotes they are catalyzed by a multifunctional enzyme. Vertebrate versions of this enzyme transfer the carboxylate group directly to the final position in carboxyaminoimidazole ribonucleotide (CAIR). The vertebrate enzymes are much more efficient. There doesn't seem to be any reason why the enzymes in other species have to undergo a two-step reaction.

The amino group of aspartate is incorporated into the growing purine ring system in steps 7 and 8. First, aspartate condenses with the newly added carboxylate group to form an amide, specifically a succinylocarboxamide. Then adenylosuccinate lyase catalyzes a nonhydrolytic cleavage reaction that releases fumarate. This two-step process

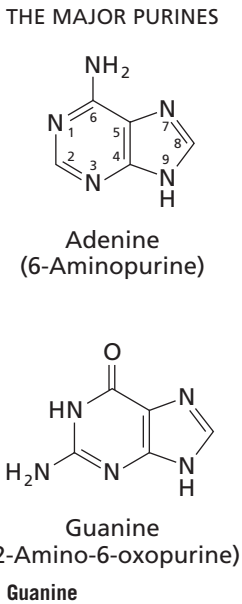


▲ **Figure 18.1**
Uric acid.

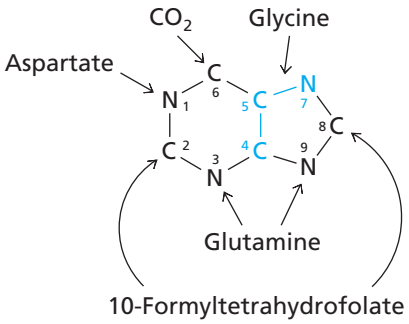


▲ **Structure of adenosine triphosphate (ATP).** The nitrogenous base adenine (blue) is attached to ribose (black). Three phosphoryl groups (red) are bound to the ribose at the 5′ position.

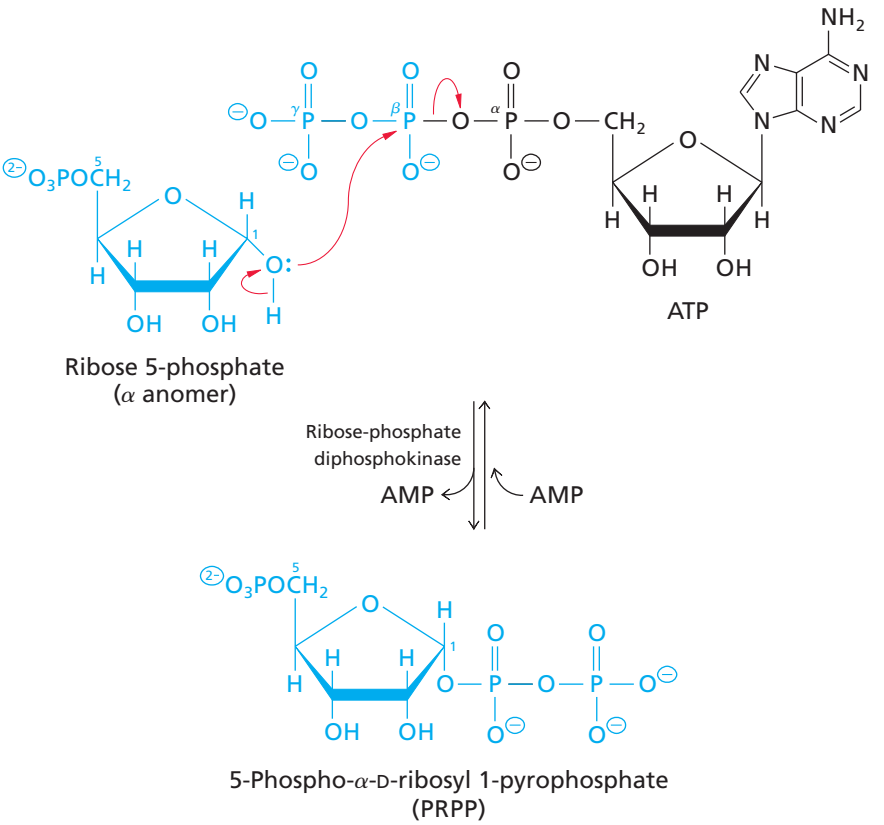
Ribose 5-phosphate is produced by the pentose phosphate pathway (Section 12.4).



▲ **Adenine, Guanine**



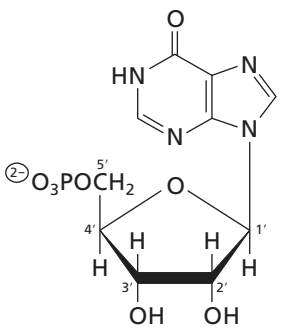
▲ **Figure 18.2**
Sources of the ring atoms in purines synthesized *de novo*.



▲ **Figure 18.3**
Synthesis of 5-phosphoribosyl 1-pyrophosphate (PRPP) from ribose 5-phosphate and ATP. Ribose-phosphate diphosphokinase catalyzes the transfer of a pyrophosphoryl group from ATP to the 1-hydroxyl oxygen of ribose 5-phosphate.



▲ **G. Robert Greenberg (1918–2005).** Greenberg's research group worked out many of the reactions of the purine biosynthesis pathway.



▲ **Figure 18.4**
Inosine 5′-monophosphate (IMP, or inosinate). IMP is converted to other purine nucleotides. Much of the IMP is degraded to uric acid in birds and primates.

results in the transfer of an amino group containing the nitrogen destined to become N-1 of IMP. The two steps are similar to steps 2 and 3 of the urea cycle (Figure 17.43) except that in this case ATP is cleaved to ADP + P_i rather than to AMP + PP_i.

In step 9, which resembles step 3, the cosubstrate 10-formyltetrahydrofolate donates a formyl group (—CH=O) to the nucleophilic amino group of aminoimidazole carboxamide ribonucleotide. The amide nitrogen of the final intermediate then condenses with the formyl group in a ring closure that completes the purine ring system of IMP.

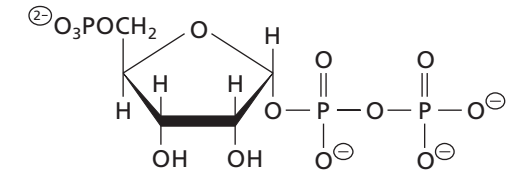
The synthesis of IMP consumes considerable energy. ATP is converted to AMP during the synthesis of PRPP and steps 2, 4, 5, 6, and 7 are driven by the conversion of ATP to ADP. Additional ATP is required for the synthesis of glutamine from glutamate and ammonia (Figure 17.4).

BOX 18.1 COMMON NAMES OF THE BASES

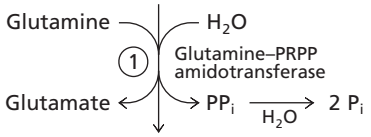
Adenine	from the Greek *adenas*, "gland": first isolated from pancreatic glands (1885)
Cytosine	derived from *cyto-* from the Greek word for "receptacle," referring to cells (1894)
Guanine	originally isolated from "guano" or bird excrement (1850)
Uracil	origin uncertain, possibly from "urea" (1890)
Thymine	first isolated from thymus glands (1894)
Xanthine	from the Greek word for "yellow" (1857)

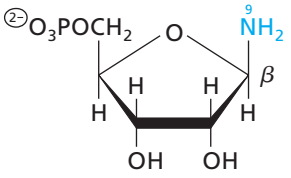
KEY CONCEPT

Nucleotide biosynthesis pathways are energetically expensive.

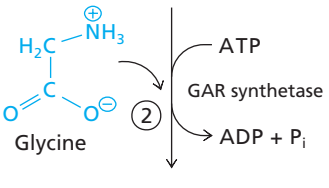


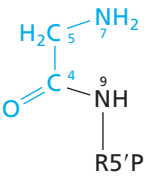
5-Phospho-α-D-ribosyl 1-pyrophosphate (PRPP)



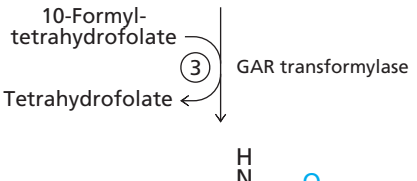


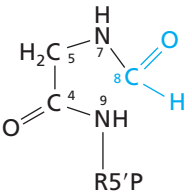
5-Phospho-β-D-ribosylamine (PRA)



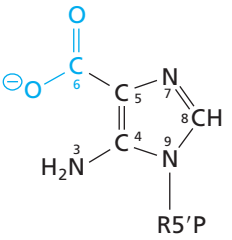


Glycinamide ribonucleotide (GAR)

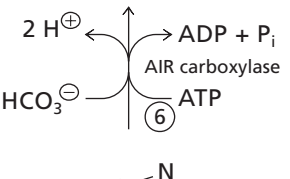


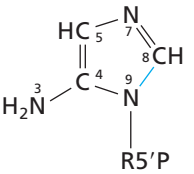


Formylglycinamide ribonucleotide (FGAR)

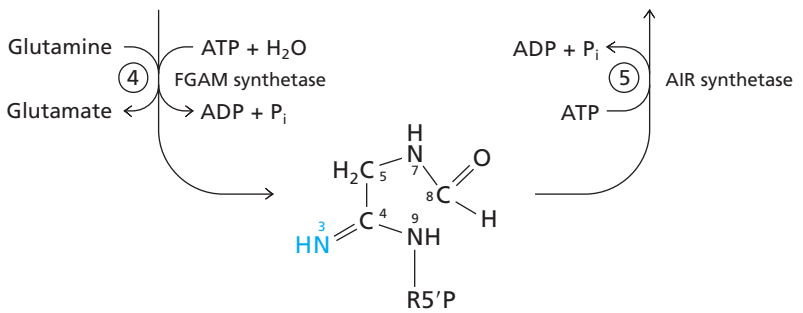


Carboxyaminoimidazole ribonucleotide (CAIR)

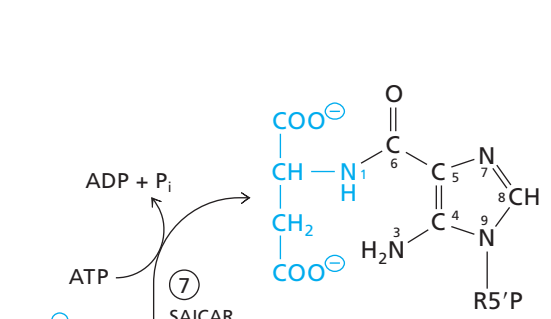


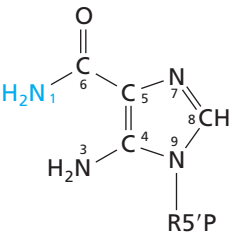


Aminoimidazole ribonucleotide (AIR)

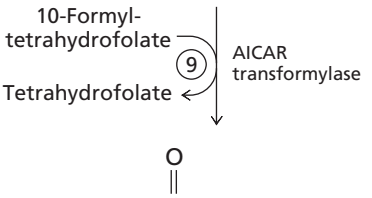


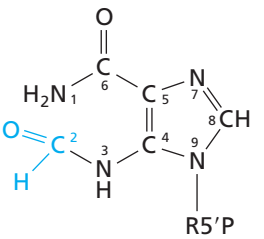
Formylglycinamidine ribonucleotide (FGAM)



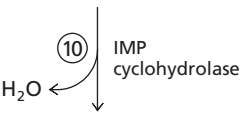


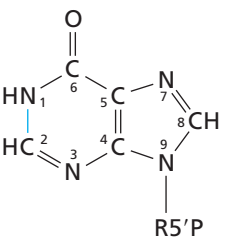
Aminoimidazole carboxamide ribonucleotide (AICAR)





Formamidoimidazole carboxamide ribonucleotide (FAICAR)

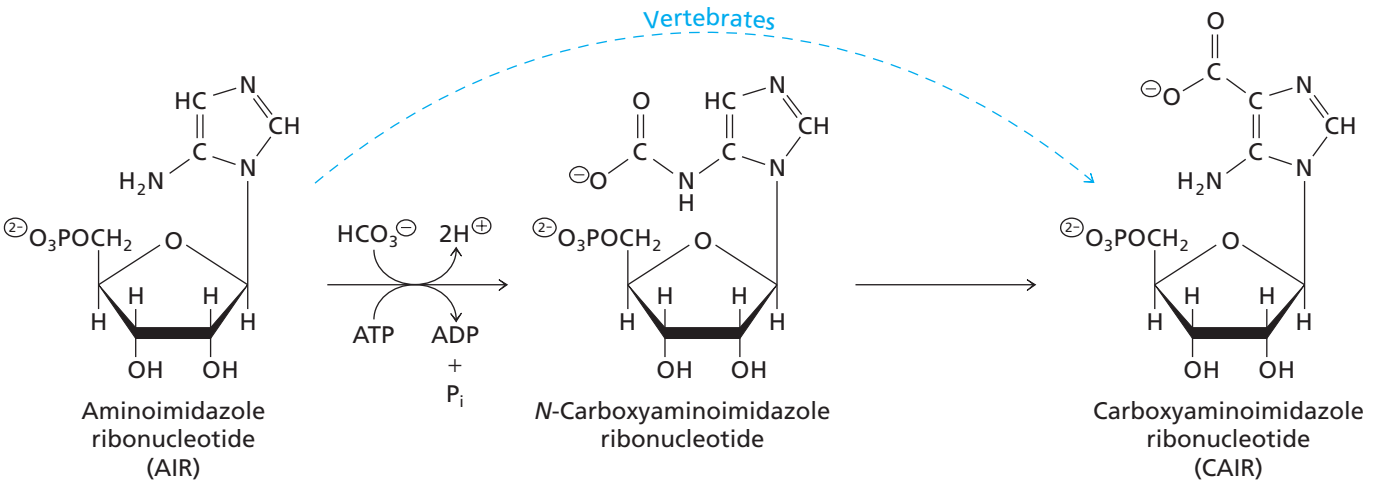




Inosine 5′-monophosphate (IMP)

▲ **Figure 18.5**

Ten-step pathway for the *de novo* synthesis of IMP. R5′P stands for ribose 5′-phosphate. The atoms are numbered according to their positions in the completed purine ring structure.



▲ **Figure 18.6**
***N*-Carboxyaminoimidazole ribonucleotide is sometimes an intermediate in the conversion of AIR to CAIR.**



▲ **John M. ("Jack") Buchanan (1917–2007).** Buchanan's group discovered many of the purine biosynthesis pathway reactions. He and Greenberg were friendly competitors sharing many of their research results.

18.2 Other Purine Nucleotides Are Synthesized from IMP

IMP can be converted to AMP or GMP (Figure 18.7). Two enzymatic reactions are required for each of these conversions. AMP and GMP can then be phosphorylated to their di- and triphosphates by the actions of specific nucleotide kinases (adenylate kinase and guanylate kinase, respectively) and the broadly specific nucleoside diphosphate kinase (Section 10.6).

The two steps that convert IMP to AMP closely resemble steps 7 and 8 in the biosynthesis of IMP. First, the amino group of aspartate condenses with the keto group of IMP in a reaction catalyzed by GTP-dependent adenylosuccinate synthetase. Next, the elimination of fumarate from adenylosuccinate is catalyzed by adenylosuccinate lyase, the same enzyme that catalyzes step 8 of the *de novo* pathway to IMP.

The first step in the conversion of IMP to GMP is the oxidation of C-2 catalyzed by $NAD^{\oplus}$-dependent IMP dehydrogenase. This reaction proceeds by the addition of a molecule of water to the double bond between C-2 and N-3 followed by oxidation of the hydrate. The product of the oxidation is xanthosine monophosphate (XMP). Next, the amide nitrogen of glutamine replaces the oxygen at C-2 of XMP in an ATP-dependent reaction catalyzed by GMP synthetase. The use of GTP as a cosubstrate in the synthesis of AMP from IMP, and of ATP in the synthesis of GMP from IMP, helps balance the formation of the two products.

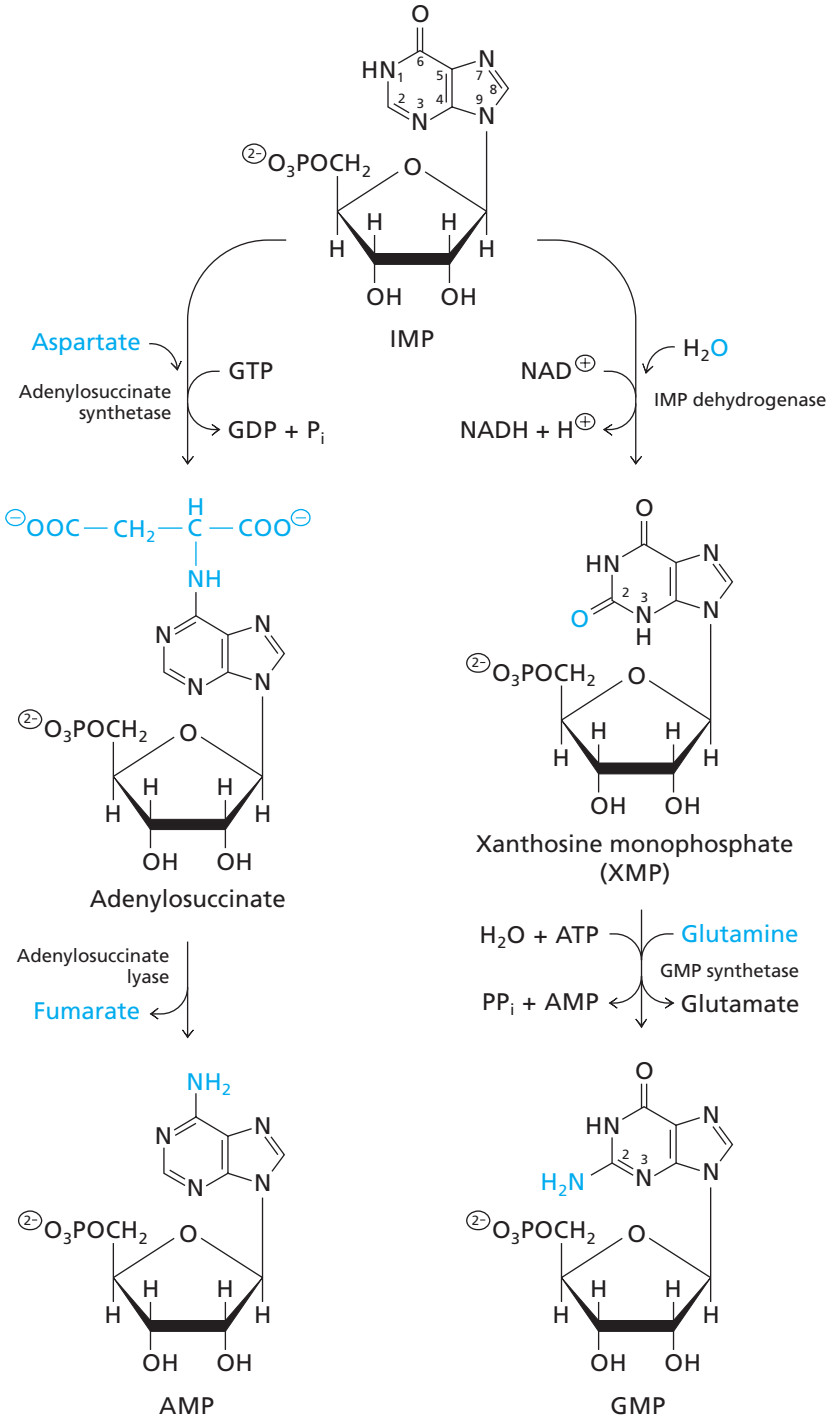
Purine nucleotide synthesis is regulated in cells by feedback inhibition. Several enzymes that catalyze steps in the biosynthesis of purine nucleotides exhibit allosteric behavior *in vitro*. Ribose-phosphate diphosphokinase is inhibited by several purine ribonucleotides but only at concentrations higher than those usually found in the cell. PRPP is a donor of ribose 5-phosphate in more than a dozen reactions so we would not expect PRPP synthesis to be regulated exclusively by the concentrations of purine nucleotides. The enzyme that catalyzes the first committed step in the pathway of purine nucleotide synthesis, glutamine–PRPP amidotransferase (step 1 in Figure 18.5), is allosterically inhibited by 5′-ribonucleotide end products (IMP, AMP, and GMP) at intracellular concentrations. This step appears to be the principal site of regulation of this pathway.

The paths leading from IMP to AMP and from IMP to GMP also appear to be regulated by feedback inhibition. Adenylosuccinate synthetase is inhibited *in vitro* by AMP, the product of this two-step branch. Both XMP and GMP inhibit IMP dehydrogenase. The pattern of feedback inhibition in the synthesis of AMP and GMP is shown in Figure 18.8. Note that the end products inhibit two of the initial common steps as well as steps leading from IMP at the branch point.

18.3 Synthesis of Pyrimidine Nucleotides

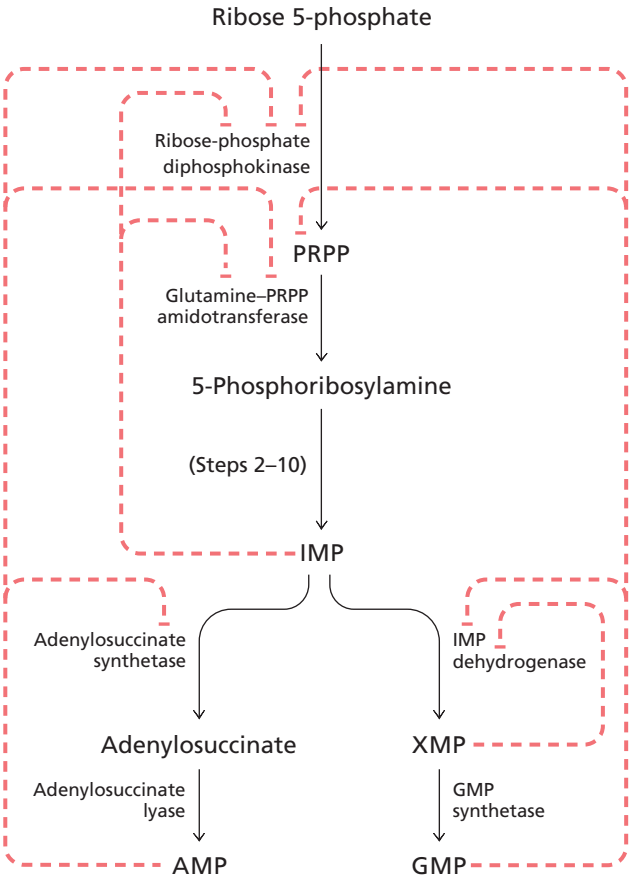
Uridine 5′-monophosphate is the precursor of other pyrimidine nucleotides. The pathway for the biosynthesis of UMP is simpler than the purine pathway and consumes fewer ATP molecules. The pyrimidine ring atoms come from bicarbonate that contributes C-2; the amide group of glutamine (N-3); and aspartate that contributes the remaining atoms (Figure 18.9). C-2 and N-3 are incorporated after formation of the intermediate carbamoyl phosphate.

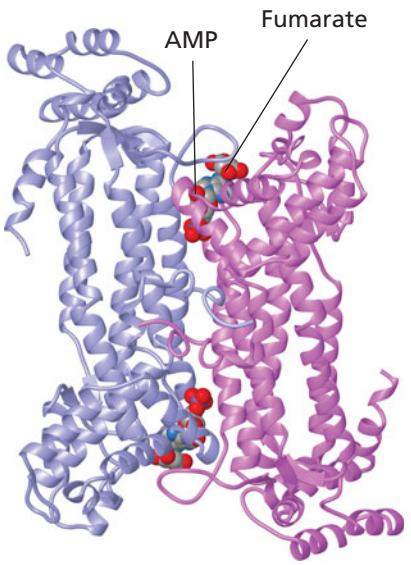
PRPP is required for the biosynthesis of pyrimidine nucleotides but the sugar-phosphate from PRPP is donated after the ring has formed rather than entering the



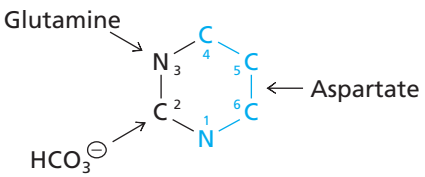
▲ **Figure 18.7**
Pathways for the conversion of IMP to AMP and to GMP.

Figure 18.8 ►
Feedback inhibition in purine nucleotide biosynthesis.





▲ **Adenylosuccinate lyase from *E. coli*.** The enzyme is a homodimer. One subunit is colored blue and the other is purple. This is a mutant enzyme (H171N) showing the two products, AMP and fumarate, bound at the active sites. Adenylosuccinate lyase catalyzes similar steps in the IMP synthesis pathway and in the conversion of IMP to AMP. [PDB 2PTQ]



▲ **Figure 18.9**
Sources of the ring atoms in pyrimidines. The immediate precursor of C-2 and N-3 is carbamoyl phosphate.

pathway in the first step. A compound with a completed pyrimidine ring—orotate (6-carboxyuracil)—reacts with PRPP to form a pyrimidine ribonucleotide in the fifth step of the six-step pathway.

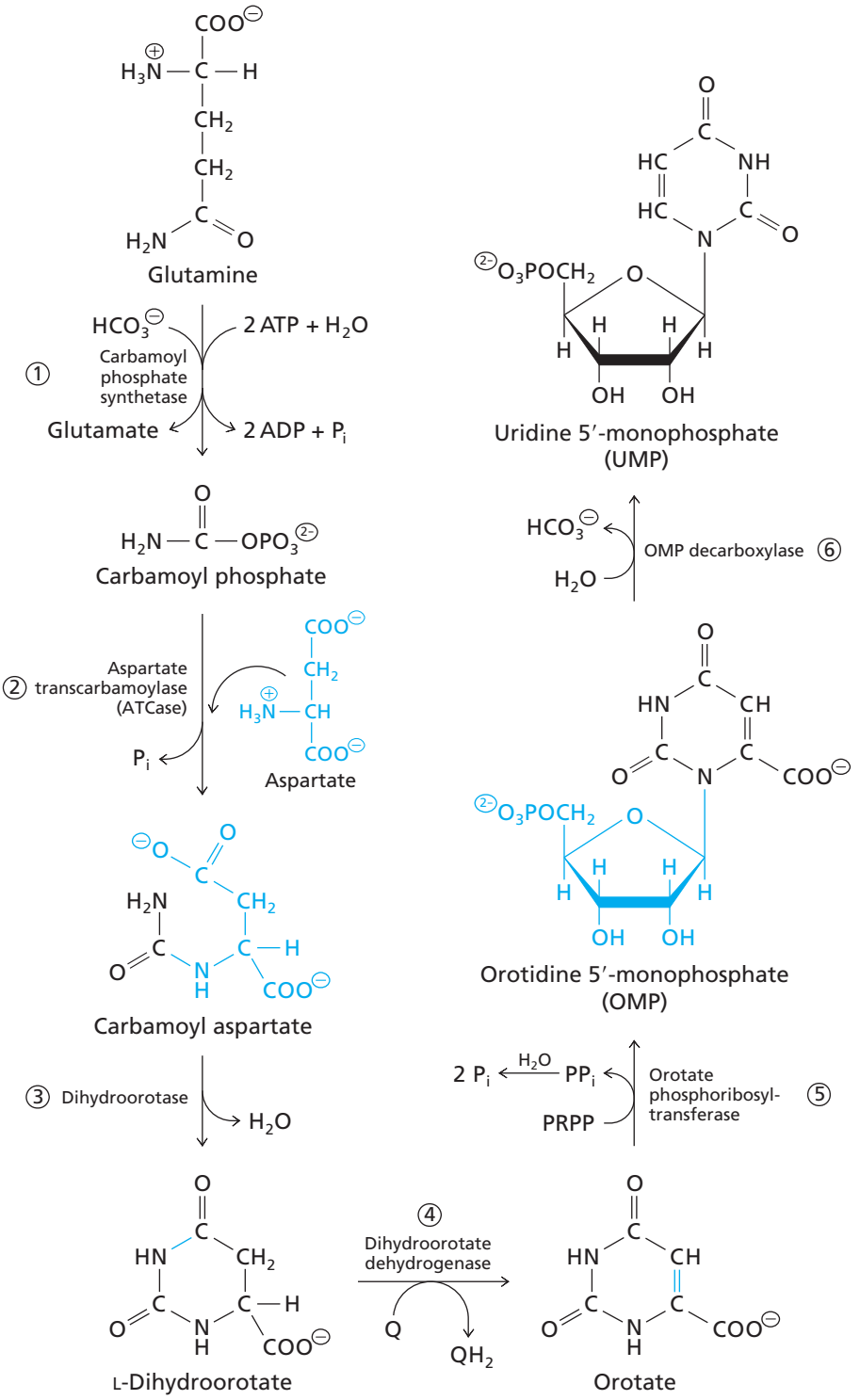
A. The Pathway for Pyrimidine Synthesis

The six-step pathway for pyrimidine synthesis is shown in Figure 18.10. The first two steps generate a noncyclic intermediate that contains all the atoms destined for the pyrimidine ring. This intermediate, carbamoyl aspartate, is enzymatically cyclized. The product is dihydroorotate and it is subsequently oxidized to orotate. Orotate is then converted to the ribonucleotide orotidine 5′-monophosphate (OMP, or orotidylate) that undergoes decarboxylation to form UMP (uridylate). This pyrimidine nucleotide is the precursor not only of all other pyrimidine ribonucleotides but also of the pyrimidine deoxyribonucleotides. The enzymes required for pyrimidine synthesis are organized and regulated differently in prokaryotes and eukaryotes.

The first step in the pathway of pyrimidine biosynthesis is the formation of carbamoyl phosphate from bicarbonate plus the amide nitrogen of glutamine and ATP. This reaction is catalyzed by carbamoyl phosphate synthetase (or by carbamoyl phosphate synthetase II activity in mammalian cells). It requires two molecules of ATP—one to drive formation of the C—N bond and the other to donate a phosphoryl group. This enzyme is not the same carbamoyl phosphate synthetase that is used in the urea cycle. That enzyme, carbamoyl phophate synthetase I, assimilates free ammonia whereas this enzyme (carbamoyl phosphate synthetase II in animals) transfers an amino group from glutamine.

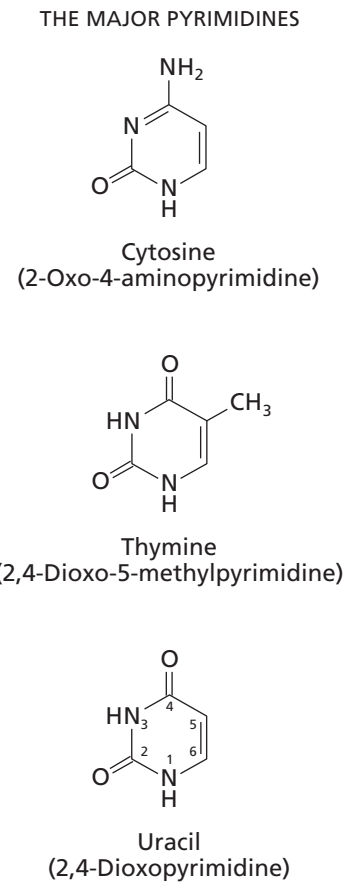
The activated carbamoyl group of carbamoyl phosphate is transferred to aspartate to form carbamoyl aspartate in the second step of UMP biosynthesis. This reaction is catalyzed by a famous enzyme, aspartate transcarbamoylase (ATCase). The mechanism involves the nucleophilic attack of the aspartate nitrogen on the carbonyl group of carbamoyl phosphate.

Dihydroorotase catalyzes the third step of UMP biosynthesis—the reversible closure of the pyrimidine ring (Figure 18.10). The product, dihydroorotate, is then oxidized by the



▲ **Figure 18.10**
Six-step pathway for the synthesis of UMP in prokaryotes. In eukaryotes, steps 1 through 3 are catalyzed by a multifunctional protein called dihydroorotate synthase, and steps 5 and 6 are catalyzed by a bifunctional enzyme, UMP synthase.

action of dihydroorotate dehydrogenase to form orotate. In eukaryotes, dihydroorotate is produced in the cytosol by steps 1 through 3. It then passes through the outer mitochondrial membrane prior to being oxidized to orotate by the action of dihydroorotate dehydrogenase. This enzyme is associated with the inner mitochondrial membrane. Its substrate binding site is located on the outer surface. The enzyme is an iron-containing



▲ **Cytosine, Thymine, Uracil**

flavoprotein that catalyzes the transfer of electrons to ubiquinone (Q) forming ubiquinol (QH_2). Electrons from QH_2 are then transferred to O_2 via the electron transport chain.

Once formed, orotate displaces the pyrophosphate group of PRPP, producing OMP in a reaction catalyzed by orotate phosphoribosyltransferase. The subsequent hydrolysis of pyrophosphate makes this reaction essentially irreversible.

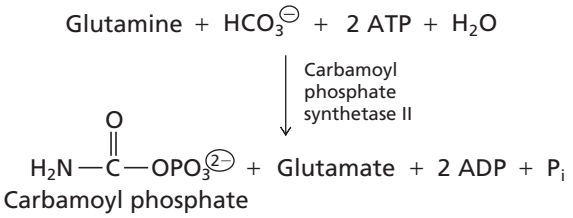
Finally, OMP is decarboxylated to form UMP in a reaction catalyzed by OMP decarboxylase. In eukaryotes, orotate produced in the mitochondria moves to the cytosol where it is converted to UMP. A bifunctional enzyme known as UMP synthase catalyzes both the reaction of orotate with PRPP to form OMP and the rapid decarboxylation of OMP to UMP.

Orotidine 5′-phosphate decarboxylase (OMP decarboxylase) is one of the most efficient enzymes known (Table 5.2).

In mammals, the intermediates formed in steps 1 and 2 (carbamoyl phosphate and carbamoyl aspartate) and OMP (from step 5) are not normally released to solvent but remain bound to enzyme complexes and are channeled from one catalytic center to the next. Several multifunctional proteins, each catalyzing several steps, also occur in the pathway of purine nucleotide biosynthesis in some organisms.

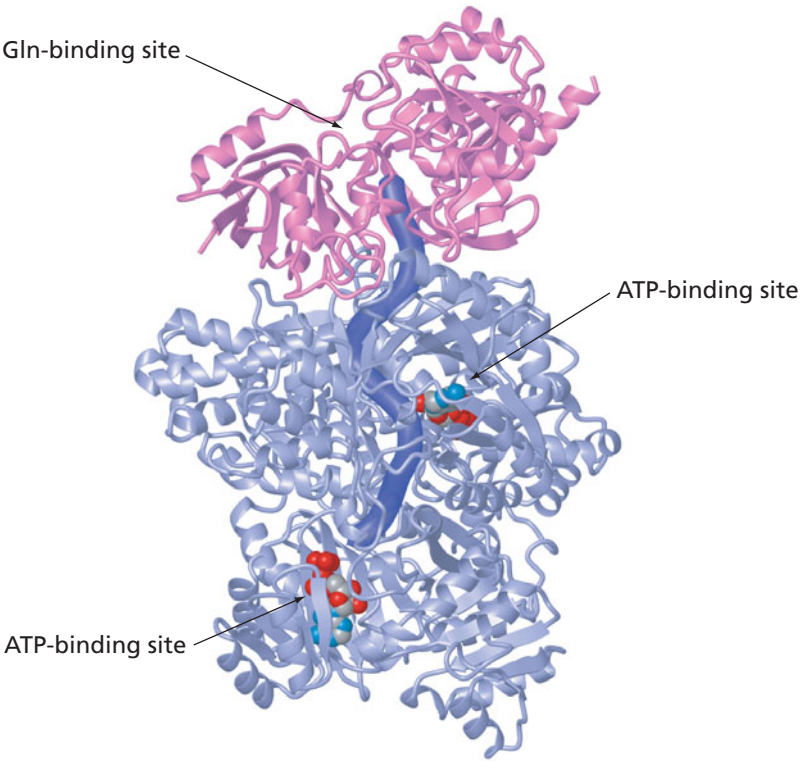
BOX 18.2 HOW SOME ENZYMES TRANSFER AMMONIA FROM GLUTAMINE

Several enzymes that use glutamine as an amide donor have a molecular tunnel running through the interior of the protein. This is an example of metabolite channeling (Section 5.11). Carbamoyl phosphate synthetase from *E. coli* is the most fully studied of these enzymes. It catalyzes the synthesis of carbamoyl phosphate from bicarbonate and glutamine:



Carbamoyl phosphate formed in this reaction is used in the synthesis of pyrimidine nucleotides. (A different carbamoyl phosphate synthetase that uses ammonia rather than glutamine as its substrate is discussed in Section 17.7A.)

Carbamoyl phosphate synthetase of *E. coli* is a heterodimer with one small subunit and one large subunit (see figure). The synthesis of carbamoyl phosphate from glutamine proceeds via three intermediates, each formed at a different active site. ATP reacts at two of these sites. The three sites are connected by a tunnel that runs from the glutamine-binding site, where a molecule of ammonia is released from glutamine, to the second ATP-binding site, where ammonia is carboxylated, and finally to the third site where carbamoyl phosphate is formed. Ammonia that is released from glutamine at the active site in the small subunit does not equilibrate with solvent but proceeds down the tunnel and undergoes the reactions that eventually produce carbamoyl phosphate. Several of the intermediates in the overall reaction are quite unstable and would be degraded by water if they were not protected by being in a tunnel.



► **Carbamoyl phosphate synthetase from *E. coli*.** The small subunit (*N*-terminal domain, purple) contains the active site for glutamine hydrolysis releasing NH_3. The large subunit is shown in blue. NH_3 is converted to the unstable intermediate carbamate ($H_2N—COOH$) at its upper ATP-binding site. Carbamate is then phosphorylated at the *C*-terminal (lower) ATP-binding site. A molecule of ADP is bound in each ATP-binding site. The molecular tunnel connecting the three active sites is shown by the thick blue wire. [PDP 1A9X]

B. Regulation of Pyrimidine Synthesis

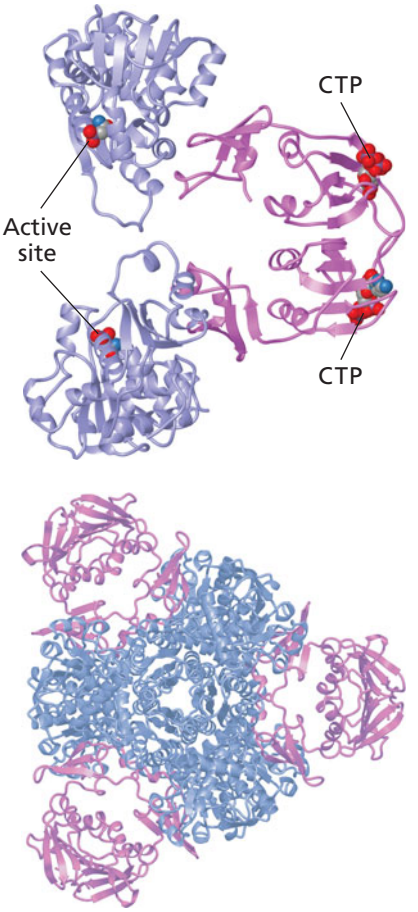
Regulation of pyrimidine biosynthesis also differs between prokaryotes and eukaryotes. Although the six enzymatic steps leading to UMP are the same in all species, the enzymes are organized differently in different organisms. In *E. coli*, each of the six reactions is catalyzed by a separate enzyme. In eukaryotes, a multifunctional protein in the cytosol known as dihydroorotate synthase contains separate catalytic sites (carbamoyl phosphate synthetase II, ATCase, and dihydroorotase) for the first three steps of the pathway.

In addition to being an intermediate in pyrimidine synthesis, carbamoyl phosphate is a metabolite in the pathway for the biosynthesis of arginine via citrulline (Figure 17.43). The same carbamoyl phosphate synthetase in prokaryotes is also used in both pyrimidine and arginine biosynthetic pathways. This enzyme is allosterically inhibited by pyrimidine ribonucleotides such as UMP, the product of the pyrimidine biosynthetic pathway. It is activated by L-ornithine, a precursor of citrulline, and by purine nucleotides, the substrates (along with pyrimidine nucleotides) for the synthesis of nucleic acids. Eukaryotic carbamoyl phosphate synthetase II is also allosterically regulated. PRPP and IMP activate the enzyme and several pyrimidine nucleotides inhibit it.

The next enzyme of the pathway is aspartate transcarbamoylase (ATCase). ATCase from *E. coli* is the most thoroughly studied allosteric enzyme. ATCase catalyzes the first committed step of pyrimidine biosynthesis since carbamoyl phosphate can enter pathways leading either to pyrimidines or to arginine in bacteria. This enzyme is inhibited by pyrimidine nucleotides and activated *in vitro* by ATP. ATCase in *E. coli* is only partially inhibited (50% to 70%) by the most potent inhibitor, CTP, but inhibition can be almost total when both CTP and UTP are present. UTP alone does not inhibit the enzyme. The allosteric controls—inhibition by pyrimidine nucleotides and activation by the purine nucleotide ATP—provide a means for carbamoyl phosphate synthetase and ATCase to balance the pyrimidine nucleotide and purine nucleotide pools in *E. coli*. The ratio of the concentrations of the two types of allosteric modulators determines the activity level of ATCase.

E. coli ATCase has a complex structure with binding sites for substrates and allosteric modulators on separate subunits. The enzyme contains six catalytic subunits arranged as two trimers and six regulatory subunits arranged as three dimers (Figure 18.11). Each subunit of a catalytic trimer is connected to a subunit of the other catalytic trimer through a regulatory dimer. When one molecule of aspartate binds, in the presence of carbamoyl phosphate, all six catalytic subunits change to a conformation having increased catalytic activity.

Eukaryotic ATCase is not feedback-inhibited. Regulation by feedback inhibition is not necessary because the pyrimidine pathway can be controlled by regulating the enzyme preceding ATCase, carbamoyl phosphate synthetase II. The substrate of ATCase in eukaryotes is not a branch-point metabolite—the synthesis of carbamoyl phosphate and citrulline for the urea cycle occurs in mitochondria, and the synthesis of carbamoyl phosphate for pyrimidines occurs in the cytosol. The pools of carbamoyl phosphate are separate.



▲ **Figure 18.11**
ATCase from *Escherichia coli*. The top structure has two regulatory subunits (purple) with a bound CTP. The two catalytic subunits (blue) have a bound substrate analog that identifies the active site. Note the large distance between the allosteric site where CTP binds and the active site of the enzyme. Three of these units are bound together to produce a large hexameric ring (below) and two of these hexameric rings stack together to create the complete 12-subunit enzyme. [PDB 2FZC (top) 9ATC (bottom)].

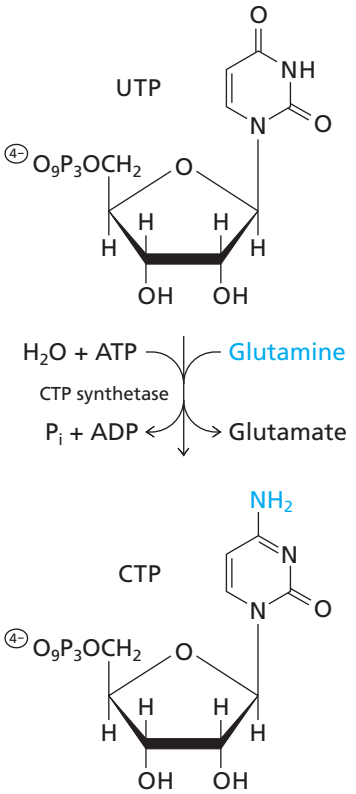
18.4 CTP Is Synthesized from UMP

UMP is converted to CTP in three steps. Uridylate kinase (UMP kinase) catalyzes the transfer of the γ-phosphoryl group of ATP to UMP to generate UDP, and then nucleoside diphosphate kinase catalyzes the transfer of the γ-phosphoryl group of a second ATP molecule to UDP to form UTP. Two molecules of ATP are converted to two molecules of ADP during the synthesis of UTP from UMP.

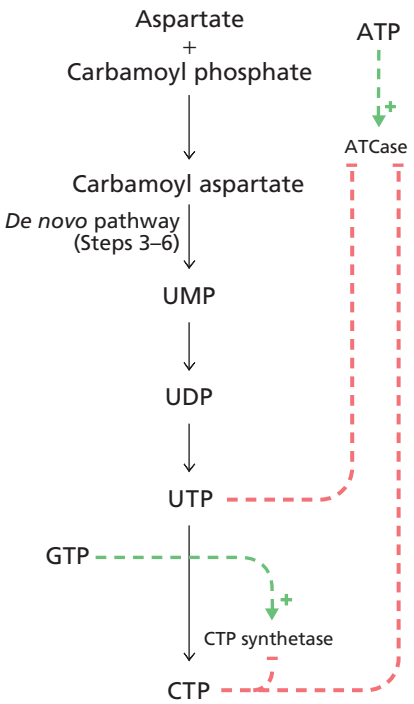
$$\text{UMP} \xrightarrow{\text{ATP}\ \rightarrow\ \text{ADP}} \text{UDP} \xrightarrow{\text{ATP}\ \rightarrow\ \text{ADP}} \text{UTP} \qquad (18.1)$$

CTP synthetase then catalyzes the ATP-dependent transfer of the amide nitrogen from glutamine to C-4 of UTP forming CTP (Figure 18.12). This reaction is chemically analogous to step 4 of purine biosynthesis (Figure 18.5) and to GMP synthesis from XMP catalyzed by GMP synthetase (Figure 18.7).

CTP synthetase is allosterically inhibited by its product, CTP, and in *E. coli* it is allosterically activated by GTP (Figure 18.13). The regulation of ATCase and CTP synthetase balances the concentrations of endogenous pyrimidine nucleotides. Elevated levels of CTP block further synthesis of CTP by inhibiting CTP synthetase. Under these



▲ **Figure 18.12**
Conversion of UTP to CTP.



▲ **Figure 18.13**
Regulation of pyrimidine nucleotide synthesis in *E. coli*. Allosteric regulation of ATCase and CTP synthetase by both purine and pyrimidine nucleotides helps balance nucleotide synthesis.

conditions, UMP synthesis will be slowed but not stopped since CTP only partially inhibits ATCase. UMP can still be used in RNA synthesis and as a precursor to dTTP (Section 18.6). ATCase is completely inhibited when the concentrations of both UTP and CTP are elevated. Elevated concentrations of the purine nucleotides ATP and GTP increase the rates of synthesis of the pyrimidine nucleotides and this helps balance the supplies of purine and pyrimidine nucleotides.

18.5 Reduction of Ribonucleotides to Deoxyribonucleotides

The 2′-deoxyribonucleoside triphosphates are synthesized by the enzymatic reduction of ribonucleotides. This reduction occurs at the nucleoside diphosphate level in most organisms. Peter Reichard and his colleagues showed that all four ribonucleoside diphosphates—ADP, GDP, CDP, and UDP—are substrates of a single, closely regulated, ribonucleoside diphosphate reductase. In some microorganisms, including species of *Lactobacillus*, *Clostridium*, and *Rhizobium*, ribonucleoside triphosphates are the substrates for reduction by a cobalamin-dependent reductase. Both types of enzymes are called ribonucleotide reductase (class I and class II, respectively), although the more precise names are ribonucleoside diphosphate reductase and ribonucleoside triphosphate reductase.

NADPH provides the reducing power for the synthesis of deoxyribonucleoside diphosphates in class I enzymes. A disulfide bond at the active site of ribonucleotide reductase is reduced to two thiol groups that reduce C-2′ of the ribose moiety of the nucleotide substrate by a complex free-radical mechanism. As shown in Figure 18.14, electrons are transferred from NADPH to ribonucleotide reductase via the flavoprotein thioredoxin reductase and the dithiol protein coenzyme thioredoxin (Figure 7.35). Thioredoxin reductase of prokaryotes and yeast has a dithiol/disulfide (cysteine pair) group in the active site. In mammalian thioredoxin reductase, the oxidation–reduction center differs by having one residue of cysteine and one of selenocysteine. Once formed, dADP, dGDP, and dCDP are phosphorylated to the triphosphate level by the action of nucleoside diphosphate kinases. dUDP, as we will see in the next section, is converted to dTMP via dUMP. A third version of ribonucleotide reductase (class III) uses *S*-adenosylmethionine as a cofactor.

Ribonucleotide reductase has a complicated mechanism of allosteric regulation that supplies a balanced pool of the deoxynucleotides required for DNA synthesis. Both the substrate specificity and the catalytic rate of ribonucleotide reductase are regulated in eukaryotic cells by the reversible binding of nucleotide metabolites. The allosteric modulators—ATP, dATP, dTTP, and dGTP—act by binding to ribonucleotide reductase at either of two regulatory sites. One allosteric site, called the *activity site*, controls the activity of the catalytic site. A second allosteric site, called the *specificity site*, controls the substrate specificity of the catalytic site (Figure 18.15). The binding of ATP to the activity site forms an activated enzyme whereas the binding of dATP to the activity site inhibits all enzymatic activity. When ATP is bound to the activity site and either ATP or dATP is bound to the specificity site, the reductase becomes pyrimidine specific, catalyzing the reduction of CDP and UDP. The binding of dTTP to the specificity site activates the reduction of GDP, and the binding of dGTP activates the reduction of ADP. The allosteric regulation of ribonucleotide reductase, summarized in Table 18.1, controls enzyme activity and ensures a balanced selection of deoxyribonucleotides for DNA synthesis.

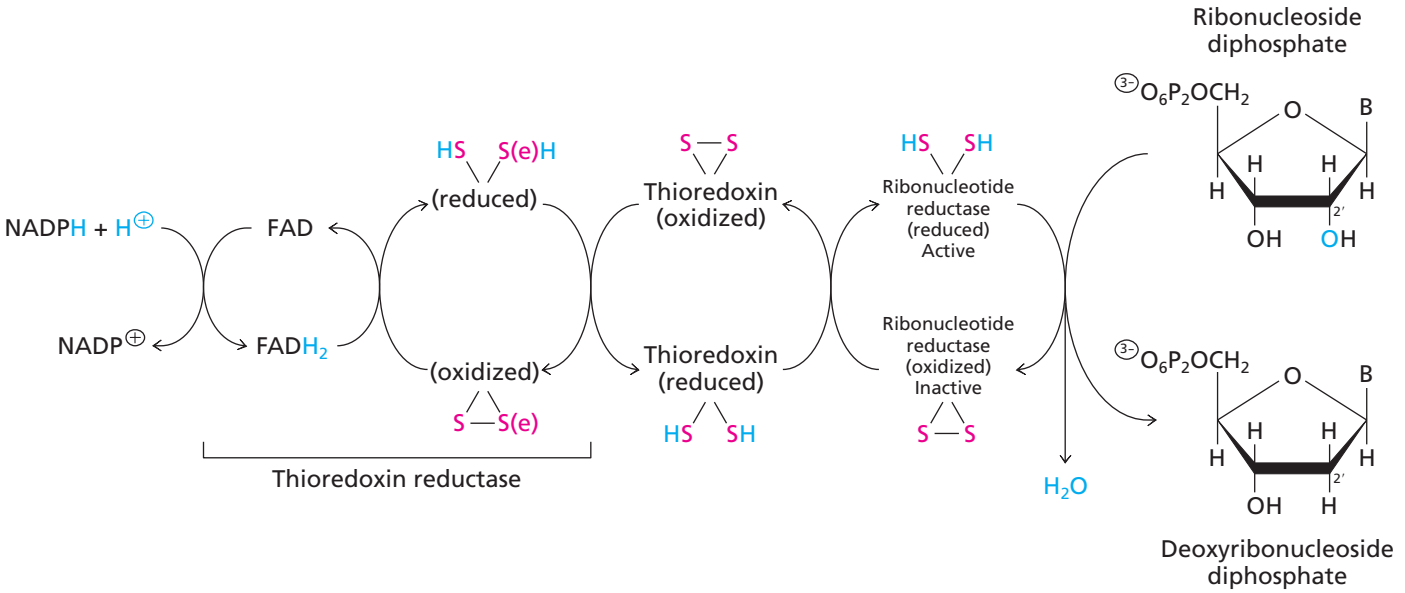
18.6 Methylation of dUMP Produces dTMP

Deoxythymidylate (dTMP) is formed from UMP in four steps. UMP is phosphorylated to UDP that is reduced to dUDP and dUDP is dephosphorylated to dUMP. dUMP is then methylated to dTMP.

$$\text{UMP} \longrightarrow \text{UDP} \longrightarrow \text{dUDP} \longrightarrow \text{dUMP} \longrightarrow \text{dTMP} \qquad (18.2)$$

The conversion of dUDP to dUMP can occur by two routes. dUDP can react with ADP in the presence of a nucleoside monophosphate kinase to form dUMP and ATP.

$$\text{dUDP} + \text{ADP} \rightleftharpoons \text{dUMP} + \text{ATP} \qquad (18.3)$$



▲ **Figure 18.14**
Reduction of ribonucleoside diphosphates. Three proteins are involved: the NADPH-dependent flavoprotein thioredoxin reductase, thioredoxin, and ribonucleotide reductase. B represents a purine or pyrimidine base. S(e) represents either sulfur or selenium.

dUDP can also be phosphorylated to dUTP at the expense of ATP through the action of nucleoside diphosphate kinases. dUTP is then rapidly hydrolyzed to dUMP + PP_i by the action of deoxyuridine triphosphate diphosphohydrolase (dUTPase).

$$\text{dUDP} + \text{ATP} \longrightarrow \text{dUTP} \longrightarrow \text{dUMP} + PP_i \qquad (18.4)$$

(ATP → dUTP: ADP; dUTP → dUMP + PP_i: H_2O)

The rapid hydrolysis of dUTP prevents it from being accidentally incorporated into DNA in place of dTTP.

dCMP can also be a source of dUMP via hydrolysis catalyzed by dCMP deaminase.

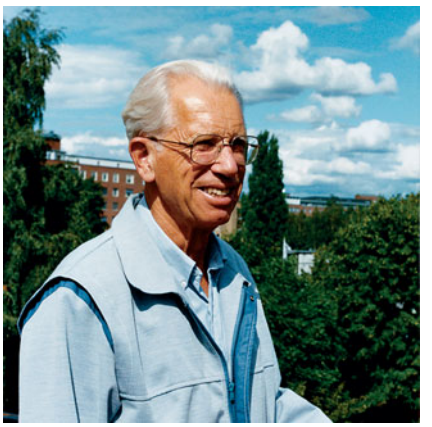
$$\text{dCMP} + H_2O \longrightarrow \text{dUMP} + HN_4^{\oplus} \qquad (18.5)$$

The conversion of dUMP to dTMP is catalyzed by the enzyme known as thymidylate synthase. (Because thymine occurs almost exclusively in DNA, the trivial names thymidine and thymidylate are commonly used instead of deoxythymidine and deoxythymidylate.) 5,10-Methylenetetrahydrofolate is the donor of the one-carbon group in this reaction (Figure 18.16). The carbon-bound methyl group (C—CH_3) in dTMP is more reduced than the nitrogen-bridged methylene group (N—CH_2—N) in 5,10-methylenetetrahydrofolate, whose oxidation state is equivalent to that of a nitrogen-bound hydroxymethyl group (N—CH_2OH) or formaldehyde. Thus, not only is methylenetetrahydrofolate a coenzyme donating a one-carbon unit but it is also the reducing agent for the reaction, furnishing a hydride ion and being oxidized to 7,8-dihydrofolate in the process. This is the only known reaction in which the transfer of a one-carbon unit from a tetrahydrofolate derivative results in its oxidation at N-5 and C-6 to produce dihydrofolate.

The structure of selenocysteine, the 22nd amino acid, is shown in Section 3.3.

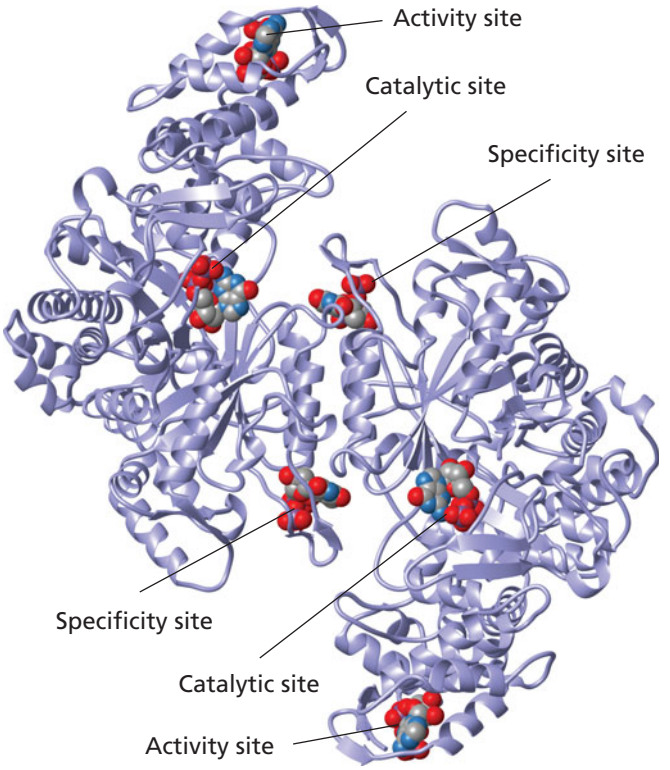
Table 18.1 Allosteric regulation of eukaryotic ribonucleotide reductase

Ligand bound to activity site	Ligand bound to specificity site	Activity of catalytic site
dATP		Enzyme inactive
ATP	ATP or dATP	Specific for CDP or UDP
ATP	dTTP	Specific for GDP
ATP	dGTP	Specific for ADP



▲ **Peter Reichard (1925–).**
Reichard worked for many years at the Karolinska Institute in Sweden. In addition to working on ribonucleotide reductase, he was an active member of the Nobel Committee that selects candidates to receive the Nobel Prize.

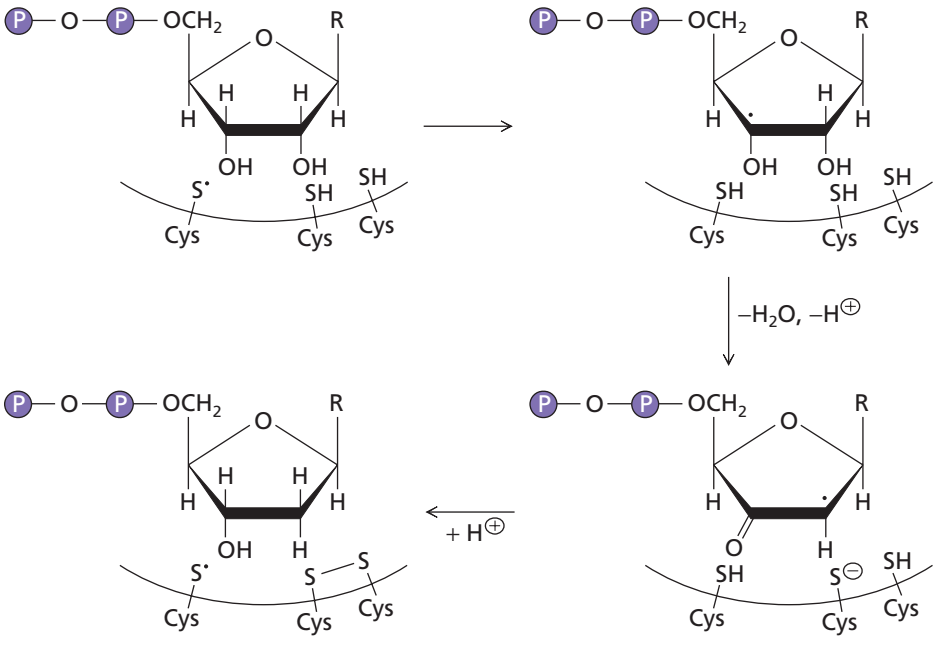
Figure 18.15 ►
Ribonucleotide reductase. The complete enzyme is an $\alpha_2\beta_2$ tetramer. The structure shown here (from *E. coli*) shows only the α_2 dimer of catalytic subunits. The activity site is occupied by an ATP analog. A molecule of TTP is bound to the specificity site and a molecule of GDP is bound at the active site. [PDB 3R1R + 4R1R]

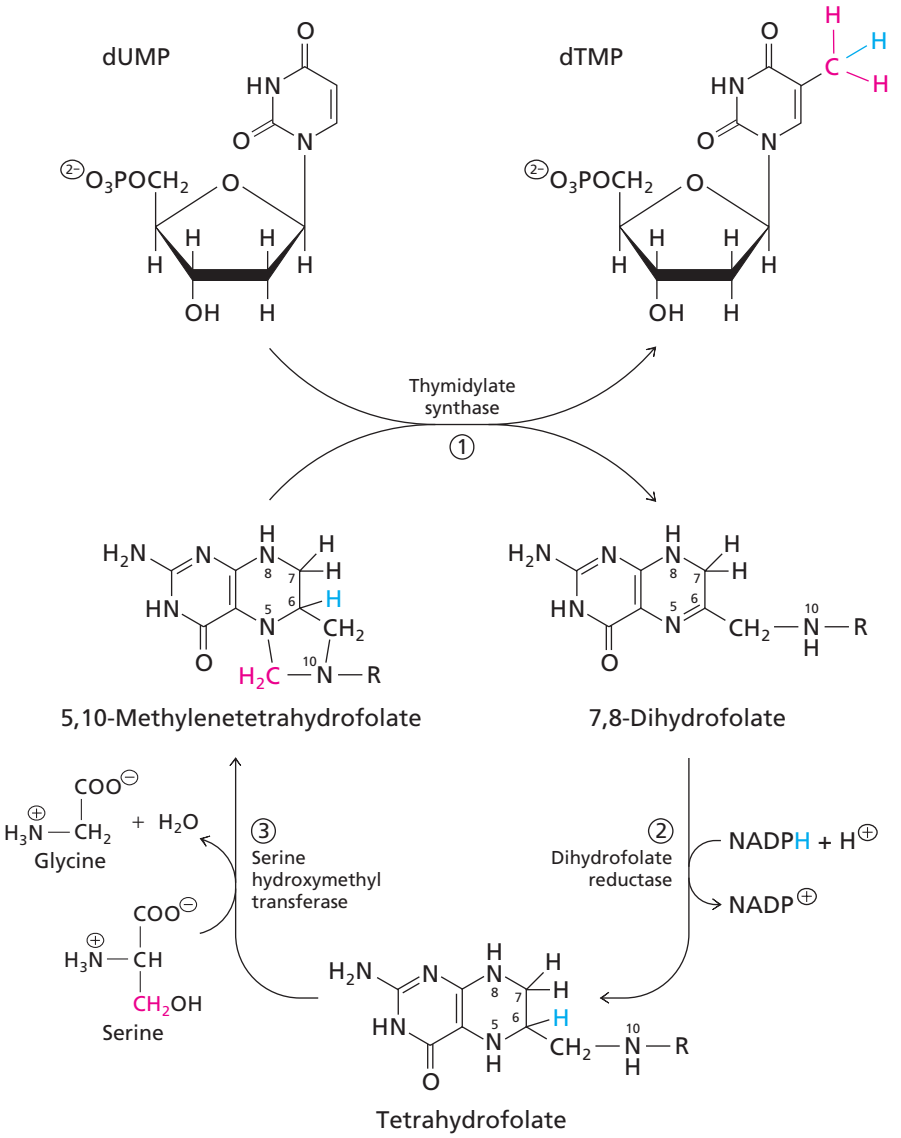


BOX 18.3 FREE RADICALS IN THE REDUCTION OF RIBONUCLEOTIDES

The ribonucleotide reductase reaction is an unusual reaction because it proceeds by a free radical mechanism. The first clue to the free radical nature of the reaction was the observation that the reductase from *E. coli* could be isolated with a tyrosine residue in the free radical form. This was the first free radical protein to be discovered. The role of the tyrosine radical is to convert the thiol group of an active-site cysteine residue to a thiyl radical. (In the *Lactobacillus* enzyme, cobalamin serves to convert the active-site thiol to a radical.)

The proposed mechanism is shown in the accompanying figure. The active site of the reductase has three cysteine residues—one forms the free radical and the other two are an oxidation–reduction group. The thiyl radical removes a hydrogen atom from the C-3′ position of the ribonucleotide forming a substrate radical. This substrate radical is first dehydrated (losing the C-2′—OH) and then reduced by the cysteine reduction pair. A hydrogen atom is returned to C-3′, regenerating the thiyl radical.





◀ **Figure 18.16**
Cycle of reactions in the synthesis of thymidylate (dTMP) from dUMP. Thymidylate synthase catalyzes the first reaction of this cycle producing dTMP. The other product of the reaction, dihydrofolate, must be reduced by NADPH in a reaction catalyzed by dihydrofolate reductase before a methylene group can be added to regenerate 5,10-methylenetetrahydrofolate. Methylenetetrahydrofolate is regenerated in a reaction catalyzed by serine hydroxymethyltransferase.

Dihydrofolate must be converted to tetrahydrofolate before the coenzyme can accept another one-carbon unit for further transfer reactions. The 5,6 double bond of dihydrofolate is reduced by NADPH in a reaction catalyzed by dihydrofolate reductase. Serine hydroxymethyltransferase (Figure 17.16) then catalyzes the transfer of the β-CH_2OH group of serine to tetrahydrofolate to regenerate 5,10-methylenetetrahydrofolate.

Thymidylate synthase and dihydrofolate reductase are distinct polypeptides in most organisms but in protozoa the two enzyme activities are contained on the same polypeptide chain. The dihydrofolate product of the first reaction is channeled from the thymidylate synthase active site to the dihydrofolate reductase active site. Charge–charge interactions between a positively charged region on the surface of the bifunctional enzyme and the negatively charged dihydrofolate (recall that it contains several γ-glutamate residues; Section 7.11) steer the product to the next active site.

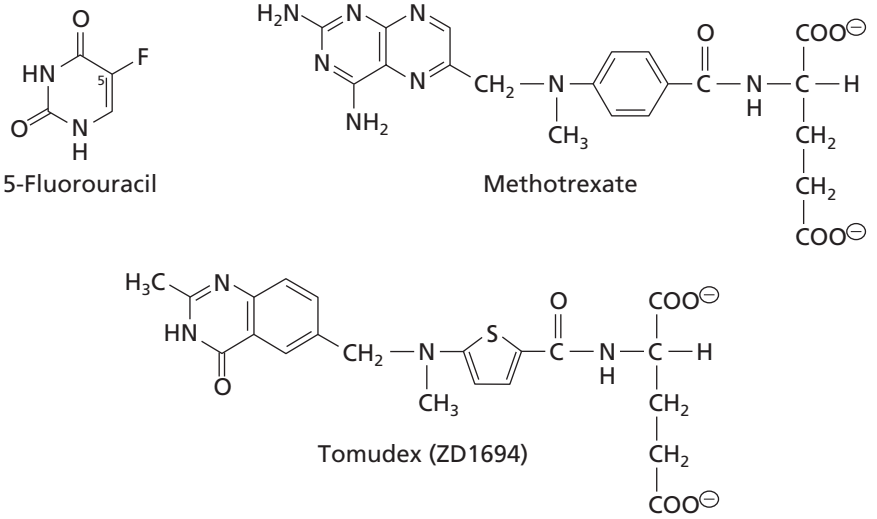
dTMP can also be synthesized via the salvage of thymidine (deoxythymidine), catalyzed by ATP-dependent thymidine kinase.

$$\text{Deoxythymidine (Thymidine)} \xrightarrow[\text{Thymidine kinase}]{\text{ATP ADP}} \text{dTMP} \qquad (18.6)$$

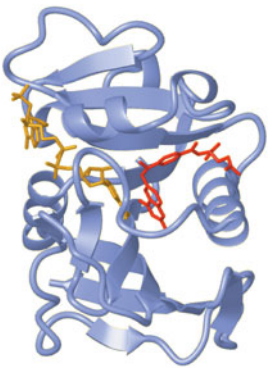
BOX 18.4 CANCER DRUGS INHIBIT dTTP SYNTHESIS

Since dTMP is an essential precursor of DNA, any agent that lowers dTMP levels drastically affects cell division. Thymidylate synthase and dihydrofolate reductase have been major targets for anticancer drugs because rapidly dividing cells are particularly dependent on the activities of these enzymes. The inhibition of either or both of these enzymes blocks the synthesis of dTMP and therefore the synthesis of DNA.

5-Fluorouracil, methotrexate, and Tomudex are effective in combating some types of cancer. 5-Fluorouracil is converted to its deoxyribonucleotide, 5-fluorodeoxyuridylate, which binds tightly to thymidylate synthase inhibiting the enzyme and bringing the three-reaction cycle shown in Figure 18.16 to a halt. Methotrexate, an analog of folate, is a potent, relatively specific inhibitor of dihydrofolate reductase that catalyzes step 2 of the cycle shown in Figure 18.16. The resulting decrease in tetrahydrofolate levels greatly diminishes the formation of dTMP since dTMP synthesis depends on adequate concentrations of methylenetetrahydrofolate. Tomudex is a folate-based inhibitor of human thymidylate synthase that has been approved for the treatment of cancer.



▲ 5-Fluorouracil, methotrexate, and Tomudex are drugs designed to inhibit thymidylate synthase and block the growth of rapidly dividing cells.



▲ Structure of human dihydrofolate reductase with the substrate analog methotrexate (red) and the cosubstrate NADPH (gold) bound in the active site. [PDB 1DLS]

Radioactive thymidine is often used as a highly specific tracer for monitoring intracellular synthesis of DNA because it enters cells easily and its principal metabolic fate is conversion to thymidylate and incorporation into DNA.

18.7 Modified Nucleotides

DNA and RNA contain a number of modified nucleotides. The ones present in transfer RNA are well known (Section 21.8B) but the modified nucleotides in DNA are just as important. Some of the more common modified bases in DNA are shown in Figure 18.17. Most of them are only found in a few species or in bacteriophage while others are more widespread.

We will encounter N^6-methyladenine in the next chapter when we discuss restriction endonucleases. 5-Methylcytosine is a common modified base in mammalian DNA because it plays a role in chromatin assembly and the regulation of transcription. About 3% of all deoxycytidylate residues in mammalian DNA are modified to 5-methylcytidine. The methylation occurs after DNA is synthesized and the modified residues are at CG sequences. All of these modified nucleotides are made *in situ* by enzymes that act on one of the four common nucleotides in the DNA molecule.

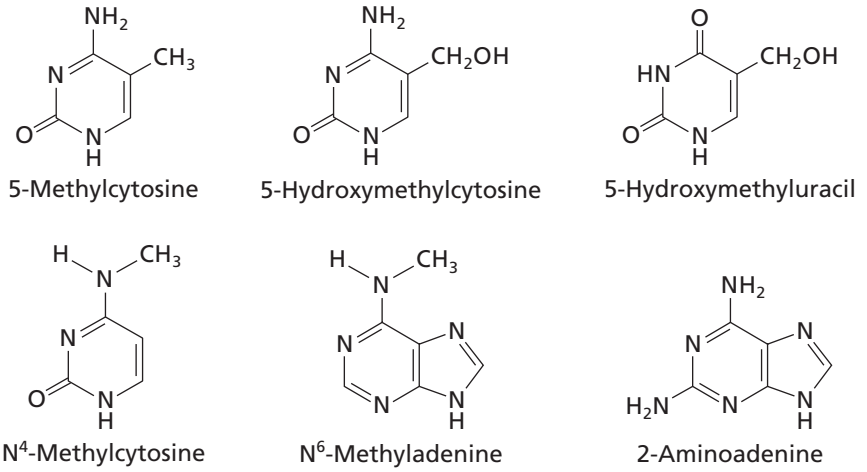


▲ **Salvage pathways are a form of biochemical recycling.**

18.8 Salvage of Purines and Pyrimidines

Nucleic acids are degraded to mononucleotides, nucleosides, and eventually, heterocyclic bases during normal cell metabolism (Figure 18.18). The catabolic reactions are catalyzed by

ribonucleases, deoxyribonucleases, and a variety of nucleotidases, nonspecific phosphatases, and nucleosidases or nucleoside phosphorylases. Some of the purine and pyrimidine bases formed in this way are further degraded (e.g., purines are converted to uric acid and other excretory products) but a considerable fraction is normally salvaged by direct conversion back to 5′-mononucleotides. PRPP is the donor of the 5-phosphoribosyl moiety for salvage reactions. The degradation pathways are part of fuel metabolism in animals. Purines and pyrimidines formed during digestion are more likely to be degraded while those formed inside the cell are usually salvaged. The recycling of intact bases conserves cellular energy.

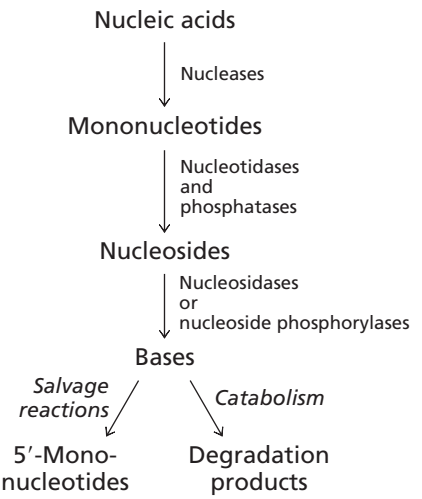


▲ Figure 18.17
Modified bases in DNA.

The degradation of purine nucleotides to their respective purines and their salvage through reaction with PRPP are outlined in Figure 18.19. Adenine phosphoribosyltransferase catalyzes the reaction of adenine with PRPP to form AMP and PP_i. The hydrolysis of PP_i, catalyzed by pyrophosphatase, renders the reaction metabolically irreversible. Hypoxanthine–guanine phosphoribosyltransferase catalyzes similar reactions—the conversion of hypoxanthine to IMP and of guanine to GMP with formation of PP_i.

Pyrimidines are salvaged by the action of orotate phosphoribosyltransferase, which catalyzes step 5 of the biosynthesis pathway (Figure 18.10). This enzyme can also catalyze the conversion of pyrimidines other than orotate to the corresponding pyrimidine nucleotides.

Nucleotides and their constituents are interconverted by many reactions, some of which we have seen already. The actions of phosphatases, nucleotidases, and nucleosidases or nucleoside phosphorylases can release bases from nucleotides. Reactions catalyzed by phosphoribosyltransferases or nucleoside phosphorylases can salvage the bases and nucleosides by converting them to the nucleotide level. Bases that are not salvaged can be catabolized. The interconversions of purine nucleotides and their constituents are summarized in Figure 18.20, and the interconversions of pyrimidine nucleotides and their constituents are summarized in Figure 18.21.

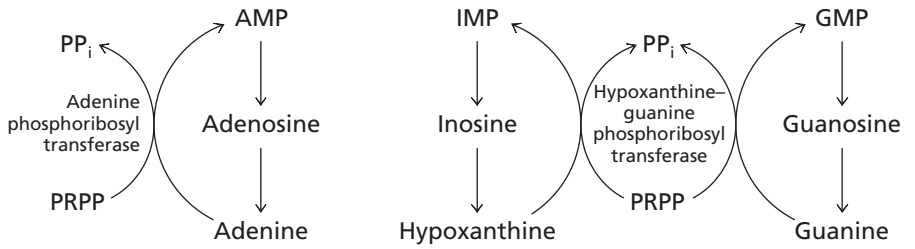


▲ Figure 18.18
Breakdown of nucleic acids.

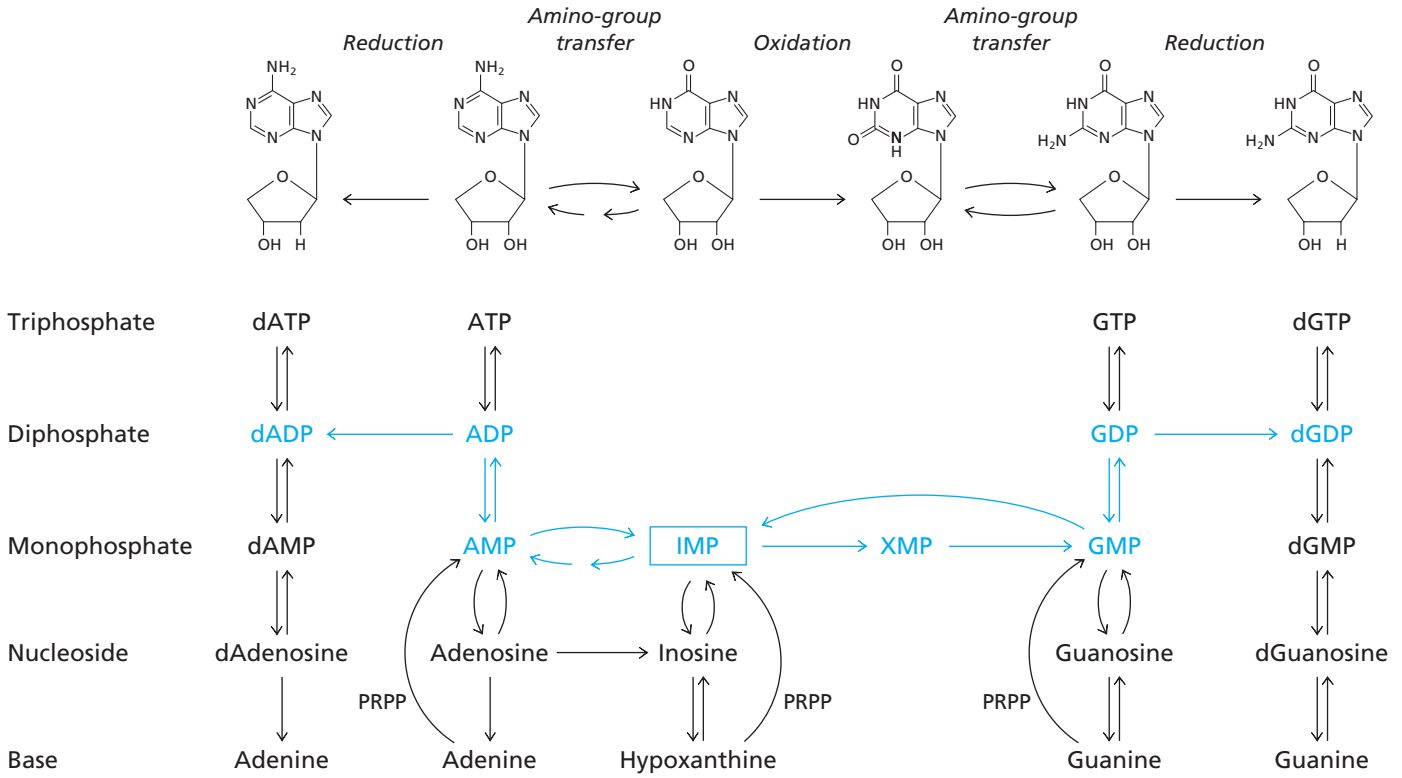
18.9 Purine Catabolism

Most free purine and pyrimidine molecules are salvaged but some are catabolized. Birds, some reptiles, and primates (including humans) convert purine nucleotides to uric acid or urate, which is then excreted. In birds and reptiles, amino acid catabolism also leads to uric acid; in mammals, surplus nitrogen from amino acid catabolism is disposed of in the form of urea. Birds and reptiles cannot further catabolize uric acid (urate) but many organisms degrade urate to other products.

As shown in Figure 18.20, AMP can be broken down to hypoxanthine and GMP is broken down to guanine. The hydrolytic removal of phosphate from AMP and GMP produces adenosine and guanosine, respectively. Adenosine can be deaminated to inosine by the action of adenosine deaminase. Alternatively, AMP can be deaminated to



◀ Figure 18.19
Degradation and salvage of purines.



▲ **Figure 18.20**
Interconversions of purine nucleotides and their constituents. IMP, the first nucleotide product of the *de novo* biosynthetic pathway, is readily converted to AMP and GMP, their di- and triphosphates, and the deoxy counterparts of these nucleotides. 5′-Phosphate groups are not shown in the abbreviated structures. [Adapted from Traut, T. W. (1988). Enzymes of nucleotide metabolism: the significance of subunit size and polymer size for biological function and regulatory properties. *Crit. Rev. Biochem.* 23:121–169.]

IMP by the action of AMP deaminase and then IMP can be hydrolyzed to inosine. The phosphorolysis of inosine produces hypoxanthine and the phosphorolysis of guanosine produces guanine. Both these reactions (as well as the phosphorolysis of several deoxynucleosides) are catalyzed by purine–nucleoside phosphorylase and produce α-D-ribose 1-phosphate (or deoxyribose 1-phosphate) and the free purine base.

$$\text{(Deoxy)Nucleoside} + P_i \rightleftharpoons \text{Base} + \text{(Deoxy)-}\alpha\text{-D-Ribose 1-phosphate} \quad \textbf{(18.7)}$$

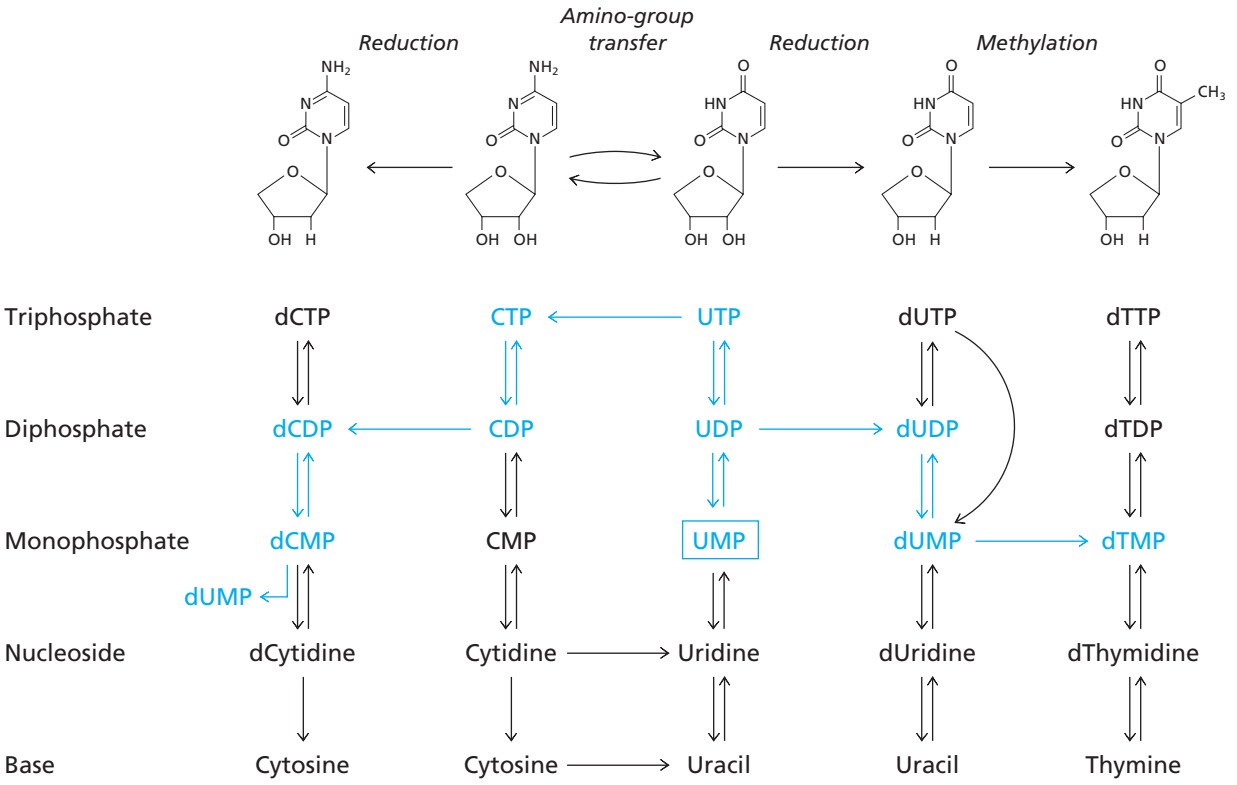
Adenosine is not a substrate of mammalian purine-nucleoside phosphorylase.

Hypoxanthine formed from inosine is oxidized to xanthine, and xanthine is oxidized to urate (Figure 18.22). Either xanthine oxidase or xanthine dehydrogenase can catalyze both reactions. Electrons are transferred to O_2 to form hydrogen peroxide (H_2O_2) in the reactions catalyzed by xanthine oxidase. (The H_2O_2 is converted to H_2O and O_2 by the action of catalase.) Xanthine oxidase is an extracellular enzyme in mammals and it appears to be an altered form of the intracellular enzyme xanthine dehydrogenase that generates the same products as xanthine oxidase but transfers electrons to $NAD^{\oplus}$ to form NADH. These two enzyme activities occur widely in nature and exhibit broad substrate specificity. Their active sites contain complex electron-transfer systems that include an iron–sulfur cluster, a pterin coenzyme with bound molybdenum, and FAD.

See Section 6.5D for a description of the adenosine deaminase mechanism.

In most cells, guanine is deaminated to xanthine in a reaction catalyzed by guanase (Figure 18.22). Animals that lack guanase excrete guanine. For example, pigs excrete guanine but metabolize adenine derivatives further to allantoin, the major end product of the catabolism of purines in most mammals.

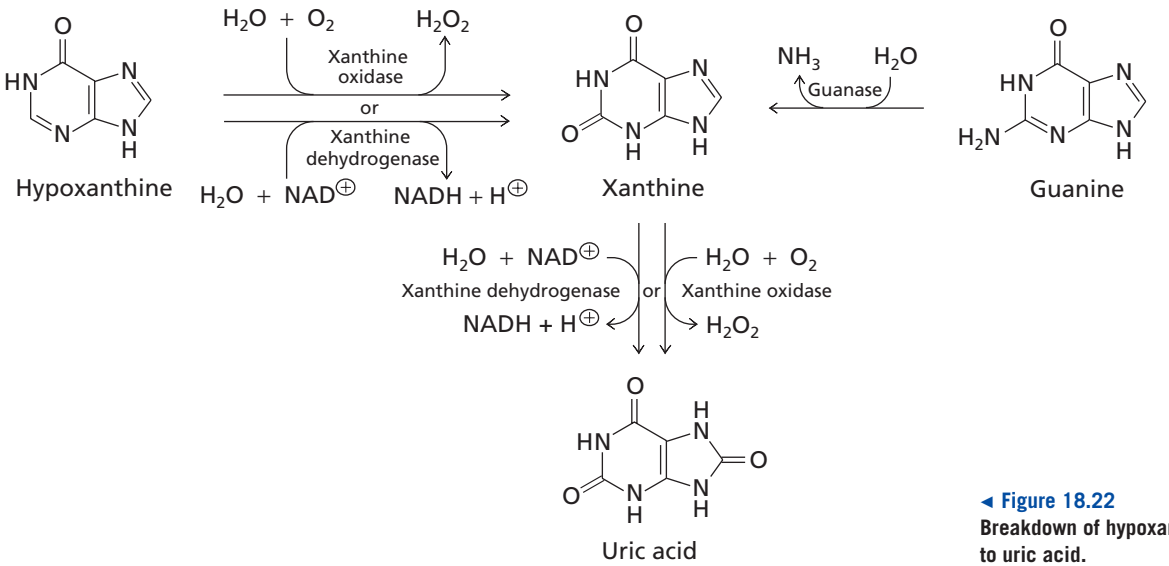
Urate can be further oxidized in most organisms. Up until recently it was thought that urate oxidase converted urate directly to allantoin but it is now known that the pathway is more complex. The conversion of urate to the stereospecific product (S)-allantoin



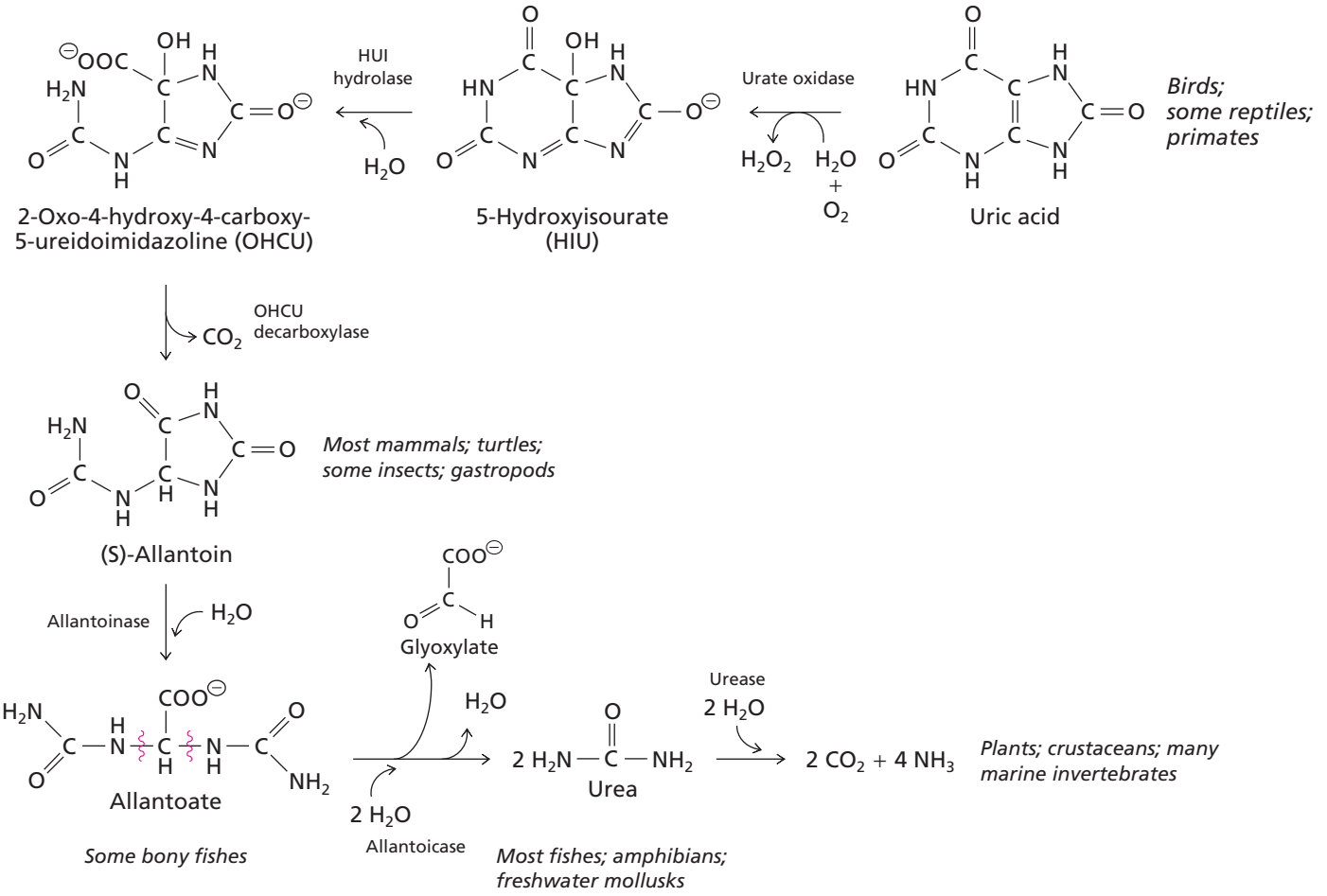
▲ **Figure 18.21**
Interconversions of pyrimidine nucleotides and their constituents. UMP formed by the *de novo* pathway can be converted to cytidine and thymidine phosphates, as well as to other uridine derivatives. 5′-Phosphate groups are not shown in the abbreviated structures. [Adapted from Traut, T. W. (1988). Enzymes of nucleotide metabolism: the significance of subunit size and polymer size for biological function and regulatory properties. *Crit. Rev. Biochem.* 23:121–169.]

requires urate oxidase plus two additional enzymes as shown in Figure 18.23. Peroxide (H_2O_2) and CO_2 are released in this series of reactions. Allantoin is the major end product of purine degradation in most mammals (though not in humans, for whom the end product is urate). It is also excreted by turtles, some insects, and also gastropods.

The enzyme allantoinase catalyzes hydrolytic opening of the imidazole ring of allantoin to produce allantoate, the conjugate base of allantoic acid. Some bony fishes (teleosts) possess allantoinase activity and excrete allantoate as the end product of purine degradation.



◀ **Figure 18.22**
Breakdown of hypoxanthine and guanine to uric acid.



▲ **Figure 18.23**
Catabolism of uric acid through oxidation and hydrolysis. To the right of each compound are listed the organisms for which it is an excretory product.



▲ When they were alive, these snails could convert urate to allantoin. Humans can't do that.

Most fishes, amphibians, and freshwater mollusks can further degrade allantoate. These species contain allantoicase that catalyzes the hydrolysis of allantoate to one molecule of glyoxylate and two molecules of urea. Urea is the nitrogenous end product of purine catabolism in these organisms.

Finally, several organisms—including plants, crustaceans, and many marine invertebrates—can hydrolyze urea in a reaction catalyzed by urease. Carbon dioxide and ammonia are the products of this reaction. Urease is found only in the cells of organisms in which the hydrolysis of urea does not lead to ammonia toxicity. For example, in plants, ammonia generated from urea is rapidly assimilated by the action of glutamine synthetase. In marine animals, ammonia is produced in surface organs such as gills and is flushed away before it can accumulate to toxic levels. Most terrestrial organisms would be poisoned by the final nitrogen-containing product, ammonia. The enzymes that catalyze urate catabolism have been lost through evolution by organisms that excrete urate.

18.10 Pyrimidine Catabolism

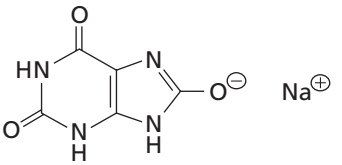
The catabolism of pyrimidine nucleotides begins with hydrolysis to the corresponding nucleosides and P_i, catalyzed by 5′-nucleotidase (Figure 18.24). Initial hydrolysis to cytidine can be followed by deamination to uridine in a reaction catalyzed by cytidine deaminase.

BOX 18.5 LESCH–NYHAN SYNDROME AND GOUT

Defects in purine metabolism can have devastating effects. In 1964 Michael Lesch and William Nyhan described a severe metabolic disease characterized by slow mental development, palsylike spasticity, and a bizarre tendency toward self-mutilation. Individuals afflicted with this disease, called Lesch–Nyhan syndrome, rarely survive past childhood. Prominent biochemical features of the disease are the excretion of up to six times the normal amount of uric acid and a greatly increased rate of purine biosynthesis.

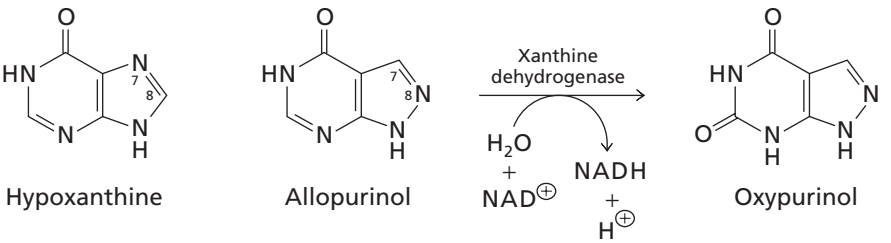
The disease is caused by a hereditary deficiency of the activity of the enzyme hypoxanthine–guanine phosphoribosyltransferase (Section 18.8). The deficiency is usually seen in males because the mutation is recessive and the gene for this enzyme is on the X chromosome. Lesch–Nyhan patients usually have less than 1% of the normal activity of the enzyme and most show a complete absence of activity. In the absence of hypoxanthine–guanine phosphoribosyltransferase, hypoxanthine and guanine are degraded to uric acid instead of being converted to IMP and GMP, respectively. The PRPP normally used for the salvage of hypoxanthine and guanine contributes to the synthesis of excessive amounts of IMP and the surplus IMP is degraded to uric acid. It is not known how this single enzyme defect causes the various behavioral symptoms. The catastrophic effects of the deficiency indicate that in some cells the purine salvage pathway in humans is not just an energy-saving addendum to the central pathways of purine nucleotide metabolism.

Gout is a disease caused by the overproduction or inadequate excretion of uric acid. Sodium urate is relatively insoluble and when its concentration in blood is elevated, it can crystallize (sometimes along with uric acid) in soft tissues, especially the kidney, and in toes and joints. Gout has several causes including a deficiency of hypoxanthine–guanine phosphoribosyltransferase activity resulting in less salvage of purines and more catabolic production of uric acid. The difference between gout and Lesch–Nyhan syndrome is due to the fact that gout patients retain up to 10% enzyme activity. Gout can also be caused by defective regulation of purine biosynthesis.



Sodium urate

Gout can be treated by giving patients allopurinol, a synthetic C-7, N-8 positional isomer of hypoxanthine. Allopurinol is converted in cells to oxypurinol, a powerful inhibitor of xanthine oxidase. Administration of allopurinol prevents the formation of abnormally high levels of uric acid. Hypoxanthine and xanthine are more soluble than sodium urate and uric acid and they are excreted when not reused by salvage reactions.



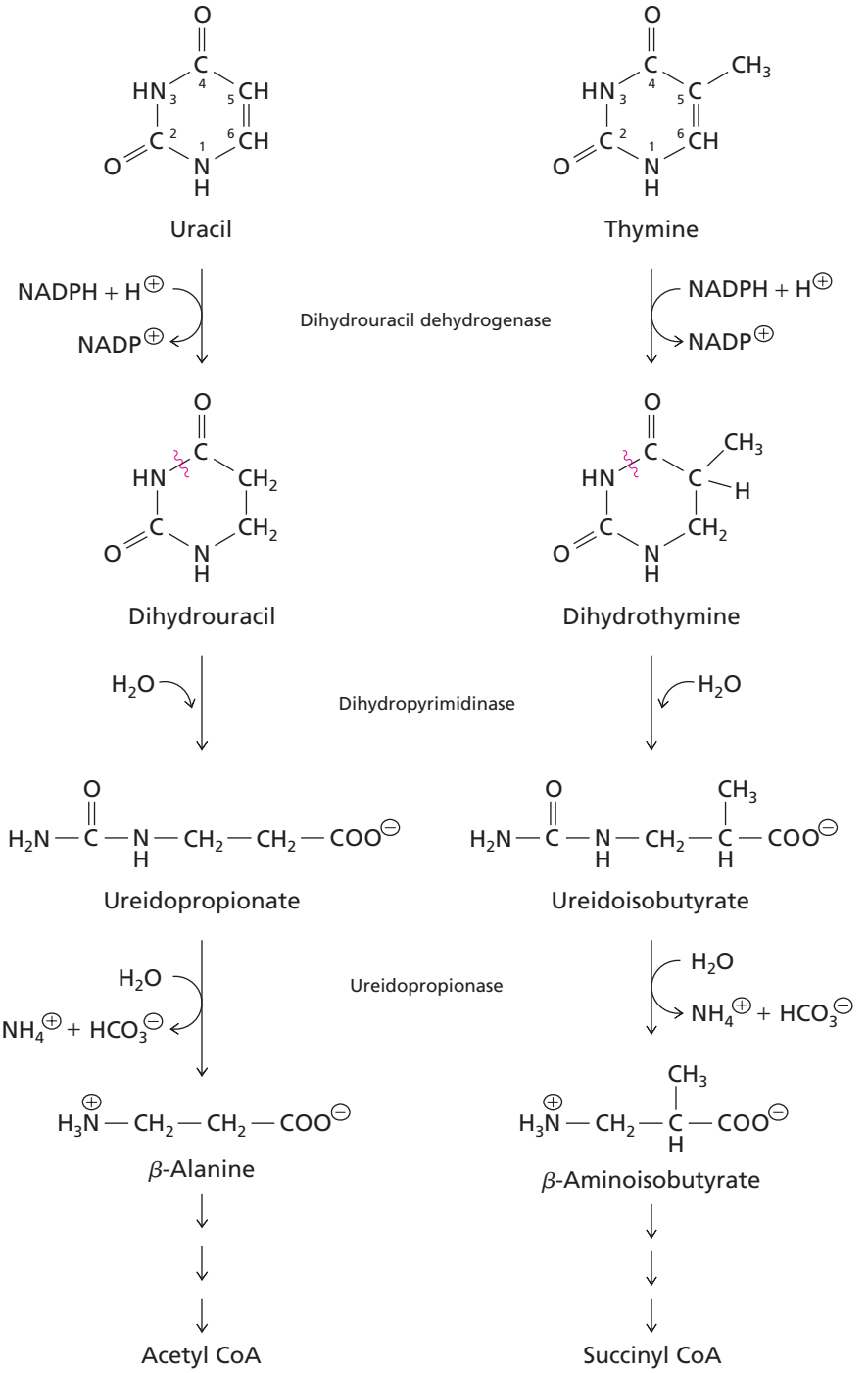
Hypoxanthine Allopurinol Oxypurinol

▲ **Allopurinol and oxypurinol.** Xanthine dehydrogenase catalyzes the oxidation of allopurinol, an isomer of hypoxanthine. The product, oxypurinol, binds tightly to xanthine dehydrogenase, inhibiting the enzyme.

The glycosidic bonds of uridine and thymidine are then cleaved by phosphorolysis in reactions catalyzed by uridine phosphorylase and thymidine phosphorylase, respectively. Deoxyuridine can also undergo phosphorolysis catalyzed by uridine phosphorylase. The products of these phosphorolysis reactions are α-D-ribose 1-phosphate or deoxyribose 1-phosphate, thymine, and uracil.

The catabolism of pyrimidines ends with intermediates of central metabolism, so no distinctive excretory products are formed. The breakdown of both uracil and thymine involves several steps (Figure 18.24). First, the pyrimidine ring is reduced to a 5,6-dihydropyrimidine in a reaction catalyzed by dihydrouracil dehydrogenase. The reduced ring

Figure 18.24 ►
Catabolism of uracil and thymine.



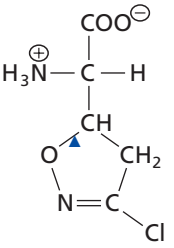
is then opened by hydrolytic cleavage of the N-3 — C-4 bond in a reaction catalyzed by dihydropyrimidinase. The resulting carbamoyl-β-amino acid derivative (ureidopropionate or ureidoisobutyrate) is further hydrolyzed to $NH_4^{\oplus}$, $HCO_3^{\ominus}$, and a β-amino acid. β-Alanine (from uracil) and β-aminoisobutyrate (from thymine) can then be converted to acetyl CoA and succinyl CoA, respectively, which can enter the citric acid cycle and be converted to other compounds. In bacteria, β-alanine can also be used in the synthesis of pantothenate, a constituent of coenzyme A.

Summary

1. The synthesis of purine nucleotides is a ten-step pathway that leads to IMP (inosinate). The purine is assembled on a foundation of ribose 5-phosphate donated by 5-phosphoribosyl 1-pyrophosphate (PRPP).
2. IMP can be converted to AMP or GMP.
3. In the six-step synthesis of the pyrimidine nucleotide UMP, PRPP enters the pathway after completion of the ring structure.
4. CTP is formed by the amination of UTP.
5. Deoxyribonucleotides are synthesized by the reduction of ribonucleotides at C-2′ in a reaction catalyzed by ribonucleotide reductase.
6. Thymidylate (dTMP) is formed from deoxyuridylate (dUMP) by a methylation reaction in which 5,10-methylenetetrahydrofolate donates both a one-carbon group and a hydride ion. 7,8-Dihydrofolate, the other product of this methylation, is recycled by NADPH-dependent reduction to the active coenzyme tetrahydrofolate.
7. PRPP reacts with pyrimidines and purines in salvage reactions to yield nucleoside monophosphates. Nucleotides and their constituents are interconverted by a variety of enzymes.
8. Nitrogen from amino acids and purine nucleotides is excreted as uric acid in birds and some reptiles. Primates degrade purines to uric acid (urate). Most other organisms further catabolize urate to allantoin, allantoate, urea, or ammonia.
9. Pyrimidines are catabolized to ammonia, bicarbonate, and either acetyl CoA (from cytosine or uracil) or succinyl CoA (from thymine).

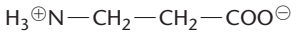
Problems

1. Indicate where the label appears in the product for each of the following precursor–product pairs:
   (a) ^{15}N-aspartate → AMP
   (b) 2-[^{14}C]-glycine → AMP
   (c) δ-[^{15}N]-glutamine → GMP
   (d) 2-[^{14}C]-aspartate → UMP
   (e) $H^{14}CO_3^{\ominus}$ → UMP
2. How many ATP equivalents are needed to synthesize one molecule of IMP, starting from ribose 5-phosphate? Assume that all necessary precursors in the pathway are present.
3. The incorporation of one-carbon units in the *de novo* pathways of purines and pyrimidines requires tetrahydrofolate (THF) derivatives as donors. List the reactions requiring THF derivatives, indicate the THF donor, and indicate which carbon of the purine or pyrimidine is derived from THF.
4. The glutamine analog acivicin, a potential anticancer agent, slows the rapid growth of cells by inhibiting nucleotide biosynthesis.
   (a) Show how acivicin structurally resembles glutamine.
   (b) What intermediate accumulates in the purine biosynthetic pathway when acivicin is present?
   (c) What enzyme is inhibited in the pyrimidine biosynthetic pathway when acivicin is present?



Acivicin

5. A hypothetical bacterium synthesizes UMP by a pathway analogous to the pathway in *E. coli*, except that β-alanine is used instead of aspartate.



β-Alanine

   (a) Why would this pathway be shorter than the pathway in *E. coli*?
   (b) When β-alanine uniformly labeled with ^{14}C is used, where would the label appear in UMP?
6. (a) The enzyme dCMP deaminase can provide a major route from cytidine to uridine nucleotides. What is the product of the action of dCMP deaminase on dCMP?
   (b) This allosteric enzyme is subject to inhibition by dTTP and activation by dCTP. Explain why this is reasonable in terms of the overall cellular needs of nucleoside triphosphates.
7. In eukaryotes, how many ATP equivalents are needed to synthesize one molecule of UMP from $HCO_3^{\ominus}$, aspartate, glutamine, and ribose 5-phosphate? (Ignore any ATP that might be produced by oxidizing the QH_2 generated in the pathway.)
8. Severe combined immunodeficiency syndrome (SCIDS) is characterized by the lack of an immune response to infectious diseases. One form of SCIDS is caused by a deficiency of adenosine deaminase (ADA), an enzyme that catalyzes the deamination of adenosine and deoxyadenosine to produce inosine and deoxyinosine, respectively. The enzyme deficiency increases dATP levels but decreases the levels of other deoxynucleotides, thereby inhibiting DNA replication and cell division in certain rapidly dividing cells. Explain how an adenosine deaminase deficiency affects the levels of deoxynucleotides. (The first effective gene therapy in humans was carried out by transforming a patient's T-cells with a normal ADA gene.)
9. One cause of gout is a deficiency in hypoxanthine–guanine phosphoribosyltransferase activity (Box 18.4). Another cause is due to an *increase* in PRPP synthetase activity. If PRPP is a positive effector of glutamine–PRPP amidotransferase in humans, how does this affect purine synthesis?
10. Identify the nucleotides involved in the following pathways:
    (a) the nucleoside triphosphate required as a substrate in the synthesis of NAD
    (b) the nucleoside triphosphate required in the synthesis of FMN
    (c) the nucleoside triphosphate that serves as a substrate in the synthesis of coenzyme A
    (d) the substrate for G proteins
    (e) the nucleotide used in the synthesis of glycogen from glucose 6-phosphate
    (f) the cofactor required in the reaction catalyzed by mammalian succinyl-CoA synthetase
    (g) the cosubstrate required for the synthesis of phosphatidylserine from phosphatidate

(h) the nucleotide required for activation of galactose in cerebroside biosynthesis
(i) the nucleotide substrate used in histidine biosynthesis
(j) the common precursor of AMP and GMP
(k) the precursor of hypoxanthine

11. The catabolism of fats and carbohydrates provides considerable metabolic energy in the form of ATP. Does the degradation of purines and pyrimidines provide a significant source of energy in eukaryotic cells?

12. PPRP synthetase uses α-D-ribose 5-phosphate as a substrate. How is the α isomer formed inside the cell?

13. The systematic names of the common bases are given in Sections 18.1 and 18.2. What are the systematic names of xanthine, hypoxanthine, and orotate?

14. The sequential action of adenylosuccinate synthetase and adenylosuccinate lyase results in the transfer of an amino group from aspartate and the release of fumarate. Identify two other pairs of enzymes that accomplish the same goal.

Selected Readings

Purine Metabolism

Honzatko, R. B., Stayton, M. M., and Fromm, H. J. (1999). Adenylosuccinate synthetase: recent developments. *Adv. Enzymol. Relat. Areas Mol. Biol.* 73:57–102.

Cendron, L., Berni, R., Folli, C., Ramazzina, I., Percudani, R., and Zanotti, G. (2007). The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J Biol. Chem.* 282:18182–18189.

Kresge, N., Simoni, R. D., and Hill, R. L. (2006). Biosynthesis of purines: the work of John M. Buchanan. *J. Biol. Chem.* 281:e35–e36.

Ramazzina, I., Folli, C., Secchi, A., Berni, R., and Percudani, R. (2006). Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat Chem Biol.* 2:144–148.

Tipton, P. A. (2006). Urate to allantoin, specifically (S)-allantoin. *Nat. Chem. Biol.* 2:124–125.

Tsai. M., Koo, J., Yip, P., Colman, R. F., Segall, M. L., Howell, P. L. (2007). Substrate and product complexes of *Escherichia coli* adenylosuccinate lyase provide new insights into the enzymatic mechanism. *J. Mol. Biol.* 370:541–554.

Zhang, R.-G., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A., and Collart, F. R. (1999). Characteristics and crystal structure of bacterial inosine-5′-monophosphate dehydrogenase. *Biochem.* 38:4691–4700.

Pyrimidine Metabolism

Blakley, R. L. (1995). Eukaryotic dihydrofolate reductase. *Adv. Enzymol. Relat. Areas Mol. Biol.* 70:23–102.

Carreras, C. W., and Santi, D. V. (1995). The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.* 64:721–762.

Chan, R. S., Sakash, J. B., Macol, C. P., West, J. M., Tsuruta, H., and Kantrowitz, E. R. (2002). The role of intersubunit interactions for the stabilization of the T state of *Escherichia coli* aspartate transcarbamoylase. *J. Biol. Chem.* 277:49755–49760.

Lipscomb, W. N. (1994). Aspartate transcarbamoylase from *Escherichia coli*: activity and regulation. *Adv. Enzymol. Relat. Areas Mol. Biol.* 68:67–151.

Raushel, F. M., Thoden, J. B., and Holden, H. M. (1999). The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia. *Biochem.* 38:7891–7899.

Stroud, R. M. (1994). An electrostatic highway. *Struct. Biol.* 1:131–134.

Ribonucleotide Reduction

Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B. M., and Eklund, H. (1997). Binding of allosteric effectors to ribonucleotide reductase protein R1: reduction of active-site cysteines promotes substrate binding. *Structure* 5:1077–1092.

Gorlatov, S. N., and Stadtman, T. C. (1998). Human thioredoxin reductase from HeLa cells: selective alkylation of selenocysteine in the protein inhibits enzyme activity and reduction with NADPH influences affinity to heparin. *Proc. Natl. Acad. Sci. USA.* 95:8520–8525.

Jordan, A., and Reichard, P. (1998). Ribonucleotide reductases. *Annu. Rev. Biochem.* 67:71–98.

Kresge, N., Simoni, R. D., and Hill, R. L. (2006). Peter Reichard and the reduction of ribonucleosides. *J. Biol. Chem.* 281:e13–e15.

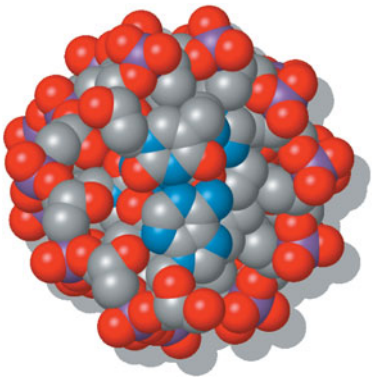
Nordland, P. and Reichard, P. (2006). Ribonucleotide reductases. *Annu. Rev. Biochem.* 75:681–706.

Sjöberg, B.M. (2010). A never-ending story. *Science* 329:1475–1476.

Stubbe, J. (1998). Ribonucleotide reductases in the twenty-first century. *Proc. Natl. Acad. Sci. USA.* 95:2723–2724.

Uppsten, M., Färnegårdh, M., Domkin, V., and Uhlin, U. (2006). The first holocomplex structure of ribonucleotide reductase gives new insight into its mechanism of action. *J. Mol. Biol.* 359:365–377.





Nucleic Acids

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

—J.D. Watson and F.H.C. Crick (1953)

The discovery of the substance that proved to be deoxyribonucleic acid (DNA) was made in 1869 by Friedrich Miescher, a young Swiss physician working in the laboratory of the German physiological chemist Felix Hoppe-Seyler. Miescher treated white blood cells (which came from the pus on discarded surgical bandages) with hydrochloric acid to obtain nuclei for study. When the nuclei were subsequently treated with acid, a precipitate formed that contained carbon, hydrogen, oxygen, nitrogen, and a high percentage of phosphorus. Miescher called the precipitate "nuclein" because it came from nuclei. Later, when it was found to be strongly acidic, its name was changed to nucleic acid. Although he did not know it, Miescher had discovered DNA. Soon afterward, Hoppe-Seyler isolated a similar substance from yeast cells—this substance is now known to be ribonucleic acid (RNA). Both DNA and RNA are polymers of nucleotides, or polynucleotides.

In 1944 Oswald Avery, Colin MacLeod, and Maclyn McCarty demonstrated that DNA is the molecule that carries genetic information. At the time, very little was known about the three-dimensional structure of this important molecule. Over the next few years, the structures of nucleotides were determined and in 1953 James D. Watson and Francis H. C. Crick proposed their model for the structure of double-stranded DNA.

The study of nucleic acid biochemistry has advanced considerably in the past few decades. Today it is possible not only to determine the sequence of your genome but also to synthesize large chromosomes in the laboratory. It has become routine to clone and manipulate DNA molecules. This has led to spectacular advances in our understanding of molecular biology and the ways information contained in DNA is expressed in living cells.

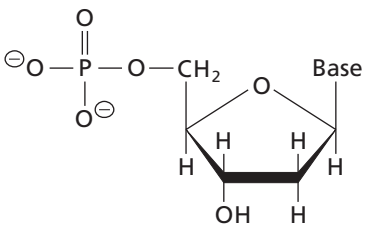
We now know that a living organism contains a set of instructions for every step required to construct a replica of itself. This information resides in the genetic material, or **genome**, of the organism. The genomes of all cells are composed of DNA but some viral genomes are composed of RNA. A genome may consist of a single molecule of DNA, as in many species of bacteria. The genome of eukaryotes is one complete set of DNA molecules found in the nucleus (i.e., the haploid set of chromosomes in diploid organisms). By convention, the genome of a species does not include mitochondrial and chloroplast DNA. With rare exception, no two individuals in a species have exactly

Top: Space-filling model of DNA, viewed along the helix axis.

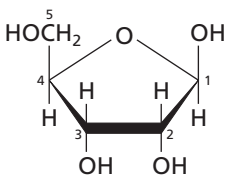


▲ James D. Watson (1928–) (left) and Francis H. C. Crick (1916–2004) (right) describing the structure of DNA in 1953.

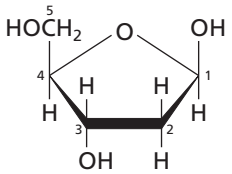
The distinction between the normal flow of information and the Central Dogma of Molecular Biology is explained in Section 1.1 and the introduction to Chapter 21.



▲ **Figure 19.1**
Chemical structure of a nucleotide. Nucleotides contain a five-carbon sugar, a nitrogenous base, and at least one phosphate group. The sugar can be either deoxyribose (shown here) or ribose.



Ribose
(β-D-Ribofuranose)



Deoxyribose
(2-Deoxy-β-D-ribofuranose)

▲ **Figure 19.2**
Chemical structures of the two sugars found in nucleotides. (a) Ribose (β-D-ribofuranose). **(b)** Deoxyribose (2-deoxy-β-D-ribofuranose).

the same genome sequence. If they were alive today, Miescher and Hoppe-Seyler would be astonished to learn that criminals could be convicted by DNA fingerprinting and that we have sequenced the complete genomes of thousands of species, including humans.

In general, the information that specifies the primary structure of a protein is encoded in the sequence of nucleotides in DNA. This information is enzymatically copied during the synthesis of RNA, a process known as transcription. Some of the information contained in the transcribed RNA molecules is translated during the synthesis of polypeptide chains that are then folded and assembled to form protein molecules. Thus, we can generalize that the biological information stored in a cell's DNA flows from DNA to RNA to protein.

Nucleic acids are the fourth major class of macromolecules that we study in this book. Like proteins and polysaccharides, they contain multiple similar monomeric units that are covalently joined to produce large polymers. In this chapter we describe the structure of nucleic acids and how they are packaged in cells. We also describe some of the enzymes that use DNA and RNA as substrates. Many other proteins and enzymes interact with DNA and RNA in order to ensure that genetic information is correctly interpreted. We will consider the biochemistry and the regulation of this flow of information in Chapters 20 to 22.

19.1 Nucleotides Are the Building Blocks of Nucleic Acids

Nucleic acids are polynucleotides, or polymers of nucleotides. As we saw in the previous chapter, nucleotides have three components: a five-carbon sugar, one or more phosphate groups, and a weakly basic nitrogenous compound called a base (Figure 19.1). The bases found in nucleotides are substituted pyrimidines and purines. The pentose is either ribose (D-ribofuranose) or 2-deoxyribose (2-deoxy-D-ribofuranose). The pyrimidine or purine *N*-glycosides of these sugars are called nucleosides. Nucleotides are the phosphate esters of nucleosides—the common nucleotides contain from one to three phosphoryl groups. Nucleotides containing ribose are called ribonucleotides and nucleotides containing deoxyribose are called deoxyribonucleotides (Section 18.5).

A. Ribose and Deoxyribose

The sugar components of the nucleotides found in nucleic acids are shown in Figure 19.2. Both sugars are shown as Haworth projections of the β-conformation of the furanose ring forms (Section 8.2). This is the stable conformation found in nucleotides and polynucleotides. Each of these furanose rings can adopt different conformations such as the envelope forms discussed in Chapter 8. The 2′-endo conformation of deoxyribose predominates in double-stranded DNA (Figure 8.11).

B. Purines and Pyrimidines

The bases found in nucleotides are derivatives of either pyrimidine or purine (Chapter 18). The structures of these heterocyclic compounds and the numbering systems for the carbon and nitrogen atoms of each base are shown in Figure 19.3. Pyrimidine has a single ring containing four carbon and two nitrogen atoms. Purine has a fused pyrimidine-imidazole ring system. Both types of bases are unsaturated, with conjugated double bonds. This feature makes the rings planar and also accounts for their ability to absorb ultraviolet light.

Substituted purines and pyrimidines are ubiquitous in living cells but the unsubstituted bases are seldom encountered in biological systems. The major pyrimidines that occur in nucleotides are uracil (2,4-dioxopyrimidine, U), thymine (2,4-dioxo-5-methylpyrimidine, T), and cytosine (2-oxo-4-aminopyrimidine, C). The major purines are adenine (6-aminopurine, A) and guanine (2-amino-6-oxopurine, G). The chemical structures of these five major bases are shown in Figure 19.4. Note that thymine can also be called 5-methyluracil because it is a substituted form of uracil (Section 18.6).

Adenine, guanine, and cytosine are found in both ribonucleotides and deoxyribonucleotides. Uracil is found mainly in ribonucleotides and thymine is found mainly in deoxyribonucleotides.

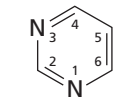
Purines and pyrimidines are weak bases and are relatively insoluble in water at physiological pH. Within cells, however, most pyrimidine and purine bases occur as constituents of nucleotides and polynucleotides and these compounds are highly soluble.

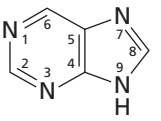
Each heterocyclic base can exist in at least two tautomeric forms. Adenine and cytosine can exist in either amino or imino forms. Guanine, thymine, and uracil can exist in either lactam (keto) or lactim (enol) forms (Figure 19.5). The tautomeric forms of each base exist in equilibrium but the amino and lactam tautomers are more stable and therefore predominate under the conditions found inside most cells. Note that the rings remain unsaturated and planar in each tautomer.

All of the bases in the common nucleotides can participate in hydrogen bonding. The amino groups of adenine and cytosine are hydrogen donors and the ring nitrogen atoms (N-1 in adenine and N-3 in cytosine) are hydrogen acceptors (Figure 19.6). Cytosine also has a hydrogen acceptor group at C-2. Guanine, cytosine, and thymine can form three hydrogen bonds. In guanine, the group at C-6 is a hydrogen acceptor while N-1 and the amino group at C-2 are hydrogen donors. In thymine, the groups at C-4 and C-2 are hydrogen acceptors and N-3 is a hydrogen donor. (Only two of these sites, C-4 and N-3, are used to form base pairs in DNA.) The hydrogen-bonding ability of uracil, a base found in RNA, is similar to that of thymine. The hydrogen-bonding patterns of bases have important consequences for the three-dimensional structure of nucleic acids.

Biochemistry textbooks in the 1940s usually depicted the bases in their imino and lactim forms. These were the structures that Jim Watson was using in 1953 to build a model of DNA. Shortly after being told by Jerry Donohue that the textbooks were wrong, Watson discovered the now-famous A/T and G/C base pairs.

Additional hydrogen bonding occurs in some nucleic acids and in nucleic acid–protein interactions. For example, N-7 of adenine and guanine can be a hydrogen acceptor and both amino hydrogen atoms of adenine, guanine, and cytosine can be donated to form hydrogen bonds.

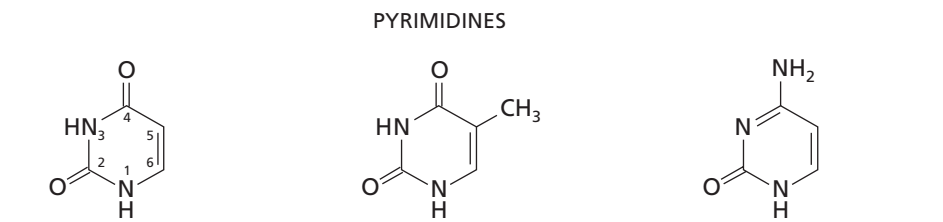


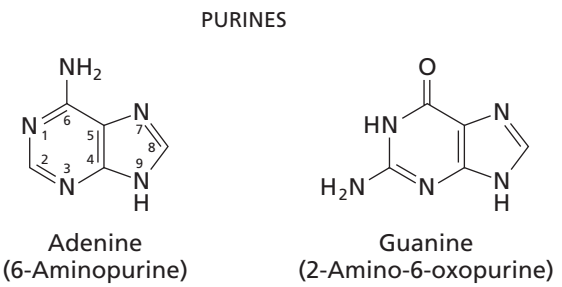


▲ **Figure 19.3**
Chemical structures of pyrimidine and purine.

C. Nucleosides

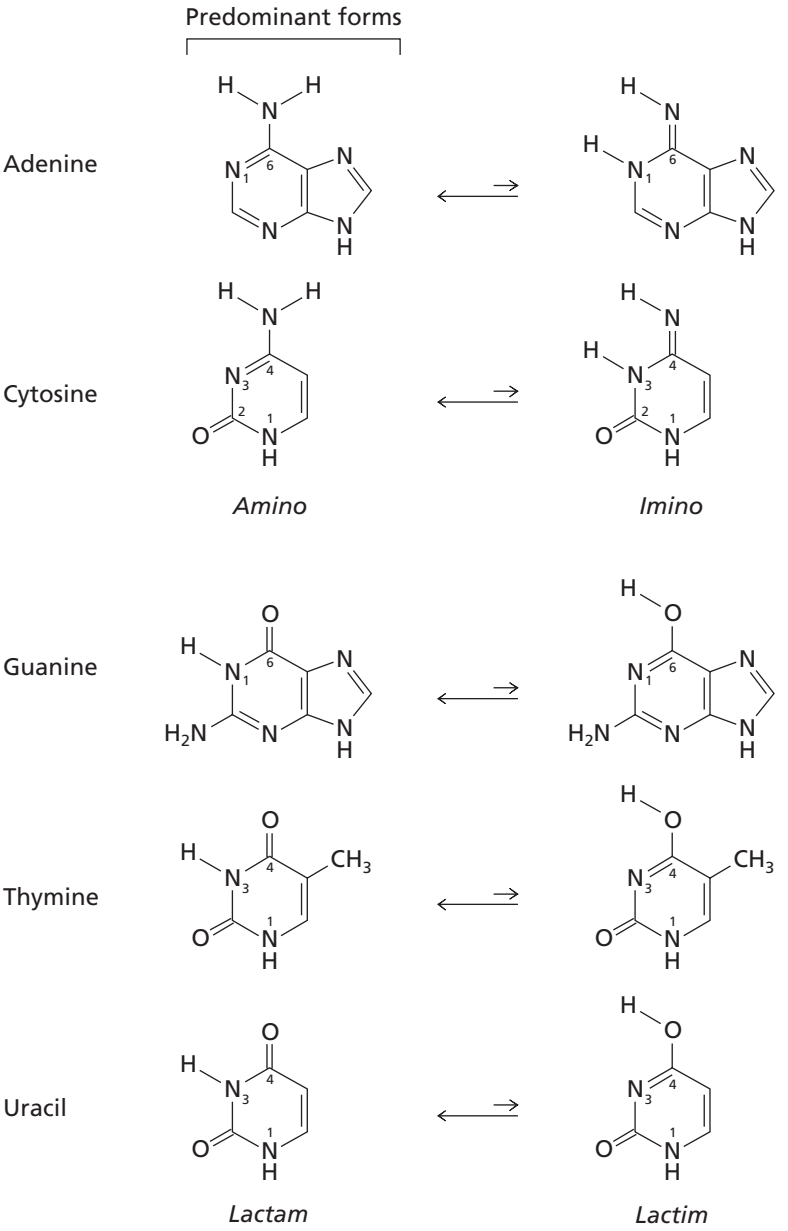
Nucleosides are composed of ribose or deoxyribose and a heterocyclic base. In each nucleoside, a *β*-*N*-glycosidic bond connects C-1 of the sugar to N-1 of the pyrimidine or N-9 of the purine. Nucleosides are therefore *N*-ribosyl or *N*-deoxyribosyl derivatives of pyrimidines or purines. The numbering convention for carbon and nitrogen atoms in

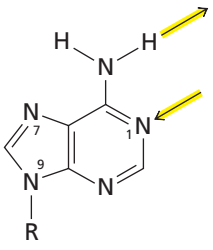


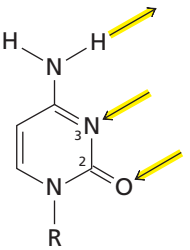


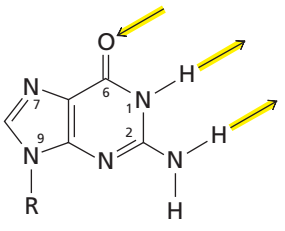
◀ **Figure 19.4**
Chemical structures of the major pyrimidines and purines.

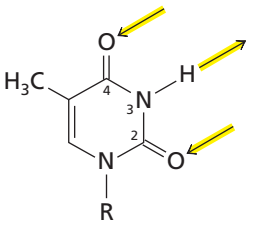
Figure 19.5 ►
Tautomers of adenine, cytosine, guanine, thymine, and uracil. At physiological pH, the equilibria of these tautomerization reactions lie far in the direction of the amino and lactam forms.




(Deoxy)Adenosine


(Deoxy)Cytidine

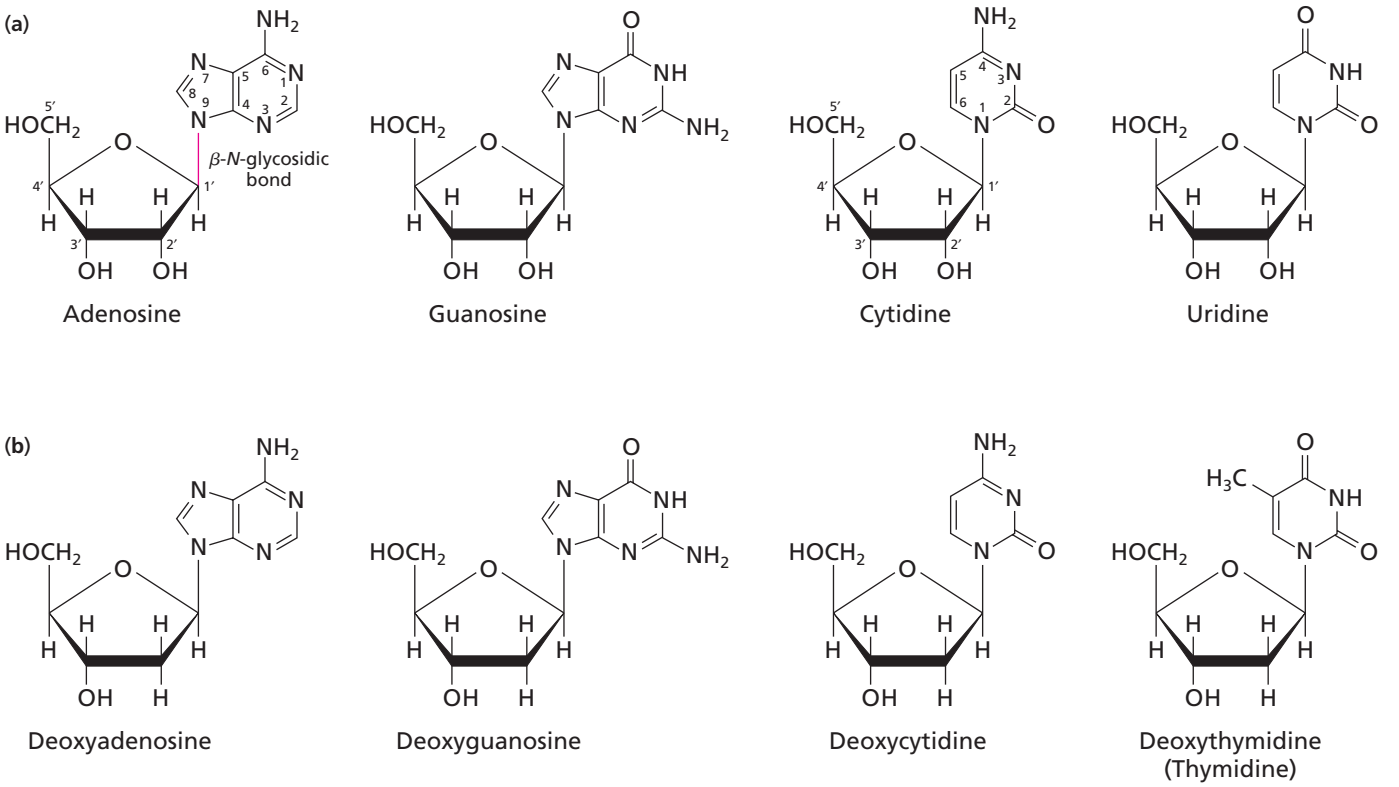

(Deoxy)Guanosine


(Deoxy)Thymidine

nucleosides reflects the fact that they are composed of a base and a five-carbon sugar, each of which has its own numbering scheme. The designation of atoms in the purine or pyrimidine moieties takes precedence. Hence the atoms in the bases are numbered 1, 2, 3, and so on, while those in the furanose ring are distinguished by adding primes (′). Thus, the β-*N*-glycosidic bond connects the C-1′, or 1′, atom of the sugar moiety to the base. Ribose and deoxyribose differ at the C-2′, or 2′, position. The chemical structures of the major ribonucleosides and deoxyribonucleosides are shown in Figure 19.7.

The names of nucleosides are derived from the names of their bases. The ribonucleoside containing adenine is called adenosine (the systematic name, 9-β-D-ribofuranosyladenine, is seldom used)

◄ Figure 19.6
Hydrogen bond sites of bases in nucleic acids. Each base contains atoms and functional groups that can serve as hydrogen donors or acceptors. The common tautomeric forms of the bases are shown. Hydrogen donor and acceptor groups differ in the other tautomers. R represents the sugar moiety.



▲ **Figure 19.7**
Chemical structures of nucleosides. Note that the carbon atoms of the sugars are numbered with primes to distinguish them from the atoms of the bases. **(a)** Ribonucleosides. The sugar in ribonucleosides is ribose, which contains a hydroxyl group at C-2′, as shown here. The *β*-*N*-glycosidic bond of adenosine is shown in red. **(b)** Deoxyribonucleosides. In deoxyribonucleosides, there is an additional hydrogen atom at C-2′ instead of a hydroxyl group.

and its deoxy counterpart is called deoxyadenosine. Similarly, the ribonucleosides of guanine, cytosine, and uracil are guanosine, cytidine, and uridine, respectively. The deoxyribonucleosides of guanine, cytosine, and thymine are deoxyguanosine, deoxycytidine, and deoxythymidine, respectively. Deoxythymidine is often simply called thymidine because thymine rarely occurs in ribonucleosides. The single-letter abbreviations for pyrimidine and purine bases are also commonly used to designate ribonucleosides: A, G, C, and U (for adenosine, guanosine, cytidine, and uridine, respectively). The deoxyribonucleosides are abbreviated dA, dG, dC, and dT when it is necessary to distinguish them from ribonucleosides.

Rotation around the glycosidic bonds of nucleosides and nucleotides is sometimes hindered. There are two relatively stable conformations, *syn* and *anti*, that are in rapid equilibrium (Figure 19.8). In the common pyrimidine nucleosides, the *anti* conformation predominates. The *anti* conformations of all nucleotides predominate in nucleic acids, the polymers of nucleotides.

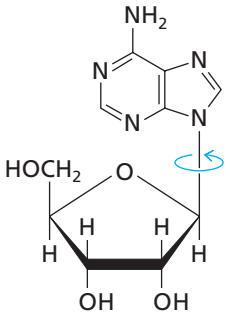
D. Nucleotides

Nucleotides are phosphorylated derivatives of nucleosides. Ribonucleosides contain three hydroxyl groups that can be phosphorylated (2′, 3′, and 5′), and deoxyribonucleosides contain two such hydroxyl groups (3′ and 5′). The phosphoryl groups in naturally occurring nucleotides are usually attached to the oxygen atom of the 5′-hydroxyl group. By convention, a nucleotide is always assumed to be a 5′-phosphate ester unless otherwise designated.

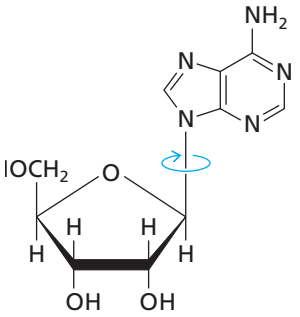
The systematic names for nucleotides indicate the number of phosphate groups present. For example, the 5′-monophosphate ester of adenosine is known as adenosine monophosphate (AMP). It is also simply called adenylate. Similarly, the 5′-monophosphate ester of deoxycytidine can be referred to as deoxycytidine monophosphate (dCMP) or deoxycytidylate. The 5′-monophosphate ester of the deoxyribonucleoside of thymine is usually known as thymidylate but is sometimes called deoxythymidylate to avoid ambiguity. Table 19.1 presents an overview of the nomenclature of bases, nucleosides, and 5′-nucleotides. Nucleotides with the phosphate esterified at the 5′ position are abbreviated AMP, dCMP, and so on. Nucleotides with the phosphate esterified at a position other than 5′ are given similar abbreviations but with position numbers designated (e.g., 3′-AMP).

KEY CONCEPT

By convention the numbering of the atoms in the base takes precedence so the carbon atoms in the sugar are numbered 1′ (“one prime”), 2′ (“two prime”), etc.



syn Adenosine



anti Adenosine

▲ Figure 19.8
***Syn* and *anti* conformations of adenosine.** Some nucleosides assume either the *syn* or *anti* conformation. The *anti* form is usually more stable in pyrimidine nucleosides.

Table 19.1 Nomenclature of bases, nucleosides, and nucleotides

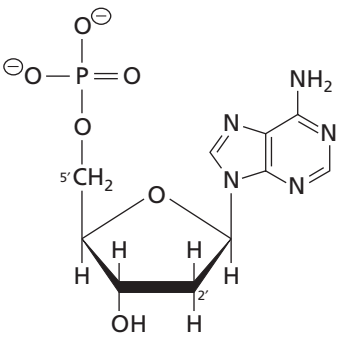
Base	Ribonucleoside	Ribonucleotide (5′-monophosphate)
Adenine (A)	Adenosine	Adenosine 5′-monophoshate (AMP); adenylate[a]
Guanine (G)	Guanosine	Guanosine 5′-monophosphate (GMP); guanylate[a]
Cytosine (C)	Cytidine	Cytidine 5′-monophosphate (CMP); cytidylate[a]
Uracil (U)	Uridine	Uridine 5′-monophosphate (UMP); uridylate[a]
Base	**Deoxyribonucleoside**	**Deoxyribonucleotide (5′-monophosphate)**
Adenine (A)	Deoxyadenosine	Deoxyadenosine 5′-monophosphate (dAMP); deoxyadenylate[a]
Guanine (G)	Deoxyguanosine	Deoxyguanosine 5′-monophosphate (dGMP); deoxyguanylate[a]
Cytosine (C)	Deoxycytidine	Deoxycytidine 5′-monophosphate (dCMP); deoxycytidylate[a]
Thymine (T)	Deoxythymidine or thymidine	Deoxythymidine 5′-monophosphate (dTMP); deoxythymidylate[a] or thymidylate[a]

[a]Anionic forms of phosphate esters predominant at pH 7.4.

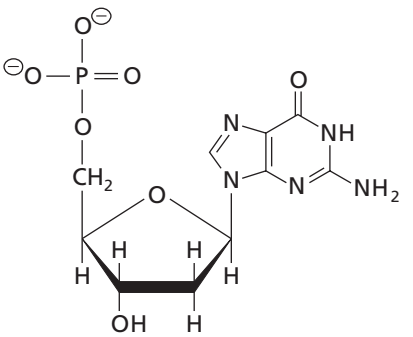
Nucleoside monophosphates, which are derivatives of phosphoric acid, are anionic at physiological pH. They are dibasic acids under physiological conditions since the pK_a values are approximately 1 and 6. The nitrogen atoms of the heterocyclic rings can also ionize.

Nucleoside monophosphates can be further phosphorylated to form nucleoside diphosphates and nucleoside triphosphates. These additional phosphoryl groups are present as phosphoanhydrides. The chemical structures of the deoxyribonucleoside-5′-monophosphates are shown in Figure 19.9. A three-dimensional view of the structure

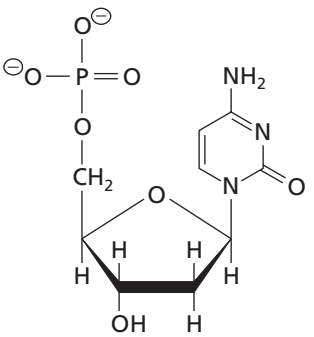
Figure 19.9 ►
Chemical structures of the deoxyribonucleoside-5′-monophosphates.



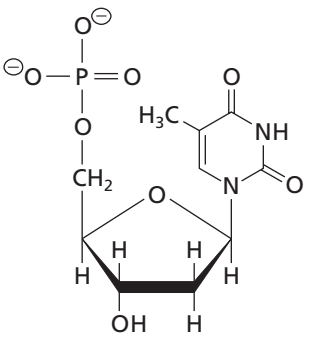
2′-Deoxyadenosine 5′-monophosphate (Deoxyadenylate, dAMP)



2′-Deoxyguanosine 5′-monophosphate (Deoxyguanylate, dGMP)



2′-Deoxycytidine 5′-monophosphate (Deoxycytidylate, dCMP)



2′-Deoxythymidine 5′-monophosphate (Thymidylate, dTMP)

of dGMP is shown in Figure 19.10. The base in dGMP is in the *anti* conformation and the sugar ring is puckered. The plane of the purine ring is almost perpendicular to that of the furanose ring. The phosphoryl group attached to the 5′-carbon atom is positioned well above the sugar and far away from the base.

Nucleoside polyphosphates and polymers of nucleotides can also be abbreviated using a scheme in which the phosphate groups are represented by "p" and the nucleosides are represented by their one-letter abbreviations. The position of the "p" relative to the nucleoside abbreviation indicates the position of the phosphate—for a 5′ phosphate, the p precedes the nucleoside abbreviation and for a 3′ phosphate, the "p" follows the nucleoside abbreviation. Thus, 5′-adenylate (AMP) can be abbreviated as pA, 3′-deoxyadenylate as dAp, and ATP as pppA.



▲ **Figure 19.10**
Deoxyguanosine-5′-monophosphate (dGMP). Hydrogen atoms have been omitted for clarity. Color key: carbon, black; nitrogen, blue; oxygen, red; phosphorus, purple.

19.2 DNA Is Double-Stranded

By 1950 it was clear that DNA is a linear polymer of 2′-deoxyribonucleotide residues linked by 3′–5′ phosphodiesters. Moreover, Erwin Chargaff had deduced certain regularities in the nucleotide compositions of DNA samples obtained from a wide variety of prokaryotes and eukaryotes. Among other things, Chargaff observed that in the DNA of a given cell, A and T are present in equimolar amounts, as are G and C. An example of modern DNA composition data showing these ratios is presented in Table 19.2. Although A = T and G = C for each species, the total mole percent of (G + C) may differ considerably from that of (A + T). The DNA of some organisms, such as the yeast *Saccharomyces cerevisiae*, is relatively deficient in (G + C) whereas the DNA of other organisms, such as the bacterium *Mycobacterium tuberculosis*, is rich in (G + C). In general, the DNAs of closely related species, such as cows, pigs, and humans, have similar base compositions. The data also shows that the ratio of purines to pyrimidines is always 1:1 in the DNA of all species.

The model of DNA proposed by Watson and Crick in 1953 was based on the known structures of the nucleosides and on X-ray diffraction patterns that Rosalind Franklin and Maurice Wilkins obtained from DNA fibers. The Watson–Crick model accounted for the equal amounts of purines and pyrimidines by suggesting that DNA was double-stranded and that bases on one strand paired specifically with bases on the other strand: A with T and G with C. Watson and Crick's proposed structure is now referred to as the B conformation of DNA, or simply B-DNA.

An appreciation of DNA structure is important for understanding the processes of DNA replication (Chapter 20) and transcription (Chapter 21). DNA is the storehouse of biological information. Every cell contains dozens of enzymes and proteins that bind to DNA recognizing certain structural features, such as the sequence of nucleotides. In the following sections we will see how the structure of DNA allows these proteins to gain access to the stored information.

Table 19.2 Base composition of DNA (mole %) and ratios of bases

Source	A	G	C	T	A/T[a]	G/C[a]	(G + C)	Purine/pyrimidine[a]
Escherichia coli	26.0	24.9	25.2	23.9	1.09	0.99	50.1	1.04
Mycobacterium tuberculosis	15.1	34.9	35.4	14.6	1.03	0.99	70.3	1.00
Yeast	31.7	18.3	17.4	32.6	0.97	1.05	35.7	1.00
Cow	29.0	21.2	21.2	28.7	1.01	1.00	42.4	1.01
Pig	29.8	20.7	20.7	29.1	1.02	1.00	41.4	1.01
Human	30.4	19.9	19.9	30.1	1.01	1.00	39.8	1.01

[a]Deviations from a 1:1 ratio are due to experimental variations.

A. Nucleotides Are Joined by 3′–5′ Phosphodiester Linkages

A linkage group consists of several different covalent bonds.

We have seen that the primary structure of a protein refers to the sequence of its amino acid residues linked by peptide bonds. Similarly, the primary structure of a nucleic acid is the sequence of its nucleotide residues connected by 3′–5′ phosphodiester linkages. A tetranucleotide representing a segment of a DNA chain illustrates such linkages (Figure 19.11). The backbone of the polynucleotide chain consists of the phosphoryl groups, the 5′, 4′, and 3′ carbon atoms, and the 3′ oxygen atom of each deoxyribose. These backbone atoms are arranged in an extended conformation. This makes double-stranded DNA a long, thin molecule, unlike polypeptide chains that can easily fold back on themselves.

All the nucleotide residues within a polynucleotide chain have the same orientation. Thus, polynucleotide chains have directionality, like polypeptide chains. One end of a linear polynucleotide chain is said to be 5′ (because no residue is attached to its 5′-carbon) and the other is said to be 3′ (because no residue is attached to its 3′-carbon). By convention, the direction of a DNA strand is defined by reading across the atoms that make up the sugar residue. Thus, going from the top to the bottom of the strand in

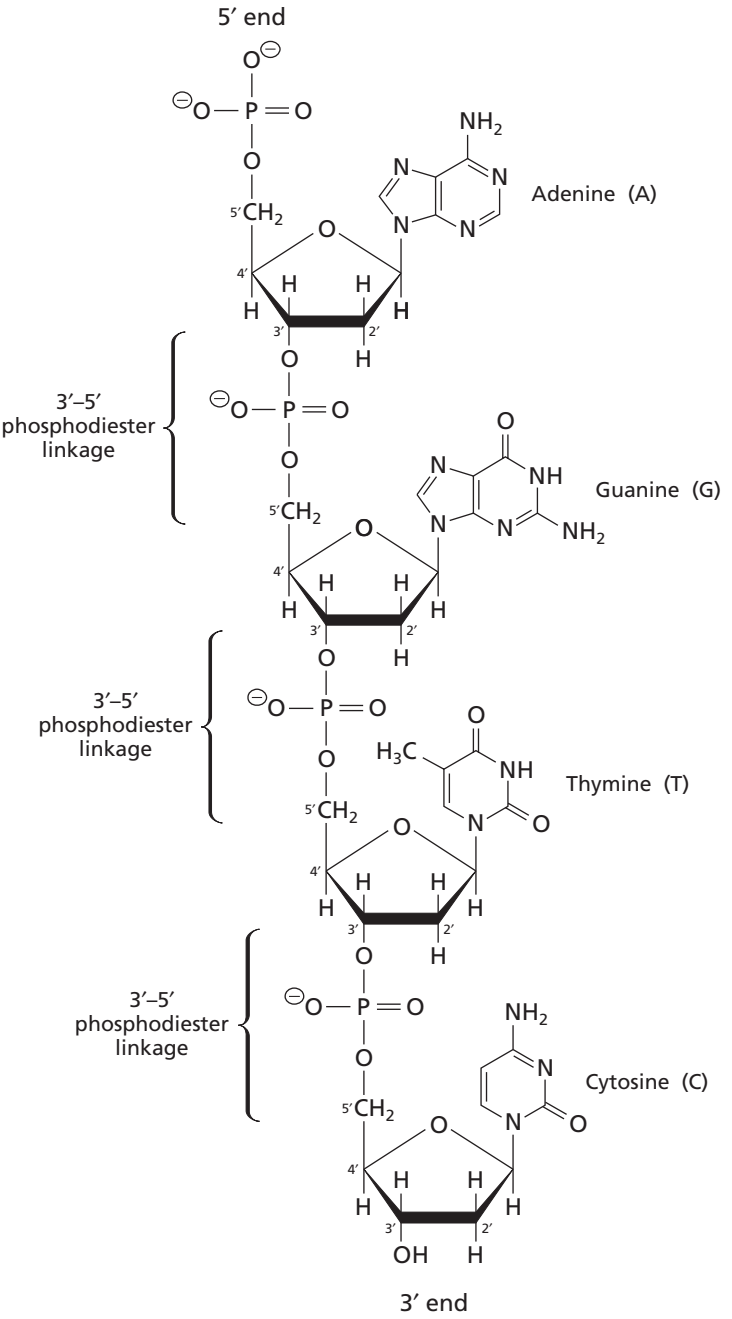


Figure 19.11 ►
Chemical structure of the tetranucleotide pdApdGpdTpdC. The nucleotide residues are joined by 3′–5′ phosphodiester linkages. The nucleotide with a free 5′-phosphoryl group is called the 5′ end, and the nucleotide with a free 3′-hydroxyl group is called the 3′ end.

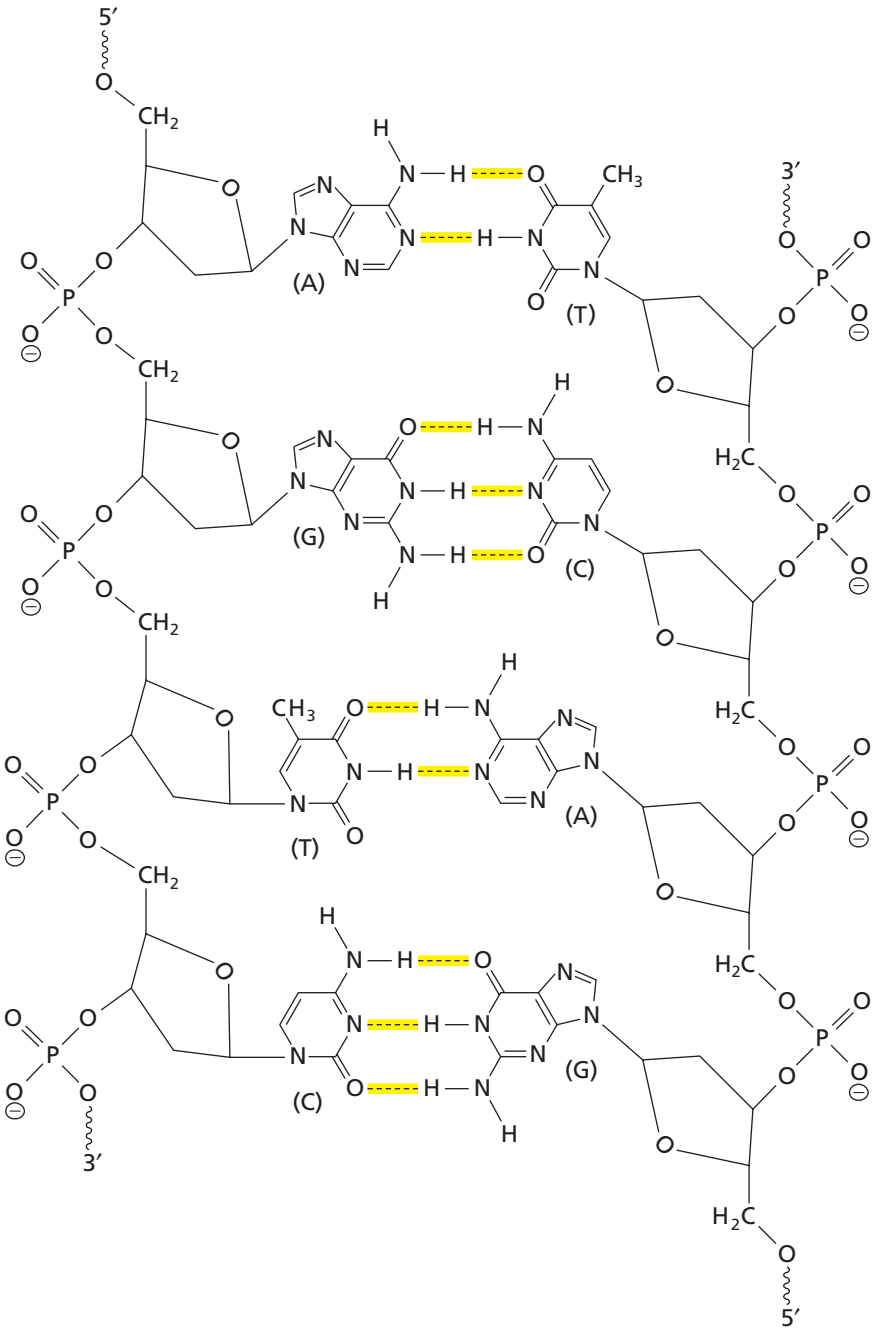
Figure 19.11 is defined as 5′ → 3′ ("five prime to three prime") because one crosses the sugar residue encountering the 5′, 4′, and 3′ carbon atoms, in that order. Similarly, going from the bottom to the top of the strand is moving in the 3′ → 5′ direction.

Structural abbreviations are assumed to read in the 5′ → 3′ direction unless otherwise specified. Because phosphates can be abbreviated as "p," the tetranucleotide in Figure 19.11 can be referred to as pdApdGpdTpdC, or shortened to AGTC when it is clear that the reference is to DNA.

Each phosphate group that participates in a phosphodiester linkage has a pK_a of about 2 and bears a negative charge at neutral pH. Consequently, nucleic acids are polyanions under physiological conditions. Negatively charged phosphate groups are neutralized by small cations and positively charged proteins.

B. Two Antiparallel Strands Form a Double Helix

Most DNA molecules consist of two strands of polynucleotide. Each of the bases on one strand forms hydrogen bonds with a base of the opposite strand (Figure 19.12). The



KEY CONCEPT

The direction of moving along a DNA or RNA strand can be either 5′ → 3′ or 3′ → 5′. It is defined by the direction of reading across the atoms that make up the sugar residue.



▲ **Watson and Crick's original DNA model.**

◀ **Figure 19.12**
Chemical structure of double-stranded DNA. The two strands run in opposite directions. Adenine in one strand pairs with thymine in the opposite strand, and guanine pairs with cytosine.

most common base pairs occur between the lactam and amino tautomers of the bases. Guanine pairs with cytosine and adenine with thymine in a manner that maximizes hydrogen bonding between potential sites. G/C base pairs have three hydrogen bonds and A/T base pairs have two. This feature of double-stranded DNA accounts for Chargaff's discovery that the ratio of A to T and of G to C is 1:1 for a wide variety of DNA molecules. Because A in one strand pairs with T in the other strand and G pairs with C, the strands are complementary and each one can serve as a template for the other.

The sugar–phosphate backbones of the complementary strands of double-stranded DNA have opposite orientations. In other words, they are antiparallel. This was one of the important new insights contributed by Watson and Crick when they built their model of DNA in 1953.

Each end of double-stranded DNA is made up of the 5′ end of one strand and the 3′ end of another. The distance between the two sugar–phosphate backbones is the same for each base pair. Consequently, all DNA molecules have the same regular structure in spite of the fact that their nucleotide sequences may be quite different.

The actual structure of DNA differs in two important aspects from that shown in Figure 19.12. In a true three-dimensional representation, the two strands wrap around each other to form a two-stranded helical structure, or double helix. Also, the bases are rotated so that the plane of the base pairs is nearly perpendicular to the page. (Recall that the plane of the base in dGMP is nearly perpendicular to that of the sugar, as shown in Figure 19.10.)

KEY CONCEPT

The two strands of DNA are anti-parallel.

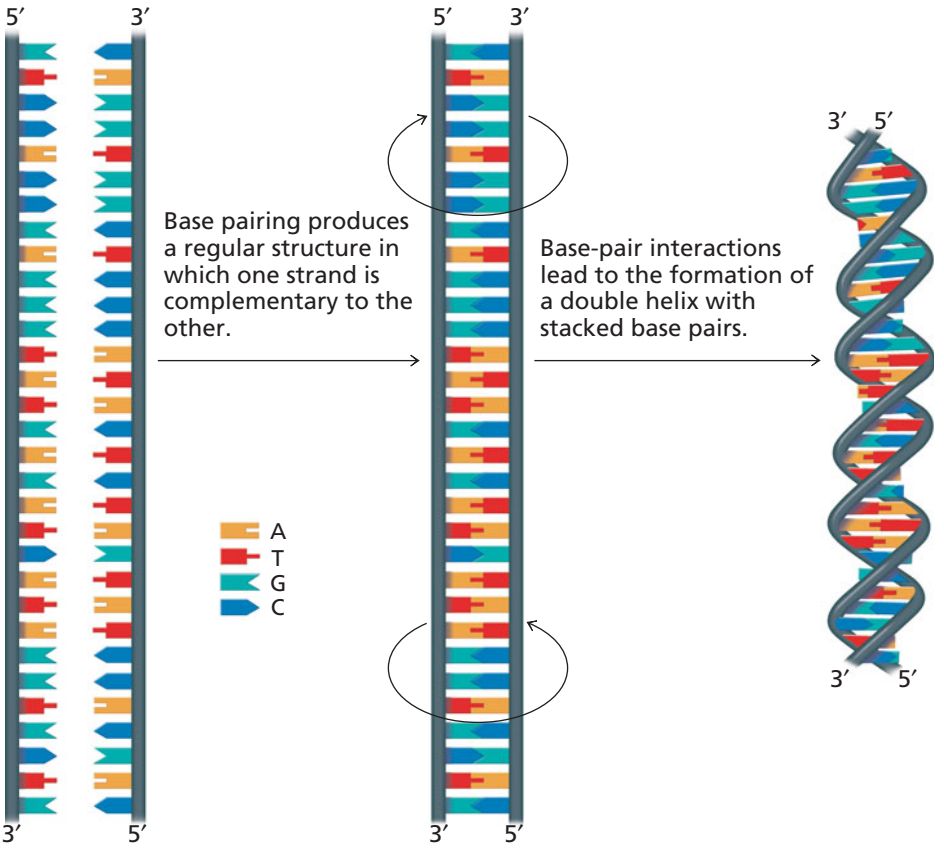
The DNA molecule can be visualized as a "ladder" that has been twisted into a helix. The paired bases represent the rungs of the ladder and the sugar–phosphate backbones represent the supports. Each complementary strand serves as a perfect template for the other. This complementarity is responsible for the overall regularity of the structure of double-stranded DNA. However, complementary base pairing alone does not produce a helix. In B-DNA, the base pairs are stacked one above the other and are nearly perpendicular to the long axis of the molecule. The cooperative, noncovalent interactions between the upper and lower surfaces of each base pair bring the pairs closer together and create a hydrophobic interior that causes the sugar–phosphate backbone to twist. It is these stacking interactions that create the familiar helix (Figure 19.13). Much of the stability of double-stranded DNA is due to the stacking interactions between base pairs.

The two hydrophilic sugar–phosphate backbones wind around the outside of the helix where they are exposed to the aqueous environment. In contrast, the stacked, relatively hydrophobic bases are located in the interior of the helix where they are largely inaccessible to water. This hydrophobic environment makes the hydrogen bonds between bases more stable since they are shielded from competition with water molecules.

The double helix has two grooves of unequal width because of the way the base pairs stack and the sugar–phosphate backbones twist. These grooves are called the **major groove** and the **minor groove** (Figure 19.14). Within each groove, functional groups on the edges of the base pairs are exposed to water. Each base pair has a distinctive pattern of chemical groups projecting into the grooves. Molecules that interact with particular base pairs can identify them by binding in the grooves without disrupting the helix. This is particularly important for proteins that must bind to double-stranded DNA and "read" a specific sequence.

B-DNA is a right-handed helix with a diameter of 2.37 nm. The **rise** of the helix (the distance between one base pair and the next along the helical axis) averages 0.33 nm, and the **pitch** of the helix (the distance to complete one turn) is about 3.40 nm. These values vary to some extent depending on the base composition. Because there are about 10.4 base pairs per turn of the helix, the angle of rotation between adjacent nucleotides within each strand is about 34.6° (360/10.4).

Two views of B-DNA are shown in Figure 19.15. The ball-and-stick model (Figure 19.15a) shows that the hydrogen bonds between base pairs are buried in the interior of the molecule where they are protected from competing interactions with water. The charged phosphate groups (purple and red atoms) are located on the outside surface. This arrangement is more evident in the space-filling model (Figure 19.15b). The space-filling model also clearly shows that functional groups of the base pairs are exposed in



◄ **Figure 19.13**
Complementary base pairing and stacking in double-stranded DNA.

the grooves. These groups can be identified by the presence of blue nitrogen atoms and red oxygen atoms.

The length of double-stranded DNA molecules is often expressed in terms of base pairs (bp). For convenience, longer structures are measured in thousands of base pairs, or **kilobase pairs**, commonly abbreviated kb. Most bacterial genomes consist of a single DNA molecule of several thousand kb; for example, the *Escherichia coli* chromosome is 4600 kb in length. The largest DNA molecules in the chromosomes of mammals and flowering plants may be several hundred thousand kb long. The human genome contains 3,200,000 kb (3.2×10^9 base pairs) of DNA.

Most bacteria have a single chromosome whose ends are joined to create a circular molecule. DNA in the mitochondria and chloroplasts of eukaryotic cells is also circular. In contrast, the chromosomes in the nucleus of a eukaryotic cell are linear. (Some bacteria also have multiple chromosomes and some have linear chromosomes.)

C. Weak Forces Stabilize the Double Helix

The forces that maintain the native conformations of complex cellular structures are strong enough to maintain the structures but weak enough to allow conformational flexibility. Covalent bonds between adjacent residues define the primary structures of proteins and nucleic acids but weak forces determine the three-dimensional shapes

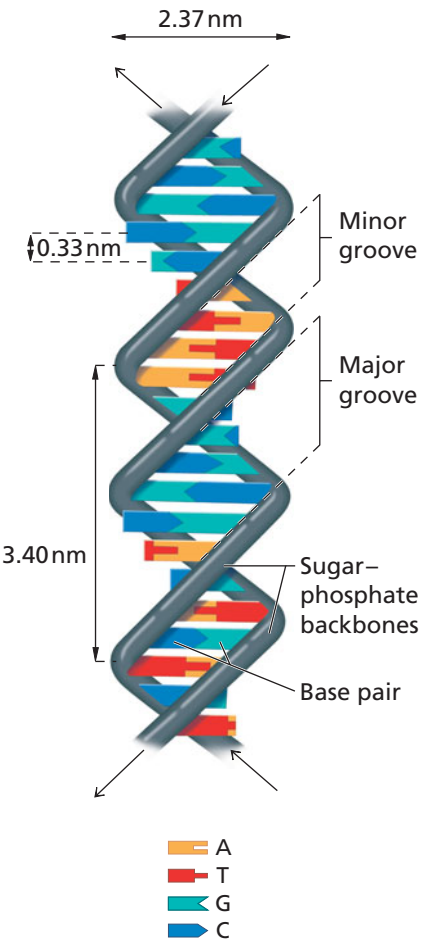
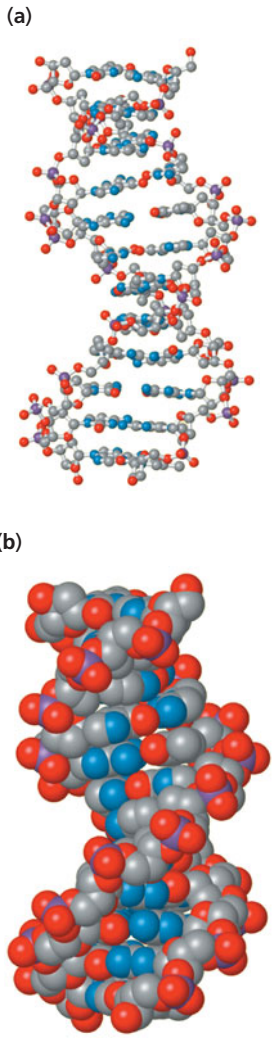


Figure 19.14 ►
Three-dimensional structure of B-DNA. This model shows the orientation of the base pairs and the sugar–phosphate backbone and the relative sizes of the pyrimidine and purine bases. The sugar–phosphate backbone winds around the outside of the helix and the bases occupy the interior. Stacking of the base pairs creates two grooves of unequal width—the major and the minor grooves. The diameter of the helix is 2.37 nm, and the distance between base pairs is 0.33 nm. The distance to complete one turn is 3.40 nm. (For clarity, a slight space has been left between the stacked base pairs and the interactions between complementary bases are shown schematically.)



▲ **Figure 19.15**
B-DNA. (a) Ball-and-stick model. The base pairs are nearly perpendicular to the sugar–phosphate backbones. **(b)** Space-filling model. Color key: carbon, gray; nitrogen, blue; oxygen, red; phosphorus, purple. [Nucleic Acids Database BD0001].

of these macromolecules. Four types of interactions affect the conformation of double-stranded DNA.

1. *Stacking interactions.* The stacked base pairs form van der Waals contacts. Although the forces between individual stacked base pairs are weak, they are additive so in large DNA molecules the van der Waals contacts are an important source of stability.
2. *Hydrogen bonds.* Hydrogen bonding between base pairs is a significant stabilizing force.
3. *Hydrophobic effects.* Burying hydrophobic purine and pyrimidine rings in the interior of the double helix increases the stability of the helix.
4. *Charge–charge interactions.* Electrostatic repulsion of the negatively charged phosphate groups of the backbone is a potential source of instability in the DNA helix. However, repulsion is minimized by the presence of cations such as Mg^{2+} and cationic proteins (proteins that contain an abundance of the basic residues arginine and lysine).

The importance of stacking interactions can be illustrated by examining the various stacking energies of the base pairs (Table 19.3). The stacking energy of two base pairs depends on the nature of the base pair (G/C or A/T) and the orientation of each base pair. Typical stacking energies are about 35 kJ mol^{-1}. Within the hydrophobic core of stacked double-stranded DNA the hydrogen bonds between base pairs have a strength of about 27 kJ mol^{-1} each (Section 2.5B). However, if the stacking interactions are weakened, the hydrogen bonds in the base pairs are exposed to competition from water molecules and the overall contribution to keeping the strands together diminishes greatly.

Under physiological conditions, double-stranded DNA is thermodynamically much more stable than the separated strands and that explains why the double-stranded form predominates *in vivo*. However, the structure of localized regions of the double helix can sometimes be disrupted by unwinding. Such disruption occurs during DNA replication, repair, recombination, and transcription. Complete unwinding and separation of the complementary single strands is called **denaturation**. Denaturation occurs only *in vitro*.

Double-stranded DNA can be denatured by heat or by a chaotropic agent such as urea or guanidinium chloride. (Recall from Section 4.10 that proteins can also be denatured.) In studies of thermal denaturation, the temperature of a solution of DNA is slowly increased. As the temperature is raised, more and more of the bases become unstacked and hydrogen bonds between base pairs are broken. Eventually, the two strands separate completely. The temperature at which half the DNA has become single-stranded is known as the **melting point** (T_m).

Absorption of ultraviolet light can be used to measure the extent of denaturation. Measurements are made at a wavelength of 260 nm—close to the absorbance maximum for nucleic acids. Single-stranded DNA absorbs 12% to 40% more light than double-stranded DNA at 260 nm (Figure 19.16). A plot of the change in absorbance of a DNA solution versus temperature is called a **melting curve** (Figure 19.17). The absorbance increases sharply at the melting point and the transition from double-stranded to single-stranded DNA takes place over a narrow range of temperature.

The sigmoid shape of the melting curve indicates that denaturation is a cooperative process as we saw in the case of protein denaturation (Section 4.10). In this case, cooperativity results from rapid unzippering of the double-stranded molecule as the many stacking interactions and hydrogen bonds are disrupted. The unzippering begins with the unwinding of a short internal stretch of DNA, forming a single-stranded "bubble." This single-stranded bubble rapidly destabilizes the adjacent stacked base pairs and this destabilization is propagated in both directions as the bubble expands.

As shown in Figure 19.17, poly (GC) denatures at a much higher temperature than poly (AT). It is easier to melt A/T-rich DNA than G/C-rich DNA because A/T base pairs have weaker stacking interactions as shown in Table 19.3. It's important to note that the stacking interactions are the first interactions to be disrupted by higher temperature. Once this process begins the hydrogen bonds—although collectively stronger in stacked

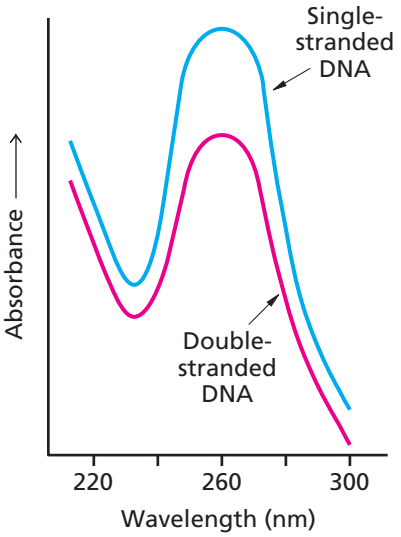


Figure 19.16 ►
Absorption spectra of double-stranded and single-stranded DNA. At pH 7.0, double-stranded DNA has an absorbance maximum near 260 nm. Denatured DNA absorbs 12% to 40% more ultraviolet light than double-stranded DNA.

DNA—become much weaker because they are exposed to water and the DNA is rapidly destabilized. Naturally occurring DNA is a mixture of regions with varying base compositions but A/T-rich regions are more easily unwound than G/C-rich regions.

At temperatures just below the melting point, a typical DNA molecule contains double-stranded regions that are G/C-rich and local single-stranded regions ("bubbles") that are A/T-rich. These *in vitro* experiments demonstrate an important point—that it is easier to unwind localized regions whose sequence consists largely of A/T base pairs rather than G/C base pairs. We will see in Chapter 21 that the initiation sites for transcription are often A/T-rich.

D. Conformations of Double-Stranded DNA

Double-stranded DNA can assume different conformations under different conditions. X-ray crystallographic studies of various synthetic oligodeoxyribonucleotides of known sequence indicate that DNA molecules inside the cell do not exist in a "pure" B conformation. Instead, DNA is a dynamic molecule whose exact conformation depends to some extent on the nucleotide sequence. The local conformation is also affected by bends in the DNA molecule and whether it is bound to protein. As a result, the number of base pairs per turn in B-DNA can fluctuate in the range of 10.2–10.6.

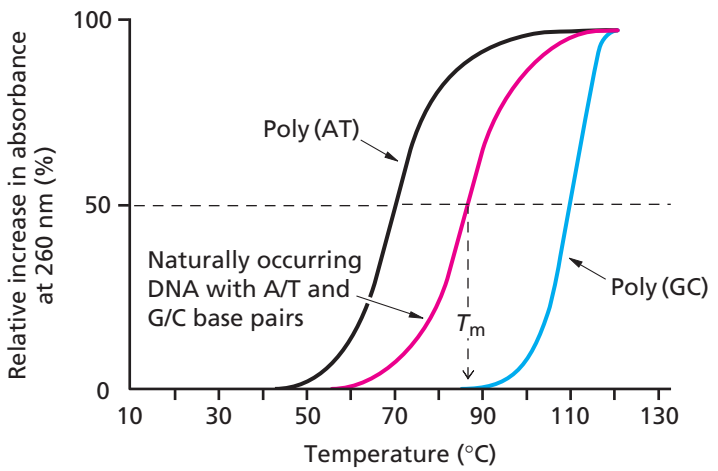
There are two other distinctly different DNA conformations in addition to the various forms of B-DNA. A-DNA forms when DNA is dehydrated and Z-DNA can form when certain sequences are present (Figure 19.18). (The A- and B-DNA forms were discovered by Rosalind Franklin in 1952.) A-DNA is more tightly wound than B-DNA and the

Table 19.3 Stacking interactions for the ten possible combinations in double-stranded DNA

Stacked dimers		Stacking energies (kJ mol^{-1})
↑C-G↓ G-C		–61.0
↑C-G↓ A-T	↑T-A↓ G-C	–44.0
↑C-G↓ T-A	↑A-T↓ G-C	–41.0
↑G-C↓ C-G		–40.5
↑G-C↓ G-C	↑C-G↓ C-G	–34.6
↑T-A↓ A-T		–27.5
↑G-C↓ T-A	↑A-T↓ C-G	–27.5
↑G-C↓ T-A	↑T-A↓ C-G	–28.4
↑A-T↓ A-T	↑T-A↓ T-A	–22.5
↑A-T↓ T-A		–16.0

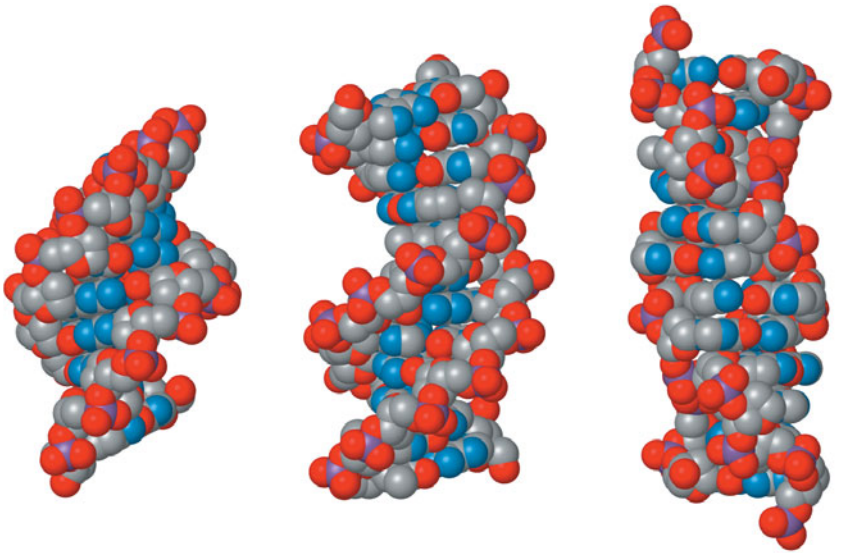
Arrows designate the direction of the sugar-phosphate backbone and point from C-3′ of one sugar unit to C-5′ of the next.

[Adapted from Omstein, R. L., Rein, R., Breen, D. L., and MacElroy, R. D. (1978). An optimized potential function for the calculation of nucleic acid interaction energies: I. Base stacking. *Biopolymers* 17: 2341–2360.]



◄ Figure 19.17 Melting curve for DNA.
In this experiment, the temperature of a DNA solution is increased while the absorbance at 260 nm is monitored. The melting point (T_m) corresponds to the inflection point of the sigmoidal curve where the increase in absorbance of the sample is one-half the increase in absorbance of completely denatured DNA. Poly (AT) melts at a lower temperature than either naturally occurring DNA or poly (GC) since more energy is required to disrupt stacked G/C base pairs.

Figure 19.18 ►
A-DNA, B-DNA, and Z-DNA. The A-DNA conformation (left) is favored when DNA is dehydrated [NDB AD0001]. B-DNA (center) is the conformation normally found inside cells [NDB BD0001]. The Z-DNA conformation (right) is favored in certain G/C-rich sequences [NDB ZDJ050].

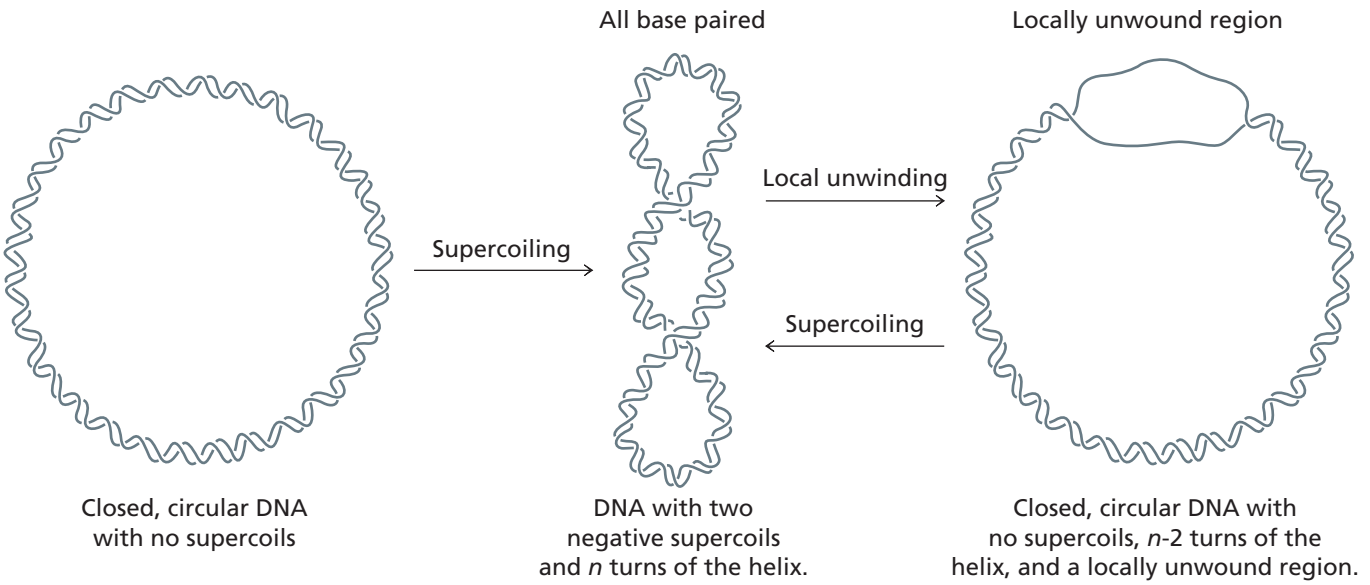


major and minor grooves of A-DNA are similar in width. There are about 11 bp per turn in A-DNA and the base pairs are tilted about 20° relative to the long axis of the helix. Z-DNA differs even more from B-DNA. There are no grooves in Z-DNA and the helix is left-handed, not right-handed. The Z-DNA conformation occurs in G/C-rich regions. Deoxyguanylate residues in Z-DNA have a different sugar conformation (3′-endo) and the base is in the *syn* conformation. A-DNA and Z-DNA conformations exist *in vivo* but they are confined to short regions of DNA.

19.3 DNA Can Be Supercoiled

A circular DNA molecule with the B conformation has an average of 10.4 base pairs per turn. It is said to be *relaxed* if such a molecule would lie flat on a surface. This relaxed double helix can be overwound or underwound if the strands of DNA are broken and the two ends of the linear molecule are twisted in opposite directions. When the strands are rejoined to create a circle, there are no longer 10.4 base pairs per turn as required to maintain the stable B conformation. The circular molecule compensates for over- or underwinding by forming supercoils that restore 10.4 base pairs per turn of the double helix (Figure 19.19). A supercoiled DNA molecule would not lie flat on a surface. Each supercoil compensates for one turn of the double helix.

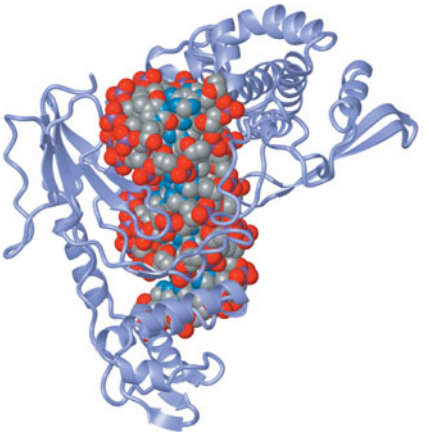
▼ Figure 19.19
Supercoiled DNA. The DNA molecule on the left is a relaxed closed circle and has the normal B conformation. Breaking the DNA helix and unwinding it by two turns before re-forming the circle produces two supercoils. The supercoils compensate for the underwinding and restore the normal B conformation. The molecule on the right has a locally unwound region of DNA. This conformation is topologically equivalent to negatively supercoiled DNA.



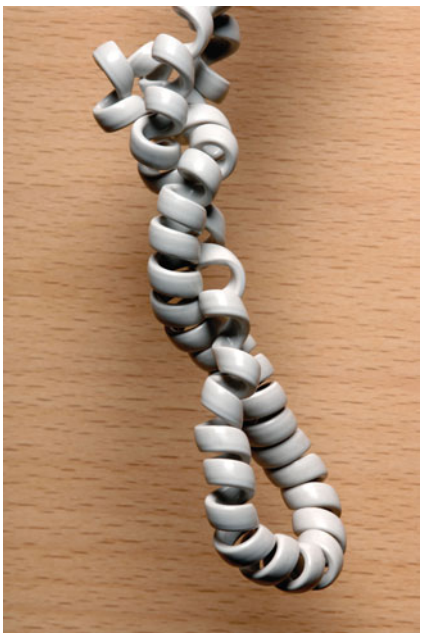
Most circular DNA molecules are supercoiled in cells but even long, linear DNA molecules contain locally supercoiled regions. Bacterial chromosomes typically have about five supercoils per 1000 base pairs of DNA. The DNA in the nuclei of eukaryotic cells is also supercoiled as we will see in Section 19.5. All organisms have enzymes that can break DNA, unwind or overwind the double helix, and rejoin the strands to alter the topology. These enzymes, called topoisomerases, are responsible for adding and removing supercoils. An example of a topoisomerase bound to DNA is shown in Figure 19.20. These remarkable enzymes cleave one or both strands of DNA, unwind or overwind DNA by rotating the cleaved ends, and then rejoin the ends to create (or remove) supercoils.

One of the important consequences of supercoiling is shown in Figure 19.19. If DNA is underwound, it compensates by forming negative supercoils in order to maintain the stable B conformation. (Overwinding produces positive supercoils.) An alternative conformation is shown on the right in Figure 19.19. In this form, most of the DNA is double-stranded but there is a locally unwound region that is due to the slight underwinding. The negatively supercoiled and locally unwound conformations are in equilibrium with the supercoiled form in excess because it is slightly more stable. The difference in free energy between the two conformations is quite small.

Most of the DNA in a cell is negatively supercoiled. This means that it is relatively easy to unwind short regions of the molecule—especially those regions that are A/T-rich. As mentioned earlier, localized unwinding is an essential step in the initiation of DNA replication, recombination, repair, and transcription. Thus, negative supercoiling plays an important biological role in these processes by storing the energy needed for local unwinding. This is why topoisomerases that catalyze supercoiling are essential enzymes in all cells.



▲ **Figure 19.20**
Human (*Homo sapiens*) topoisomerase I bound to DNA. [PDB 1A31]



▲ **Supercoiled telephone cords can be very annoying.**

19.4 Cells Contain Several Kinds of RNA

RNA molecules participate in several processes associated with gene expression. RNA molecules are found in multiple copies and in several different forms within a given cell. There are four major classes of RNA in all living cells:

1. *Ribosomal RNA* (rRNA) molecules are an integral part of ribosomes (intracellular ribonucleoproteins that are the sites of protein synthesis). Ribosomal RNA is the most abundant class of ribonucleic acid accounting for about 80% of the total cellular RNA.
2. *Transfer RNA* (tRNA) molecules carry activated amino acids to the ribosomes for incorporation into growing peptide chains during protein synthesis. tRNA molecules are only 73 to 95 nucleotide residues long. They account for about 15% of the total cellular RNA.
3. *Messenger RNA* (mRNA) molecules encode the sequences of amino acids in proteins. They are the "messengers" that carry information from DNA to the translation complex where proteins are synthesized. In general, mRNA accounts for only 3% of the total cellular RNA. These molecules are the least stable of the cellular ribonucleic acids.
4. *Small RNA* molecules are present in all cells. Some small RNA molecules have catalytic activity or contribute to catalytic activity in association with proteins. Many of these RNA molecules are associated with processing events that modify RNA after it has been synthesized. Some are required for regulating gene expression.

RNAs are single-stranded molecules, but they often have complex secondary structure. Most single-stranded polynucleotides fold back on themselves to form stable regions of base-paired, double-stranded RNA under physiological conditions. One type of secondary structure is a stem–loop which forms when short regions of complementary sequence form base pairs (Figure 19.21). The structure of the double-stranded regions of such stem–loops resembles the structure of the A form of double-stranded DNA. As we will see in Chapters 21 and 22, such structures are important in transcription and are common features in transfer RNA, ribosomal RNA, and the small RNAs.

KEY CONCEPT

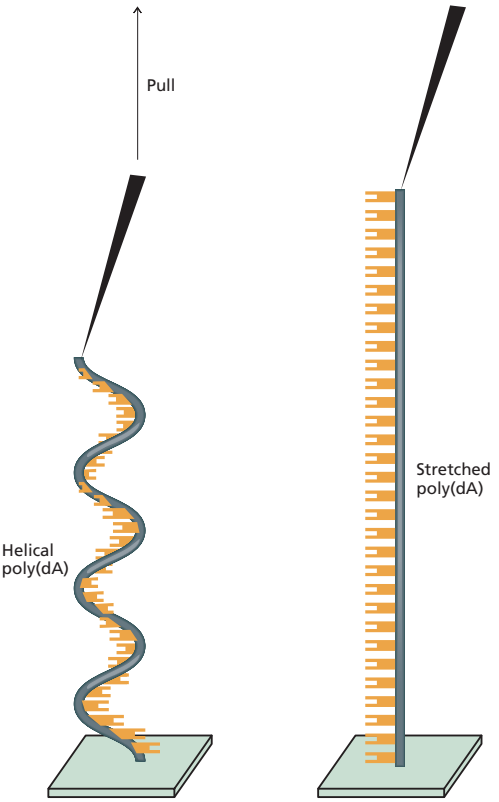
Single-stranded RNA can fold back on itself to create stable double-stranded helical regions that resemble those in DNA.

BOX 19.1 PULLING DNA

Single-molecule atomic-force spectroscopy is a powerful tool for investigating the properties of single molecules. It has been used to explore the properties of single-stranded DNA. The experiment involves fixing one end of a single-stranded DNA molecule to a solid surface and attaching the other end to a form of molecular tweezer that can be used to pull the molecule and measure its resistance.

When this experiment is done with poly(dT) there is almost no resistance until the molecule is in the fully extended form. This is because poly(dT) has no significant secondary structure. However, when poly(dA) is pulled there is initial resistance followed by a shift to the fully extended form. Poly(dA) is helical in solution because the adenylate residues stack on one another and the initial resistance is due to breaking the helix.

The resistance can be measured and the calculated energy of stacking is 15 kJ mol^{-1}, in agreement with other determinations of the stacking interactions of A bases on other A's. The experiment proves that stacking interactions are important in forming helical DNA structures-even with single-stranded polynucleotides.



Pulling poly(dA). [Adapted from Ke et al. (2007)] ▶



▲ **Figure 19.21**
Stem–loop structures in RNA. Single-stranded polynucleotides, such as RNA, can form stem–loops, or hairpins, when short regions of complementary sequence form base pairs. The stem of the structure consists of base-paired nucleotides, and the loop consists of noncomplementary nucleotides. Note that the strands in the stem are antiparallel.

19.5 Nucleosomes and Chromatin

In 1879, ten years after Miescher's discovery of nuclein, Walter Flemming observed banded objects in the nuclei of stained eukaryotic cells. He called the material **chromatin**, from the Greek *chroma*, meaning "color." Chromatin is now known to consist of DNA plus various proteins that package the DNA in a more compact form. Prokaryotic DNA is also associated with protein to form condensed structures inside the cell. These structures differ from those observed in eukaryotes and are usually not called chromatin.

In a normal resting cell, chromatin exists as 30 nm fibers—long, slender threads about 30 nm in diameter. In humans, the nucleus must accommodate 46 such chromatin fibers, or chromosomes. The largest human chromosome is about 2.4×10^8 bp; it would be about 8 cm long if it were stretched out in the B conformation. During metaphase (when chromosomes are most condensed) the largest chromosome is about 10 μm long. The difference between the length of the metaphase chromosome and the extended B form of DNA is 8000-fold. This value is referred to as the *packing ratio*.

A. Nucleosomes

The major proteins of chromatin are known as **histones**. Most eukaryotic species contain five different histones—H1, H2A, H2B, H3, and H4. All five histones are small, basic proteins containing numerous lysine and arginine residues whose positive charges allow the proteins to bind to the negatively charged sugar–phosphate backbone of DNA. The numbers of acidic and basic residues in typical mammalian histones are noted in Table 19.4. Except for H1, the amino acid sequence of each type of histone is highly conserved in all eukaryotes. For example, bovine histone H4 differs from pea histone H4 in only two residues out of 102. Such similarity in primary structure implies a corresponding conservation in tertiary structure and function.

Chromatin unfolds when it is treated with a solution of low ionic strength (<5 mM). The extended chromatin fiber looks like beads on a string in an electron micrograph (Figure 19.22). The "beads" are DNA–histone complexes called **nucleosomes** and the "string" is double-stranded DNA.

Table 19.4 Basic and acidic residues in mammalian histones

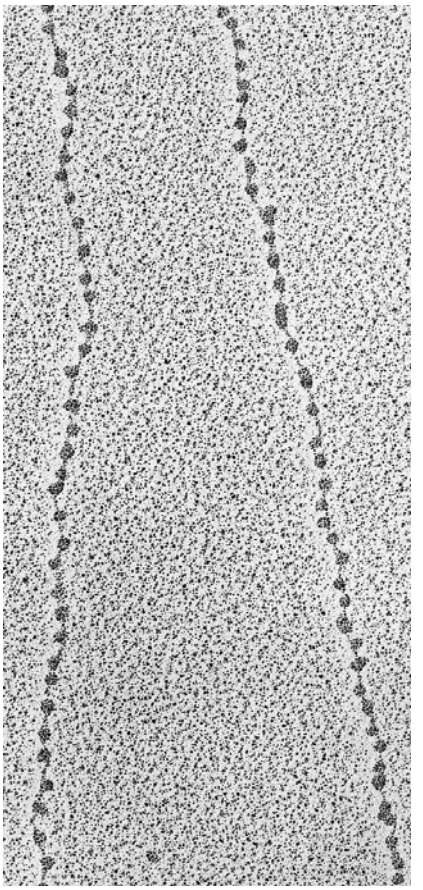
Type	Molecular weight	Number of residues	Number of basic residues	Number of acidic residues
Rabbit thymus H1	21,000	213	65	10
Calf thymus H2A	14,000	129	30	9
Calf thymus H2B	13,800	125	31	10
Calf thymus H3	15,300	135	33	11
Calf thymus H4	11,300	102	27	7

Each nucleosome is composed of one molecule of histone H1, two molecules each of histones H2A, H2B, H3, and H4, and about 200 bp of DNA (Figure 19.23). The H2A, H2B, H3, and H4 molecules form a protein complex called the histone octamer around which the DNA is wrapped. About 146 bp of DNA are in close contact with the histone octamer forming a **nucleosome core particle**. The DNA between particles is called linker DNA; it is about 54 bp long. Histone H1 can bind to the linker DNA and to the core particle but in the extended beads-on-a-string conformation H1 is often absent. Histone H1 is responsible for higher-order chromatin structures.

The structure of the nucleosome core particle has been determined by X-ray crystallography (Figure 19.24). The eight histone subunits are arranged symmetrically as four dimers: two H2A/H2B dimers and two H3/H4 dimers. The particle is shaped like a flat disk with positively charged grooves that accommodate the sugar–phosphate backbone of DNA.

DNA wraps around the core particle forming about $1\frac{3}{4}$ turns per nucleosome. If this DNA were in an extended conformation it would be about 50 nm in length but when bound to the nucleosome core particle, the overall length is reduced to the width of the disk, about 5 nm. The coils of DNA are topologically equivalent to negative supercoils and that's why eukaryotic DNA becomes supercoiled when histones are removed from chromatin.

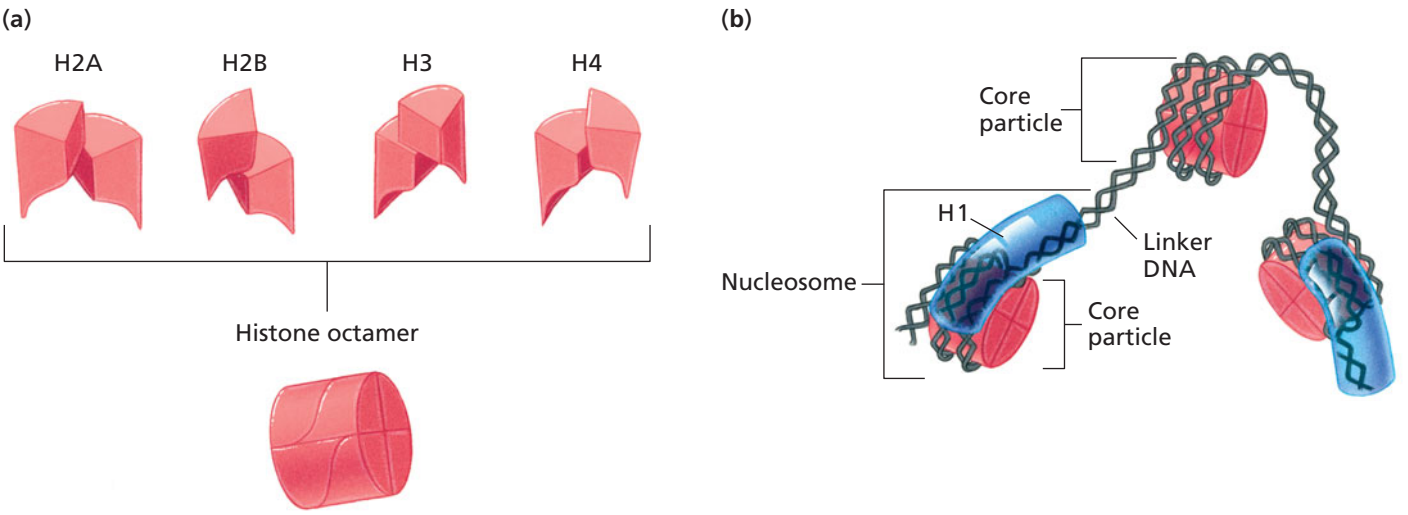
The N-termini of all four core histones are rich in positively charged lysine (K) and arginine (R) residues. These ends extend outward from the core particle where they interact with DNA and negatively charged regions of other proteins (Figure 19.24). These interactions serve to stabilize higher-order chromatin structures such as the 30 nm fiber.



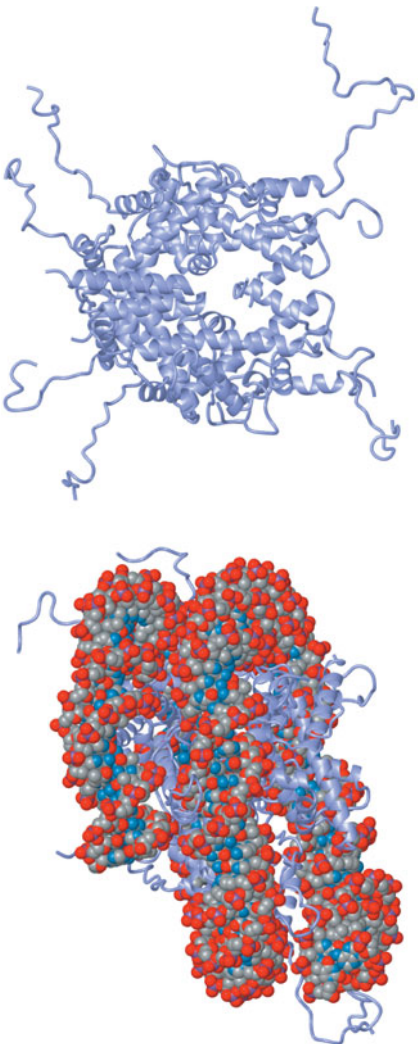
▲ **Figure 19.22**
Electron micrograph of extended chromatin showing the "beads-on-a-string" organization.

KEY CONCEPT

The vast majority of eukaryotic DNA is bound to nucleosome core particles spaced 200 bp apart.



▲ **Figure 19.23**
Diagram of nucleosome structure. (a) Histone octamer. **(b)** Nucleosomes. Each nucleosome is composed of a core particle plus histone H1 and linker DNA. The nucleosome core particle is composed of a histone octamer and about 146 bp of DNA. Linker DNA consists of about 54 bp. Histone H1 binds to the core particle and to linker DNA.



▲ **Figure 19.24**
Structure of the chicken (*Gallus gallus*) nucleosome core particle. (a) Histone octamer. **(b)** Histone octamer bound to DNA—side view showing the disk shape of the particle. [PDB 1EQZ].

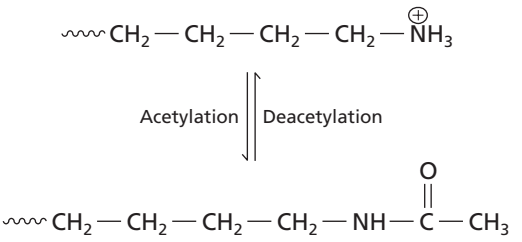
Specific lysine residues in these N-terminal ends can be acetylated by enzymes known as histone acetyltransferases (HATS). For example, residues 5, 8, 12, 16, and 20 in histone H4 can be modified by acetylation.



(19.1)

Acetylation decreases the net positive charge of the histone N-termini and weakens the interactions with other nucleosomes and proteins. The net result is a loosening up of higher-order structures. Acetylation is associated with gene expression. HATS are preferentially directed to sites where chromatin must be unraveled in order to transcribe a gene. The relationship between transcriptional activation and histone acetylation is under active investigation in many laboratories (Section 21.5C).

Histone deacetylases are responsible for removing acetyl groups from lysine residues. This restores the positively charged side chains and allows nucleosomes to adopt the more compact chromatin structure characteristic of regions that are not expressed.



(19.2)

B. Higher Levels of Chromatin Structure

The packaging of DNA into nucleosomes reduces the length of a DNA molecule about tenfold. Further reduction comes from higher levels of DNA packaging. For example, the beads-on-a-string structure is itself coiled into a solenoid to yield the 30 nm fiber. One possible model of the solenoid is shown in Figure 19.25. The 30 nm fiber forms when every nucleosome contains a molecule of histone H1 and adjacent molecules of H1 bind to each other cooperatively bringing the nucleosomes together into a more compact and stable form of chromatin. Condensation of the beads-on-a-string structure into a solenoid achieves a further fourfold reduction in chromosome length.

Finally, 30 nm fibers are themselves attached to an RNA–protein scaffold that holds the fibers in large loops. There may be as many as 2000 such loops on a large chromosome. The RNA–protein scaffold of a chromosome can be seen under an electron microscope when histones have been removed (Figure 19.26). The attachment of DNA loops to the scaffold accounts for an additional 200-fold condensation in the length of DNA.

The loops of DNA are attached to the scaffold at their base. Because the ends are not free to rotate, the loops can be supercoiled. (Some of the supercoils can be seen in Figure 19.26b, but most of the DNA is relaxed because one of the strands is broken during treatment to remove histones.)

C. Bacterial DNA Packaging

Histones are found only in eukaryotes but prokaryotic DNA is also packaged with proteins in a condensed form. Some of these proteins are referred to as histone like proteins because they resemble eukaryotic histones. In most cases, there are no defined nucleosome-like particles in prokaryotes and much of the DNA is not associated with protein. Bacterial DNA is attached to a scaffold in large loops of about 100 kb. This arrangement converts the bacterial chromosome to a structure known as the nucleoid.

19.6 Nucleases and Hydrolysis of Nucleic Acids

Enzymes that catalyze the hydrolysis of phosphodiesters in nucleic acids are collectively known as **nucleases**. There are a variety of different nucleases in all cells. Some of them are required for the synthesis or repair of DNA as we will see in Chapter 20 and others are needed for the production or degradation of cellular RNA (Chapter 21).

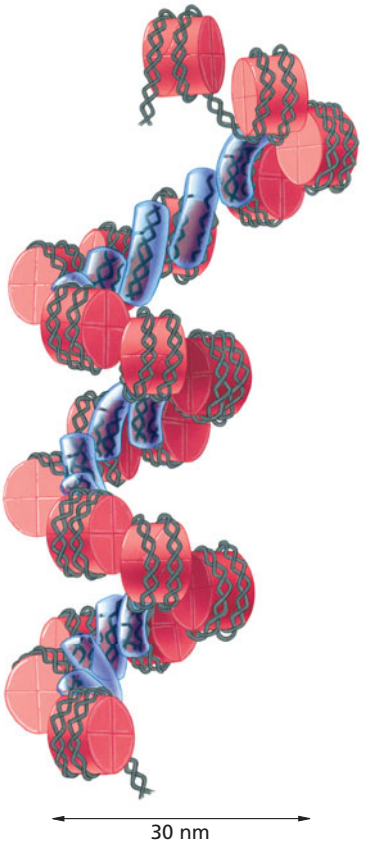
Some nucleases act on both RNA and DNA molecules but many act only on RNA and others only on DNA. The specific nucleases are called ribonucleases (RNases) and deoxyribonucleases (DNases). Nucleases can be further classified as exonucleases or endonucleases. **Exonucleases** catalyze the hydrolysis of phosphodiester linkages to release nucleotide residues from only one end of a polynucleotide chain. The most common exonucleases are the 3′ → 5′ exonucleases but there are some 5′ → 3′ exonucleases. **Endonucleases** catalyze the hydrolysis of phosphodiester linkages at various sites within a polynucleotide chain. Nucleases have a wide variety of specificities for nucleotide sequences.

Nucleases can cleave either the 3′- or the 5′-ester bond of a 3′–5′ phosphodiester linkage. One type of hydrolysis yields a 5′-monophosphate and a 3′-hydroxyl group; the other type yields a 3′-monophosphate and a 5′-hydroxyl group (see Figure 19.27). A given nuclease can catalyze one reaction or the other but not both.

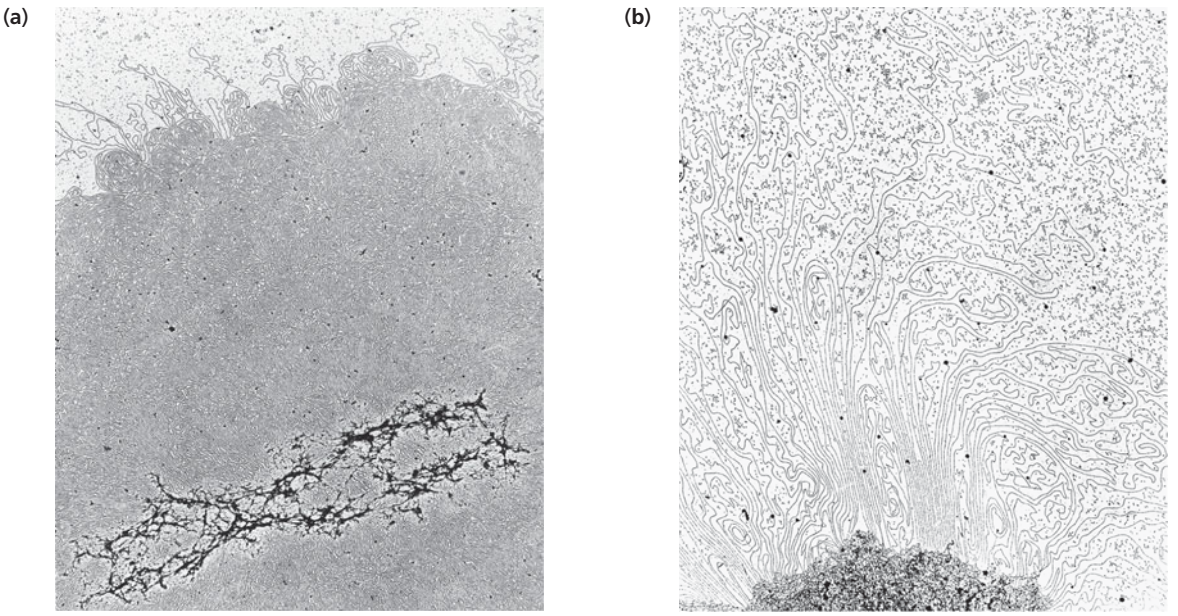
A. Alkaline Hydrolysis of RNA

The difference between ribose in RNA and 2′-deoxyribose in DNA may seem trivial but it greatly affects the properties of the nucleic acids. The 2′-hydroxyl group of ribose can form hydrogen bonds in some RNA molecules and it participates in certain chemical and enzyme-catalyzed reactions.

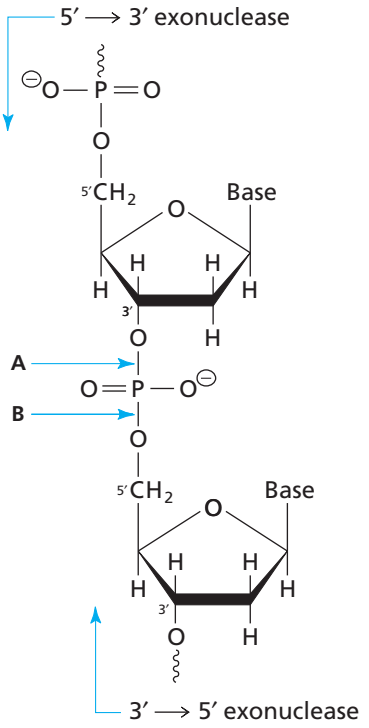
The effect of alkaline solutions on RNA and DNA illustrates the differences in chemical reactivity that result from the presence or absence of the 2′-hydroxyl group. RNA treated with 0.1 M NaOH at room temperature is degraded to a mixture of 2′- and 3′-nucleoside monophosphates within a few hours whereas DNA is stable under the same conditions. Alkaline hydrolysis of RNA (Figure 19.28) requires a 2′-hydroxyl group. In the first and second steps, hydroxide ions act only as catalysts



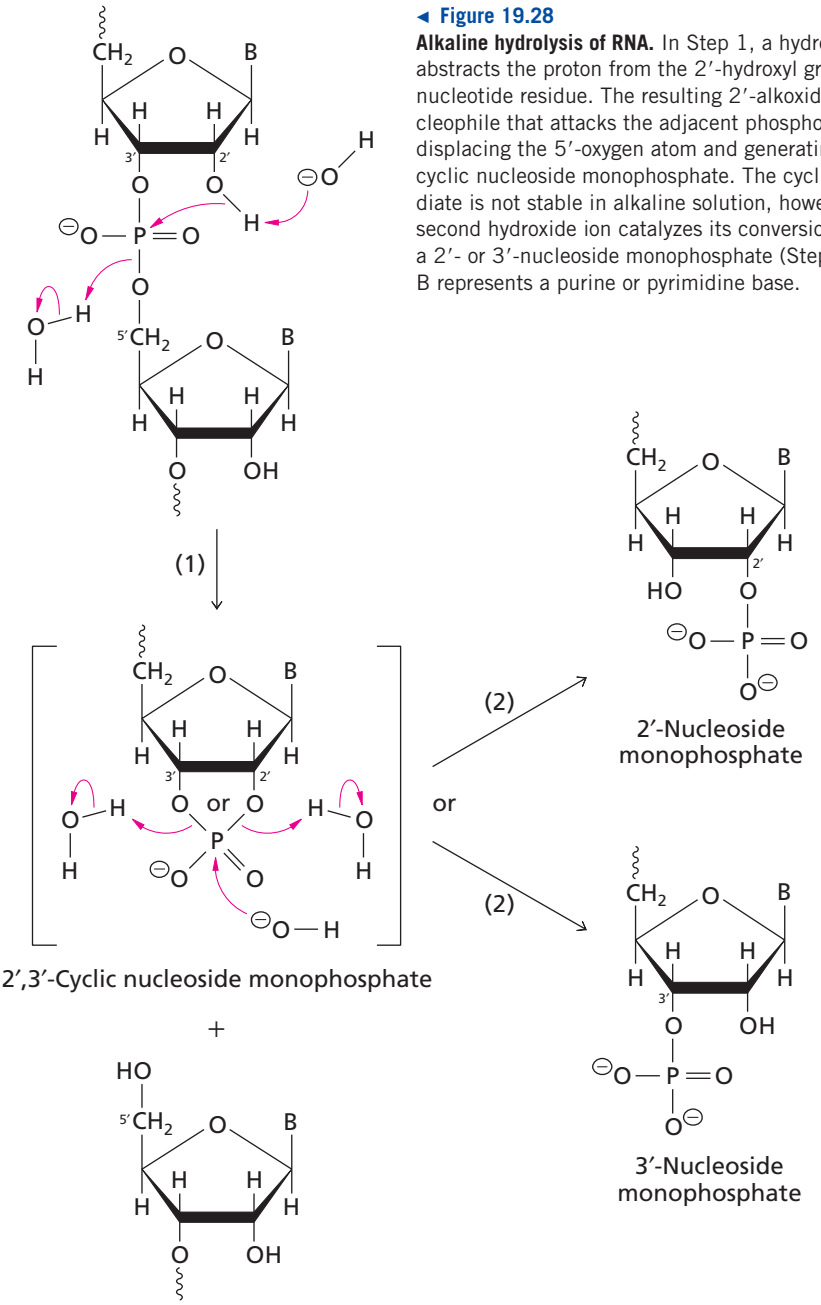
▲ **Figure 19.25**
A model of the 30 nm chromatin fiber. In this model the 30 nm fiber is shown as a solenoid, or helix, formed by individual nucleosomes. The nucleosomes associate through contacts between adjacent histone H1 molecules.



▲ **Figure 19.26**
Electron micrographs of a histone-depleted chromosome. (a) In this view, the entire protein scaffold is visible. **(b)** In this magnification of a portion of (a), individual loops attached to the protein scaffold can be seen.



▲ **Figure 19.27**
Nuclease cleavage sites. Exonucleases act on one free end of a polynucleotide and cleave the next phosphodiester linkage. Endonucleases cleave internal phosphodiester linkages. Cleavage at bond A generates a 5′-phosphate and a 3′-hydroxyl terminus. Cleavage at bond B generates a 3′-phosphate and a 5′-hydroxyl terminus. Both DNA (shown) and RNA are substrates of nucleases.



◀ **Figure 19.28**
Alkaline hydrolysis of RNA. In Step 1, a hydroxide ion abstracts the proton from the 2′-hydroxyl group of a nucleotide residue. The resulting 2′-alkoxide is a nucleophile that attacks the adjacent phosphorus atom, displacing the 5′-oxygen atom and generating a 2′,3′-cyclic nucleoside monophosphate. The cyclic intermediate is not stable in alkaline solution, however, and a second hydroxide ion catalyzes its conversion to either a 2′- or 3′-nucleoside monophosphate (Step 2). B represents a purine or pyrimidine base.

since removing a proton from water (to form the 5′-hydroxyl group in the first step or the 2′- or 3′-hydroxyl group in the second) regenerates one hydroxide ion for each hydroxide ion consumed. Note that a 2′,3′-cyclic nucleoside monophosphate intermediate forms. The polyribonucleotide chain rapidly depolymerizes as each phosphodiester linkage is cleaved. DNA is not hydrolyzed under alkaline conditions because it lacks the 2′-hydroxyl group needed to initiate intramolecular transesterification. The greater chemical stability of DNA is an important factor in its role as the primary genetic material.

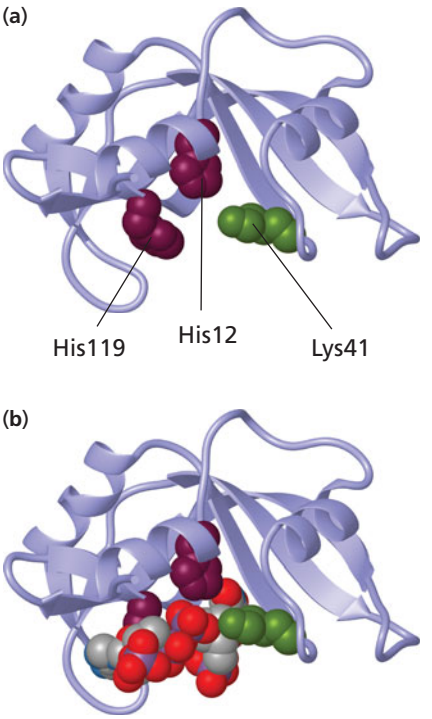
B. Hydrolysis of RNA by Ribonuclease A

Bovine pancreatic ribonuclease A (RNase A) consists of a single polypeptide chain of 124 amino acid residues cross-linked by four disulfide bridges. (This is the same enzyme that we encountered in Chapter 4 in our discussion of disulfide bond formation

and protein folding.) The enzyme has a pH optimum of about 6. RNase A catalyzes cleavage of phosphodiester linkages in RNA molecules at 5′-ester bonds. Cleavage occurs to the right of pyrimidine nucleotide residues when chains are drawn in the 5′ → 3′ direction. Thus, RNase A catalyzed hydrolysis of a strand with the sequence pApG pUpApCpGpU yields pApGpUp + ApCp + GpU.

RNase A contains three ionic amino acid residues in the active site—Lys-41, His-12, and His-119 (Figure 19.29). Many studies have led to formulation of the mechanism of catalysis shown in Figure 19.30. RNase A uses three fundamental catalytic mechanisms: proximity (in the binding and positioning of a suitable phosphodiester between the two histidine residues); acid–base catalysis (by His-119 and His-12); and transition-state stabilization (by Lys-41). As in alkaline hydrolysis of RNA, hydrolysis produces a leaving group with a 5′-hydroxyl group and a 3′-nucleoside monophosphate product. Water enters the active site on departure of the first product (P_1). Note that in the RNase A–catalyzed reaction, the phosphate atom in the transition state is pentacovalent. The pyrimidine binding pocket of the enzyme accounts for the specificity of RNase A.

Alkaline hydrolysis and the reaction catalyzed by RNase A differ in two important ways. First, alkaline hydrolysis can occur at any residue whereas enzyme-catalyzed cleavage occurs only at pyrimidine nucleotide residues. Second, hydrolysis of the cyclic intermediate is random in alkaline hydrolysis (producing mixtures of 2′- and 3′-nucleotides) but specific for RNase A–catalyzed cleavage (producing only 3′-nucleotides).



▲ **Figure 19.29**
The active site of bovine pancreatic RNase A. **(a)** The active site of the enzyme has three catalytic residues, His-12, His-119, and Lys-41, whose side chains project into the site where RNA will bind. **(b)** This figure shows RNase A bound to an artificial substrate (3′-phosphothymidine (3′-5′)-pyrophosphate adenosine 3′-phosphate) that mimics RNA. [PDB 1U1B]

C. Restriction Endonucleases

Restriction endonucleases are an important subclass of endonucleases that act on DNA. The term **restriction endonuclease** is derived from the observation that certain bacteria can block bacteriophage (virus) infections by specifically destroying the incoming bacteriophage DNA. Such bacteria *restrict* the expression of foreign DNA.

Many species of bacteria synthesize restriction endonucleases that bind to and cleave foreign DNA. These endonucleases recognize specific DNA sequences and they cut both strands of DNA at the binding site producing large fragments that are rapidly degraded by exonucleases. The bacteriophage DNA is cleaved and degraded before the genes can be expressed.

The host cell has to protect its own DNA from cleavage by restriction endonucleases. This is accomplished by covalent modification of the bases that make up the potential restriction endonuclease binding site. The most common covalent modification is specific methylation of adenine or cytosine residues within the recognition sequence (Section 18.7). The presence of methylated bases at the potential binding site inhibits cleavage of the host DNA by the restriction endonuclease. Methylation is catalyzed by a specific methylase that binds to the same sequence of DNA recognized by the restriction endonuclease. Thus, cells that contain a restriction endonuclease also contain a methylase with the same specificity.

Normally, all DNA of the host cell is specifically methylated and therefore protected from cleavage. Any unmethylated DNA that enters the cell is cleaved by restriction endonucleases. Following DNA replication, each site in the host DNA is hemimethylated—bases on only one strand are methylated. Hemimethylated sites are high affinity substrates for the methylase but are not recognized by the restriction endonuclease. Thus, hemimethylated sites are rapidly converted to fully methylated sites in the host DNA (Figure 19.31).

Most restriction endonucleases (also called restriction enzymes) can be classified as either type I or type II. Type I restriction endonucleases catalyze both the methylation of host DNA and the cleavage of unmethylated DNA at a specific recognition sequence. Type II restriction endonucleases are simpler in that they can only cleave double-stranded DNA at or near an unmethylated recognition sequence—they do not possess a methylase activity. Separate restriction methylases catalyze methylation of host DNA at the same recognition sequences. The source of the methyl group in these reactions is *S*-adenosylmethionine.

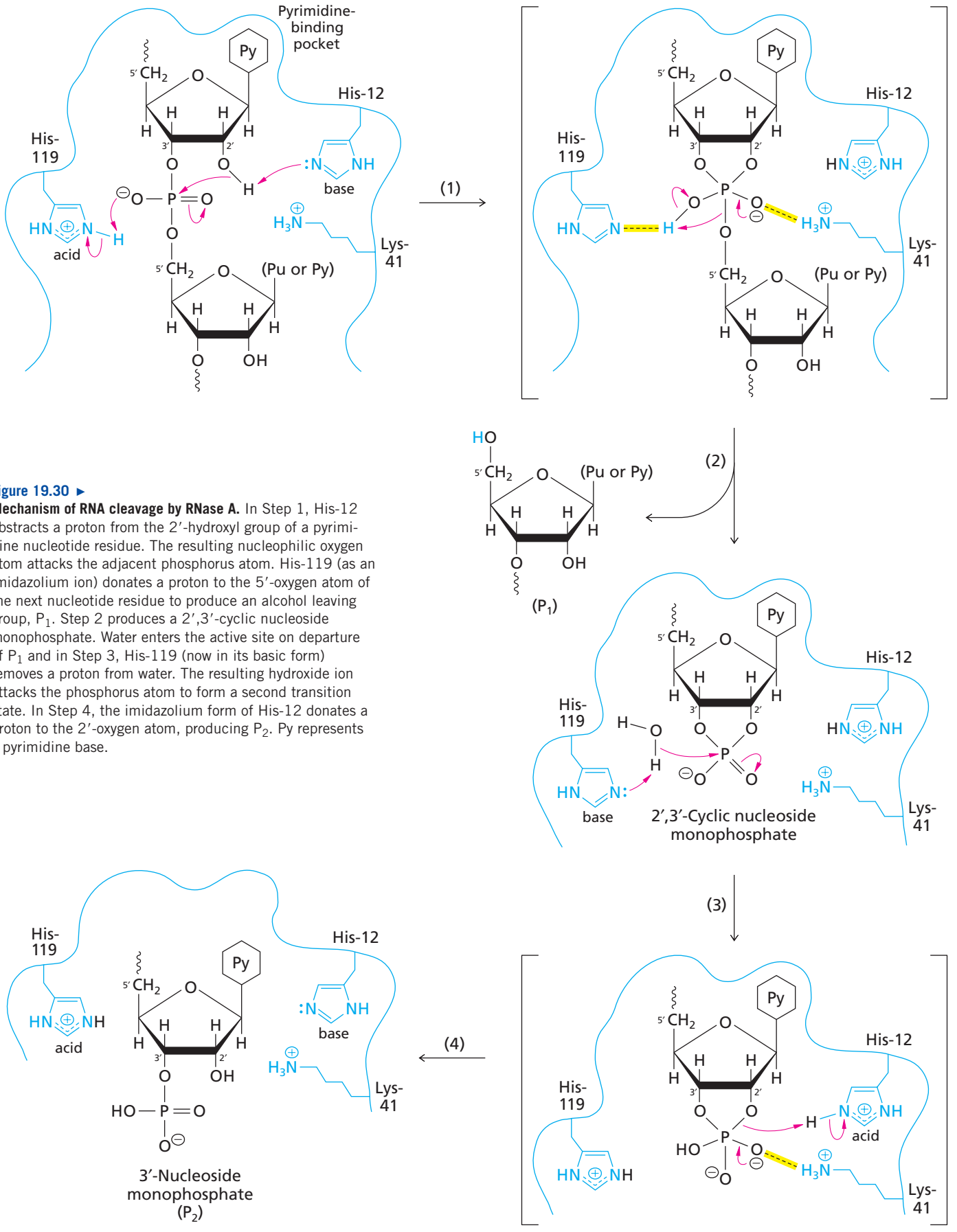


Figure 19.30 ▶
Mechanism of RNA cleavage by RNase A. In Step 1, His-12 abstracts a proton from the 2′-hydroxyl group of a pyrimidine nucleotide residue. The resulting nucleophilic oxygen atom attacks the adjacent phosphorus atom. His-119 (as an imidazolium ion) donates a proton to the 5′-oxygen atom of the next nucleotide residue to produce an alcohol leaving group, P_1. Step 2 produces a 2′,3′-cyclic nucleoside monophosphate. Water enters the active site on departure of P_1 and in Step 3, His-119 (now in its basic form) removes a proton from water. The resulting hydroxide ion attacks the phosphorus atom to form a second transition state. In Step 4, the imidazolium form of His-12 donates a proton to the 2′-oxygen atom, producing P_2. Py represents a pyrimidine base.

Table 19.5 Specificities of some common restriction endonucleases

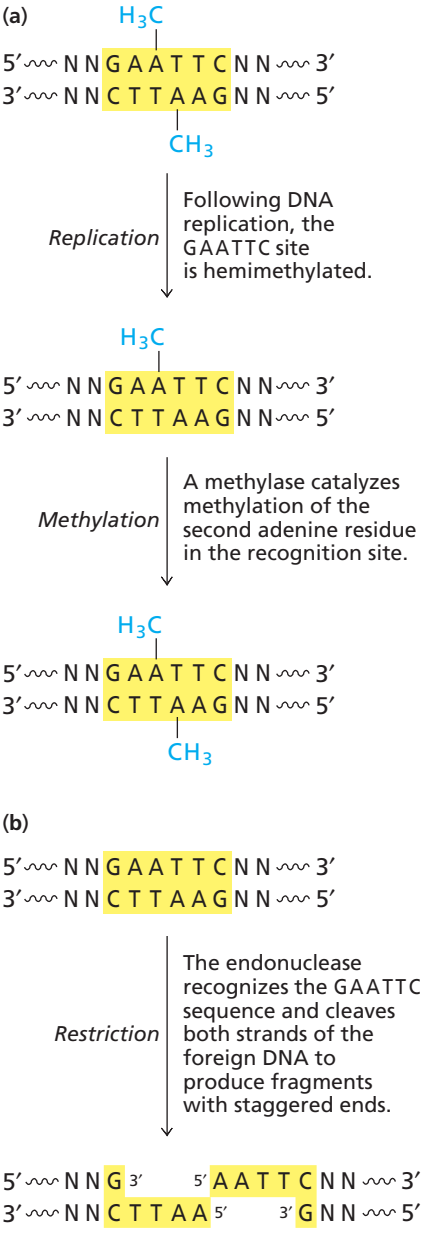
Source	Enzyme[a]	Recognition sequence[b]
Acetobacter pasteurianus	*Apa*I	GGGCC↓C
Bacillus amyloliquefaciens H	*Bam*HI	G↓GATCC
Eschericia coli RY13	*Eco*RI	G↓AA*TTC
Eschericia coli R245	*Eco*RII	↓CC*TGG
Haemophilus aegyptius	*Hae*III	GG↓CC
Haemophilus influenzae R_d	*Hind*III	A*↓AGCTT
Haemophilus parainfluenzae	*Hpa*II	C↓CGG
Klebsiella pneumoniae	*Kpn*I	GGTAC↓C
Nocardia otitidis-caviarum	*Not*I	GC↓GGCCGC
Providencia stuartii 164	*Pst*I	CTGCA↓G
Serratia marcescens S_b	*Sma*I	CCC↓GGG
Xanthomonas badrii	*Xba*I	T↓CTAGA
Xanthomonas holcicola	*Xho*I	C↓TCGAG

[a]The names of restriction endonucleases are abbreviations of the names of the organisms that produce them. Some abbreviated names are followed by a letter denoting the strain. Roman numerals indicate the order of discovery of the enzyme in that strain.
[b]Recognition sequences are written 5′ to 3′. Only one strand is represented. The arrows indicate cleavage sites. Asterisks represent known positions where bases can be methylated.

Hundreds of type I and type II restriction endonucleases have been characterized. The specificities of a few representative enzymes are listed in Table 19.5. In nearly all cases, the recognition sites have a twofold axis of symmetry; that is, the 5′ ⟶ 3′ sequence of residues is the same in both strands of the DNA molecule. Consequently, the paired sequences "read" the same in either direction—such sequences are known as palindromes. (Palindromes in English include BIB, DEED, RADAR, and even MADAM I'M ADAM, provided we ignore punctuation and spacing.)

*Eco*RI was one of the first restriction endonucleases to be discovered. It is present in many strains of *E. coli*. As shown in Table 19.5 and Figure 19.31, *Eco*RI has a palindromic recognition sequence of 6 bp (the 5′ ⟶ 3′ sequence is GAATTC on each strand). *Eco*RI is a homodimer. It possesses a twofold axis of symmetry like its substrate (see next section). In *E. coli*, the companion methylase to *Eco*RI converts the second adenine within the recognition sequence to N^6-methyladenine. Any double-stranded DNA molecule with an unmethylated GAATTC sequence is a substrate for *Eco*RI. The endonuclease catalyzes hydrolysis of the phosphodiesters that link G to A in each strand, thus cleaving the DNA.

Some restriction endonucleases (including *Eco*RI, *Bam*HI, and *Hin*dIII) catalyze staggered cleavage, producing DNA fragments with single-stranded extensions (Table 19.5 and Figure 19.31). These single-stranded regions are called sticky ends because they are complementary and can thus re-form a double-stranded structure. Other enzymes, such as *Hae*III and *Sma*I, produce blunt ends with no single-stranded extensions.

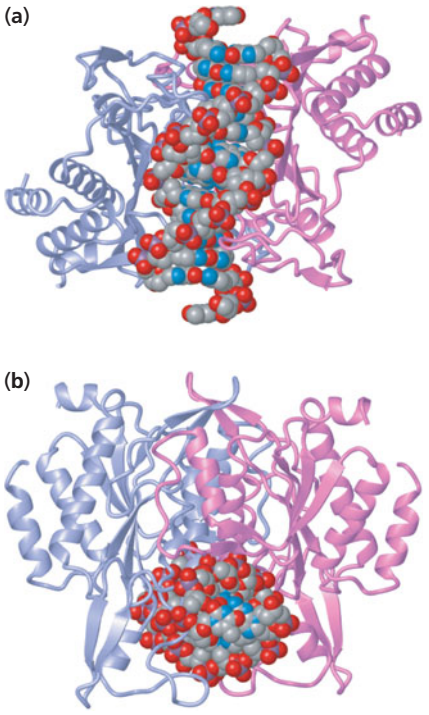


▲ **Figure 19.31**
Methylation and restriction at the *Eco*RI *site.* **(a)** Methylation of adenine residues at the recognition site. **(b)** Cleavage of unmethylated DNA to produce sticky ends.

D. *Eco*RI Binds Tightly to DNA

Restriction endonucleases must bind tightly to DNA in order to recognize a specific sequence and cleave at a specific site. The structure of *Eco*RI bound to DNA has been determined by X-ray crystallography. As shown in Figure 19.32, each half of the *Eco*RI homodimer binds to one side of the DNA molecule so that the DNA molecule is almost surrounded. The enzyme recognizes the specific nucleotide sequence by contacting base pairs in the major groove. The minor groove (in the middle of the structure shown in Figure 19.32) is exposed to the aqueous environment.

Several basic amino acid residues line the cleft that is formed by the two *Eco*RI monomers. The side chains of these residues interact electrostatically with the



▲ **Figure 19.32**
***Eco*RI bound to DNA.** *Eco*RI is composed of two identical subunits (purple and blue). The enzyme is bound to a fragment of DNA with the sequence CGCGAATTCGCG (recognition sequence underlined). **(a)** Side view. **(b)** Top view.

sugar–phosphate backbones of DNA. In addition, two arginine residues (Arg-145 and Arg-200) and one glutamate residue (Glu-144) in each *Eco*RI monomer form hydrogen bonds with base pairs in the recognition sequence thus ensuring specific binding. Other nonspecific interactions with the backbones further stabilize the complex.

*Eco*RI is typical of proteins that recognize and bind to a *specific* DNA sequence. The DNA retains its B conformation although in some cases the helix is slightly bent. Recognition of a specific nucleotide sequence depends on interactions between the protein and the functional groups on the bases that are exposed in the grooves. In contrast, histones are examples of proteins that bind *nonspecifically* to nucleic acids. Binding of such proteins depends largely on weak interactions between the protein and the sugar–phosphate backbones and not on direct contact with the bases. All proteins that bind to specific DNA sequences will also bind non-specifically to DNA with lower affinity (Sections 21.3, 21.7A).

19.7 Uses of Restriction Endonucleases

Restriction endonucleases were discovered more than 40 years ago earning Nobel Prizes in 1978 for Werner Arber, Daniel Nathans, and Hamilton Smith "for the discovery of restriction enzymes and their application to problems of molecular genetics." The first purified enzymes rapidly became important tools used to manipulate DNA in the laboratory.

A. Restriction Maps

One of the first uses of restriction enzymes was in developing restriction maps of DNA, that is, diagrams of DNA molecules that show specific sites of cleavage. Such maps are useful for identifying fragments of DNA that contain specific genes.

An example of a restriction map of bacteriophage λ DNA is shown in Figure 19.33. The DNA of bacteriophage λ is a linear, double-stranded molecule approximately 48,400 bp (48 kb) long. By treating this DNA with various restriction enzymes and measuring the sizes of the resulting fragments, it is possible to develop a map of the cleavage sites. An example of such restriction digests is shown in Figure 19.34. The information from many restriction digests is combined to produce a complete and accurate map.

B. DNA Fingerprints

The technology required for mapping restriction endonuclease cleavage sites was developed in the 1970s. It soon became apparent that the procedure could be used to identify the sites of mutations, or variations, in the genome of a population. For example, different strains of bacteriophage λ have slightly different restriction maps because their DNA sequences are not identical. One strain may have the sequence GGGCCC near the left-hand end of its DNA and it is cleaved by *Apa*I, producing the two fragments shown in Figure 19.34. Another strain may have the sequence GGACCC at the same site. Since this sequence is not a cleavage site for *Apa*I, the restriction map of this strain differs from that shown in Figure 19.33.

Variations in DNA sequence can be used to identify individuals in a large heterogeneous population. In humans, for example, regions of the genome that are highly variable give restriction fragments that are as unique as fingerprints. Such DNA fingerprints can be used in paternity disputes or criminal investigations to identify or exonerate suspects.

An example of the use of DNA fingerprinting in a rape case is shown in Figure 19.35. DNAs isolated from the victim, from the evidence (semen), and from two suspects are

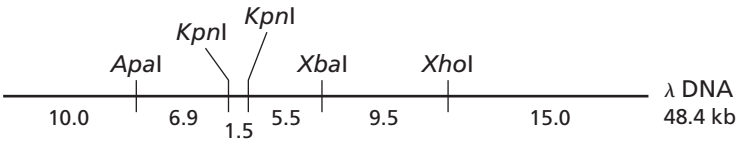
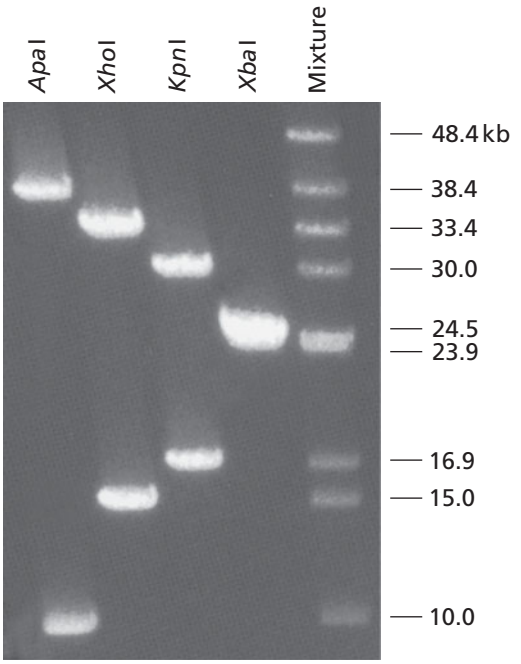
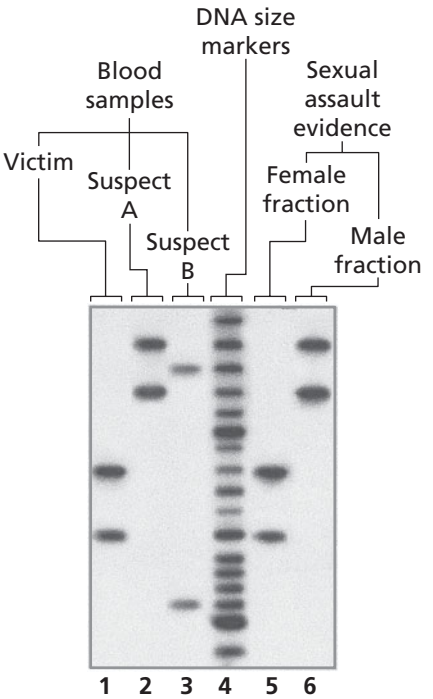


Figure 19.33 ►
Restriction map of bacteriophage λ showing the sites of cleavage by some restriction enzymes. There is a single site for the enzyme *Apa*I, for example. Digestion of phage λ DNA with this enzyme yields two fragments of 10.0 and 38.4 kb, as shown in the first lane of Figure 19.34.



▲ **Figure 19.34**
Digestion of bacteriophage λ DNA by four restriction endonucleases. A solution of DNA is treated with an enzyme and then electrophoresed on an agarose gel, which separates fragments according to size. The smallest fragments move fastest and are found at the bottom of the gel. (A fragment of 1.5 kb is not visible in this figure.) The restriction enzyme for each digest is indicated at the top of the lane. The lane at the right contains intact phage λ DNA and a mixture of fragments from the four digests. In the *Xba*I digest, two fragments of 23.9 and 24.5 kb are not well resolved.



▲ **Figure 19.35**
DNA fingerprinting.

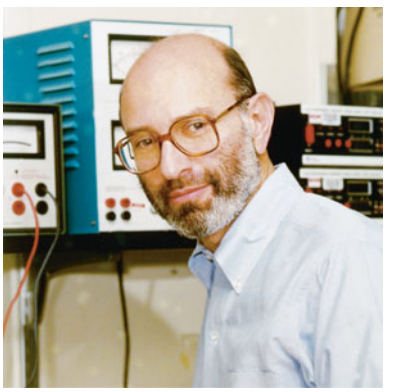
digested with a restriction endonuclease. The fragments are separated on an agarose gel as described in Figure 19.34. This DNA is then transferred (blotted) to a membrane of nylon. The bound DNA is denatured and exposed to small fragments of radioactively labeled DNA from a variable region of the human genome. The labeled DNA probe hybridizes specifically to the restriction fragments on the nylon membrane that are derived from this region. The labeled fragments are identified by autoradiography.

The technique identifies suspect A as the rapist. In actual criminal investigations, a number of different probes are used in combination with different restriction digests in order to ensure that the pattern detected is unique. Modern techniques are powerful and accurate enough to conclusively rule out some suspects and convict others. When combined with polymerase chain reaction (PCR) amplification of DNA (Chapter 22), a fingerprint can be obtained from a hair follicle or a tiny speck of blood.

C. Recombinant DNA

The discovery of restriction endonucleases soon led to the creation of recombinant DNA molecules by joining, or recombining, different fragments of DNA produced by the enzymes. A common experiment involves excising a DNA fragment containing a target gene of interest and inserting it into a cloning vector. Cloning vectors can be plasmids, bacteriophage, viruses, or even small artificial chromosomes. Most vectors contain sequences that allow them to be replicated autonomously within a compatible host cell.

All cloning vectors have in common at least one unique cloning site, a sequence that can be cut by a restriction endonuclease to allow site-specific insertion of foreign DNA. The most useful vectors have several restriction sites grouped together in a multiple cloning site called a *polylinker*.





► Stanley N. Cohen (1935–) (top) and Herbert Boyer (1936–) (bottom), who constructed the first recombinant DNA using bacterial DNA and plasmids.

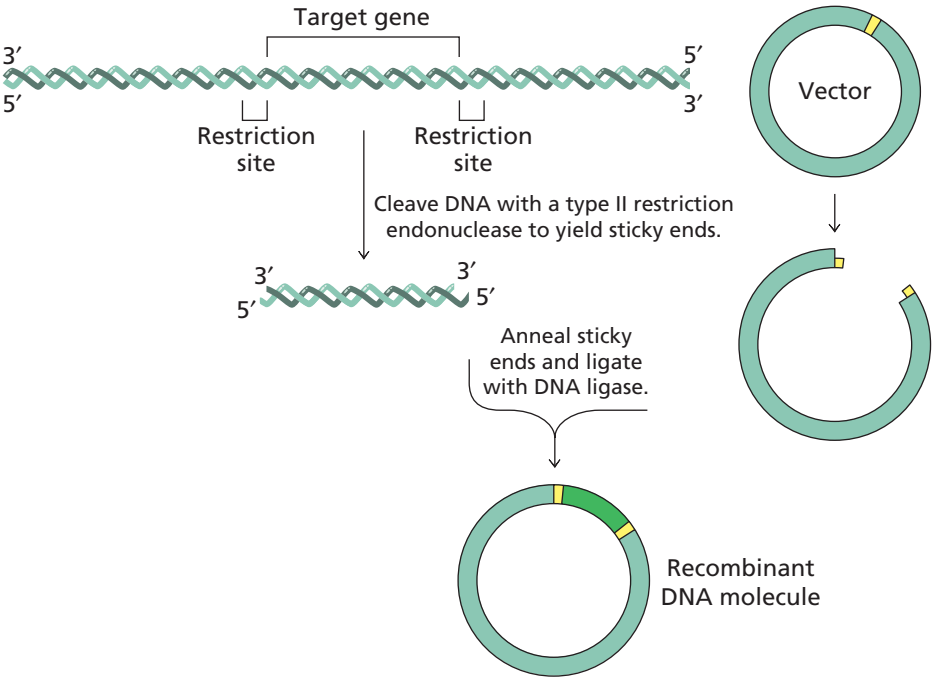


Figure 19.36 ►
Use of restriction enzymes to generate recombinant DNA. The vector DNA and the target DNA are cleaved by restriction endonucleases to generate ends that can be joined together. In cases where sticky ends are produced, the two molecules join by annealing (base pairing) of the complementary ends. The molecules are then covalently attached to one another in a reaction catalyzed by DNA ligase.

Fragments of DNA to be inserted into a vector can be generated by a variety of means. For example, they can be produced by the mechanical shearing of long DNA molecules or by digesting DNA with type II restriction endonucleases. Unlike shearing, which cleaves DNA randomly, restriction enzymes cleave DNA at specific sequences. For cloning purposes, this specificity offers extraordinary advantages.

The most useful restriction endonucleases produce fragments with single-stranded extensions at their 3′ or 5′ ends. These sticky ends can transiently form base pairs to complementary sticky ends on vector DNA and can be covalently joined to the vector in a reaction catalyzed by DNA ligase (described in Section 20.3C). Thus, the simplest kinds of recombinant DNA are those constructed by digesting both the vector and the target DNA with the same enzyme because the resulting fragments can be joined directly by ligation (Figure 19.36).

Summary

1. Nucleic acids are polymers of nucleotides that are phosphate esters of nucleosides. The amino and lactam tautomers of the bases form hydrogen bonds in nucleic acids.
2. DNA contains two antiparallel strands of nucleotide residues joined by 3′–5′ phosphodiester linkages. A and G in one strand pair with T and C, respectively, in the other strand.
3. The double-helical structure of DNA is stabilized by hydrogen bonding, hydrophobic effects, stacking interactions, and charge–charge interactions. G/C-rich DNA is more difficult to denature than A/T-rich DNA because the stacking interactions of G/C base pairs are greater than those of A/T base pairs.
4. The most common conformation of DNA is called B-DNA; alternative conformations include A-DNA and Z-DNA.
5. Overwinding or underwinding the DNA helix can produce supercoils that restore the B conformation. Negatively supercoiled DNA exists in equilibrium with DNA that has locally unwound regions.
6. The four major classes of RNA are ribosomal RNA, transfer RNA, messenger RNA, and small RNA. RNA molecules are single-stranded and have extensive secondary structure.
7. Eukaryotic DNA molecules are packaged with histones to form nucleosomes. Further condensation and attachment to the scaffold of a chromosome achieves an overall 8000-fold reduction in the length of the DNA molecule in metaphase chromosomes.
8. The phosphodiester backbones of nucleic acids can be hydrolyzed by the actions of nucleases. Alkaline hydrolysis and RNase A-catalyzed hydrolysis of RNA proceed via a 2′,3′-cyclic nucleoside monophosphate intermediate.
9. Restriction endonucleases catalyze hydrolysis of DNA at specific palindromic nucleotide sequences. Specific methylases protect restriction sites from cleavage.
10. Restriction enzymes are useful for constructing restriction maps of DNA, for DNA fingerprint analysis, and for constructing recombinant DNA molecules.

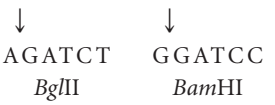
Problems

1. Compare hydrogen bonding in the α helix of proteins to hydrogen bonding in the double helix of DNA. Include in the answer the role of hydrogen bonding in stabilizing these two structures.
2. A stretch of double-stranded DNA contains 1000 bp and its base composition is 58% (G + C). How many thymine residues are in this region of DNA?
3. (a) Do the two complementary strands of a segment of DNA have the same base composition?
   (b) Does (A + G) equal (C + T)?
4. If one strand of DNA has the sequence

   ATCGCGTAACATGGATTCGG

   write the sequence of the complementary strand using the standard convention.
5. Poly A forms a single-stranded helix. What forces stabilize this structure?
6. The imino tautomer of adenine occurs infrequently in DNA but when it does it can pair with cytosine instead of thymine. Such mispairing can lead to a mutation. Draw the adenine imino tautomer/cytosine base pair.
7. Single-stranded poly-dA can hybridize to single-stranded poly-dT to form Watson–Crick base-paired double-stranded DNA. Under appropriate conditions a second strand of poly-dT can bind in the major groove and form a triple-stranded DNA helix with hydrogen bonds between the thymine and the N7 and amino group in adenine. What would a plot of absorbance at 260 nm vs. temperature look like for this unusual triple-stranded DNA?
8. Write the sequence of the RNA shown in Figure 19.21. Is it a palindrome?
9. Consider a processive exonuclease that binds exclusively to double-stranded DNA and degrades one strand in the 5′ → 3′ direction. In a reaction where the substrate is a 1 kb fragment of linear DNA, what will be the predominant products after the digestion has gone to completion?
10. The average molecular weight of a base pair in double-stranded DNA is approximately 650 kDa. Using the data from Table 19.4, calculate the mass ratio of protein to DNA in a typical 30 nm chromatin fiber.
11. The human haploid genome contains 3.2×10^9 base pairs. How many nucleosomes did you inherit from your mother?
12. A DNA molecule with the sequence pdApdGpdTpdC can be cleaved by exonucleases. List the products of a single reaction catalyzed by the following enzymes:
   (a) a 3′ → 5′ exonuclease that cleaves the 3′ ester bond of a phosphodiester linkage
   (b) a 5′ → 3′ exonuclease that cleaves the 5′ ester bond of a phosphodiester linkage
   (c) a 5′ → 3′ exonuclease that cleaves the 3′ ester bond of a phosphodiester linkage
13. A non-sequence-specific endonuclease purified from *Aspergillus oryzae* digests single-stranded DNA. Predict the effect of adding this enzyme to a preparation of negatively supercoiled plasmid DNA.
14. One of the proteins in rattlesnake venom is an enzyme named phosphodiesterase. Could polynucleotides be a substrate for this enzyme? Why or why not?
15. RNase T1 cleaves RNA after G residues to leave a 3′ phosphate group. Predict the cleavage products of this substrate:

   pppApCpUpCpApUpApGpCpUpApUpGpApGpU
16. How could bacteriophages escape the effects of bacterial restriction endonucleases?
17. The free-living soil nematode *C. elegans* was the first metazoan to have its entire 100 Mb genome sequenced. Overall, the worm genome is 36% (G + C) and 64% (A + T). The restriction endonuclease *Hin*dIII recognizes and cuts the hexameric palindromic sequence AAGCTT to generate sticky ends. (a) Approximately how many *Hin*dIII sites would you expect to find in the *C. elegans* genome? (b) If the worm genome was actually 25% G and 25% A, approximately how many *Hin*dIII sites would you expect to find?
18. The recognition sites for the restriction endonucleases *Bgl*II and *Bam*HI are shown below. Why is it possible to construct recombinant DNA molecules by combining target DNA cut with *Bgl*II and a vector cut with *Bam*HI?



19. One of the *E. coli* host strains commonly used in recombinant DNA technology carries defective genes for several restriction endonucleases. Why is such a strain useful?

Selected Readings

Historical Perspective

Clayton, J., and Denis. C. (eds.) (2003). *50 Years of DNA*. (New York: Nature/Pallgrave/Macmillan).

Judson, H. F. (1996). *The Eighth Day of Creation: Makers of the Revolution in Biology*, expanded ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Maddox, B. (2002). *Rosalind Franklin: The Dark Lady of DNA* (New York: Perennial/HarperCollins).

Watson, J. D., and Berry, A. (2003). *DNA: The Secret of Life* (New York: Alfred A. Knopf).

Watson, J. D. (1968). *The Double Helix* (New York: Atheneum).

Polynucleotide Structure and Properties

Berger, J. M., and Wang, J. C. (1996). Recent developments in DNA topoisomerase II structure and mechanism. *Curr. Opin. Struct. Biol.* 6:84–90.

Ferré-D'Amaré, A. R., and Doudna, J. A. (1999). RNA FOLDS: insights from recent crystal structures. *Annu. Rev. Biophys. Biomol. Struct.* 28:57–73.

Herbert, A., and Rich, A. (1996). The biology of left-handed Z-DNA. *J. Biol. Chem.* 271:11595–11598.

Hunter, C. A. (1996). Sequence-dependent DNA structure. *BioEssays* 18:157–162.

Ke, C., Humeniuk, M., S-Gracz, H., and Marszalek, P. E. (2007). Direct measurements of base stacking interactions in DNA by single-molecule atomic-force spectroscopy. *Phys. Rev. Lett.* 99: 018302.

Kool, E. T., Morales, J. C., and Guckian, K. M. (2000). Mimicking the structure and function of DNA: insights into DNA stability and replication. *Angew. Chem. Int. Ed.* 39:990–1009.

Packer, M. J., and Hunter, C. A. (1998). Sequence-dependent DNA structure: the role of the sugar-phosphate backbone. *J. Mol. Biol.* 280:407–420.

Saenger, W. (1984). *Principles of Nucleic Acid Structure* (New York: Springer-Verlag).

Sharma, A., and Mondragón, A. (1995). DNA topoisomerases. *Curr. Biol.* 5:39–47.

Wang, J. C. (2009). A journey in the world of DNA rings and beyond. *Annu. Rev. Biochem.* 78:31–54.

Chromatin

Bendich, A. J., and Drlica, K. (2000). Prokaryotic and eukaryotic chromosome: what's the difference? *BioEssays* 22:481–486.

Burlingame, R. W., Love, W. E., Wang, B.-C., Hamlin, R., Xuong, N.-H., and Moudrianakis, E. N. (1985). Crystallographic structure of the octameric histone core of the nucleosome at a resolution of 3.3 Å. *Science* 228:546–553.

Grigoryev, S. A., Arya, G., Correll, S., Woodcock, C. L., and Schlick, T. (2009). Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc. Natl. Acad. Sci. (USA)* 106:13317–13322.

Kornberg, R. D. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryotic chromosome. *Cell* 98:285–294.

Ramakrishnan, V. (1997). Histone structure and the organization of the nucleosome. *Annu. Rev. Biophys. Biomol. Struct.* 26:83–112.

Richmond, T. J., Finch, J. T., Rushton, D., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature* 311:532–537.

Van Holde, K., and Zlatanova, J. (1999). The nucleosome core particle: does it have structural and functional relevance? *BioEssays* 21:776–780.

Workman, J. L., and Kingston, R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* 67:545–579.

Restriction Endonucleases

Kovall, R. A., and Mathews, B. W. (1999). Type II restriction endonucleases: structural, functional and evolutionary relationships. *Curr. Opin. Chem. Biol.* 3:587–583.

McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H., Grable, J., and Rosenberg, J. M. (1986). Structure of the DNA-*Eco*RI endonuclease recognition complex at 3 Å resolution. *Science* 234:1526–1541.

Ne, M. (2000). Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Rev.* 64:412–434.





DNA Replication, Repair, and Recombination

The transfer of genetic information from one generation to the next has puzzled biologists since the time of Aristotle. Today, almost 2500 years later, we can explain why "like begets like." Since genetic information is carried in DNA, it follows that the transfer of information from a parental cell to two daughter cells requires exact duplication of DNA, a process known as DNA replication.

The DNA structure proposed by Watson and Crick in 1953 immediately suggested a method of replication. The nucleotide sequence of one strand automatically specifies the sequence of the other since the two strands of double-helical DNA are complementary. Watson and Crick proposed that the two strands of the helix unwind during DNA replication and that each strand of DNA acts as a template for the synthesis of a complementary strand. In this way, DNA replication produces two double-stranded daughter molecules, each containing one parental strand and one newly synthesized strand. This mode of replication is termed semiconservative because one strand of the parental DNA is conserved in each daughter molecule (Figure 20.1, on the next page).

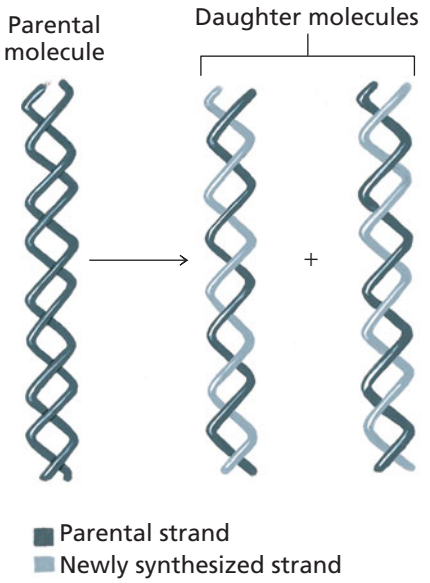
In a series of elegant experiments, Matthew Meselson and Franklin W. Stahl showed in 1958 that DNA was indeed replicated semiconservatively as predicted by Watson and Crick. About the same time, reports of the purification and properties of some of the enzymes involved in replication began to appear. The first DNA polymerase was purified in 1958 by Arthur Kornberg, who was awarded the Nobel Prize for this achievement. More recently, biochemists have isolated and characterized enzymes that catalyze all the steps in DNA replication and have identified the genes that encode these proteins. The actual mechanism of replication is much more complex—and more interesting—than the simple scheme shown in Figure 20.1.

Establishing the steps of the replication mechanism required a combination of both biochemical and genetic analysis. Much of what we know about DNA replication

The structure of DNA proposed by Watson and Crick brought forth a number of proposals as to how such a molecule might replicate. These proposals make specific predictions concerning the distribution of parental atoms among progeny molecules. The results presented here give a detailed answer to the question of this distribution and simultaneously direct our attention to other problems whose solution must be the next step in progress toward a complete understanding of the molecular basis of DNA duplication.

—Matthew Meselson and Franklin W. Stahl (1958)

Top: Holliday junction, an intermediate formed during recombination between two double-stranded DNA molecules.



◂ **Figure 20.1**
Semiconservative DNA replication. Each strand of DNA acts as a template for synthesis of a new strand. Each daughter molecule of DNA contains one parental strand and one newly synthesized strand.

has come from studies of the enzymes in *Escherichia coli* and its bacteriophages. The results of these studies have shown how large numbers of polypeptides assemble into complexes that can carry out a complicated series of reactions. The DNA replication complex is like a machine, or factory, whose parts are made of protein. Some of the component polypeptides are partially active in isolation but others are active only in association with the complete protein machine.

There are three distinct stages in DNA replication. (1) Initiation begins with the correct assembly of the replication proteins at the site where DNA replication is to start. (2) During the elongation stage, DNA is replicated semiconservatively as the complex catalyzes the incorporation of nucleotides into the growing DNA strands. (3) Finally, when replication terminates, the protein machine is disassembled and the daughter molecules separate so that they can segregate into their new cells.

Protein machines that carry out a series of biochemical reactions are not confined to the process of DNA replication but also occur in fatty acid synthesis (Section 16.1), transcription (Chapter 21), and translation (Chapter 22). All four of these processes include initiation, elongation, and termination steps. Furthermore, there is increasing evidence that other processes of cellular metabolism are also carried out by complexes of weakly associated enzymes and other macromolecules.

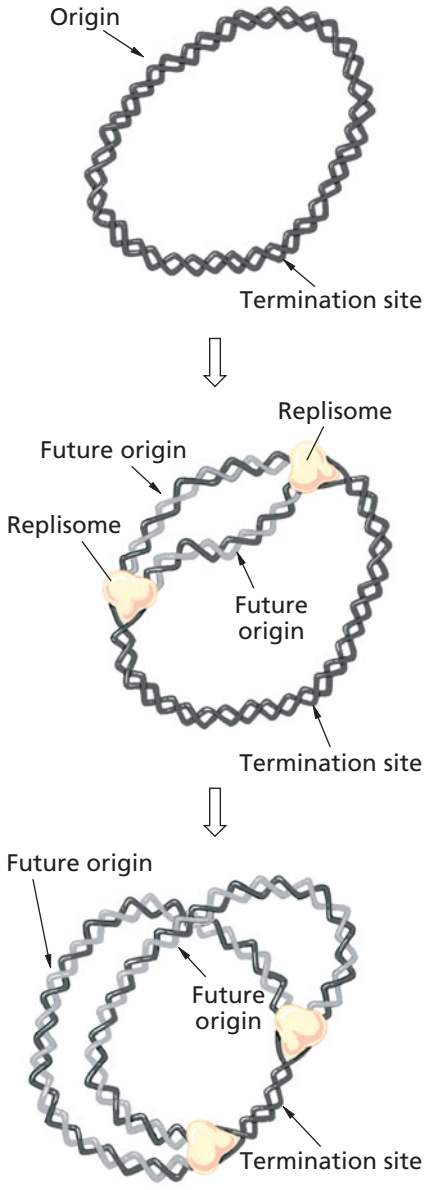
The maintenance of genetic information from generation to generation requires that DNA replication be both rapid (because the entire complement of DNA must be replicated before each cell division) and accurate. All cells have enzymes that correct replication errors and repair damaged DNA. Furthermore, all cells can shuffle pieces of DNA in a process known as genetic recombination. Both repair and recombination use many of the same enzymes and proteins that are required for DNA replication.

The overall strategy of DNA replication, repair, and recombination in prokaryotes and eukaryotes appears to be conserved, although specific enzymes vary among organisms. Just as two different makes of automobile are similar even though individual parts cannot be substituted for one another, so too are the mechanisms of DNA replication, repair, and recombination similar in all organisms, even though the individual enzymes may differ. We are going to focus on the biochemistry of these three processes in *E. coli* because of its many well-characterized enzymes.

20.1 DNA Replication Is Bidirectional

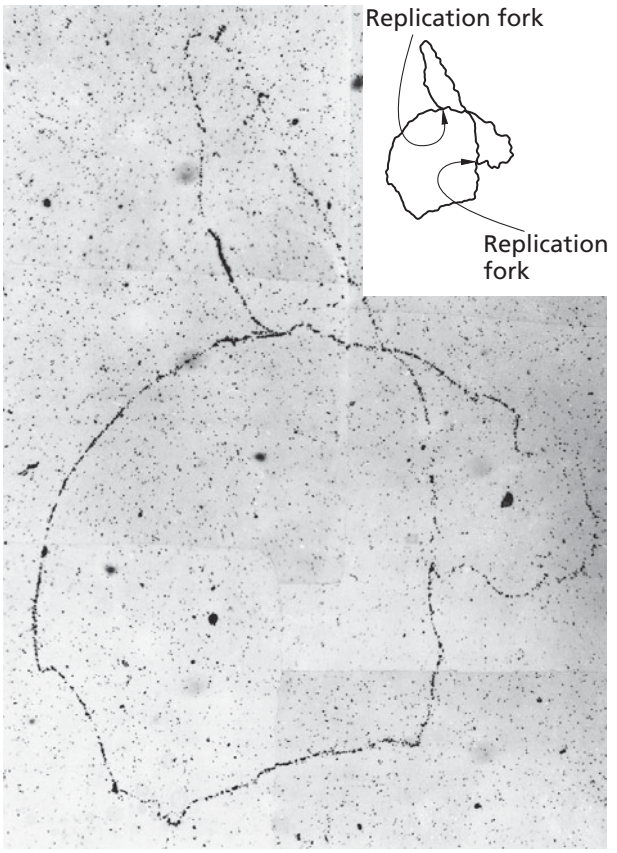
The *E. coli* chromosome is a large, circular, double-stranded DNA molecule of 4.6×10^3 kilobase pairs (kb). Replication of this chromosome begins at a unique site called the origin of replication and proceeds bidirectionally until the two replication complexes meet at the termination site, where replication stops (Figure 20.2). The protein machine that carries out the polymerization reaction is called a replisome. It contains a number of different proteins that are required for rapid and accurate DNA replication. One replisome is located at each of the two replication forks, the points where the parental DNA is unwound. Figure 20.3 shows an autoradiograph of a replicating *E. coli* chromosome.

As parental DNA is unwound at a replication fork, each strand is used as a template for the synthesis of a new strand. The rate of movement of a replication fork in *E. coli* is approximately 1000 base pairs (bp) per second. In other words, each of the two new strands is extended at the rate of 1000 nucleotides per second. Since there are two replication forks moving at this rate, the entire *E. coli* chromosome can be duplicated in about 38 minutes.



◂ **Figure 20.2**
Bidirectional DNA replication in *Escherichia coli*. Semiconservative DNA replication begins at a unique origin and proceeds in both directions. The synthesis of new strands of DNA (light gray) occurs at the two replication forks where the replisomes are located. The two double-stranded DNA molecules separate when the replication forks meet at the termination site. Note that each daughter molecule consists of one parental strand and one newly synthesized strand.

Eukaryotic chromosomes are linear, double-stranded DNA molecules that are usually much larger than the chromosomes of bacteria. The large chromosomes of the fruit fly *Drosophila melanogaster* for example, are about 5.0×10^4 kb in size or 10 times larger than the *E. coli* chromosome. Replication in eukaryotes is also bidirectional but whereas the *E. coli* chromosome has a unique origin of replication, eukaryotic chromosomes have multiple sites where DNA synthesis is initiated (Figure 20.4). The rate of fork movement in eukaryotes is slower than in bacteria but the presence of many independent origins of replication enables the larger eukaryotic genomes to be copied in approximately the same amount of time as prokaryotic genomes.

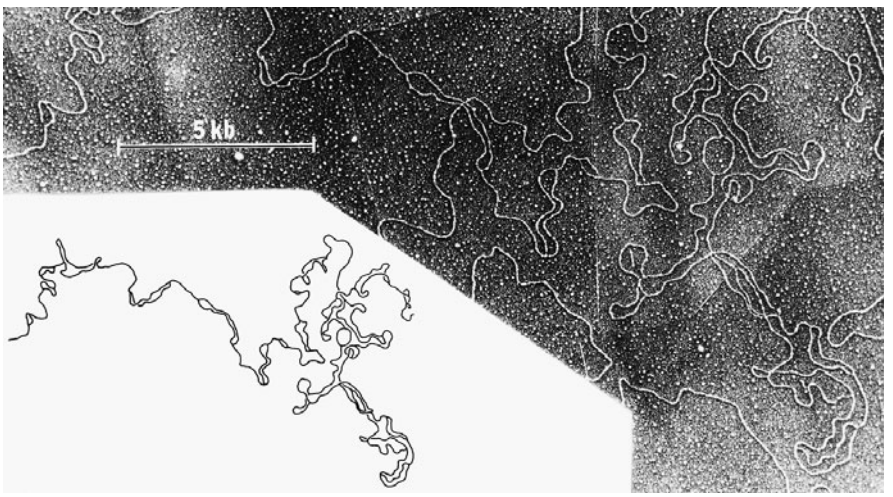


▲ **Figure 20.3**
Autoradiograph of a replicating *E. coli* chromosome. The DNA was labeled with ^{3}H-deoxythymidine, and the radioactivity detected by overlaying the replicating chromosome with photographic emulsion. The autoradiograph shows that the *E. coli* chromosome has two replication forks.

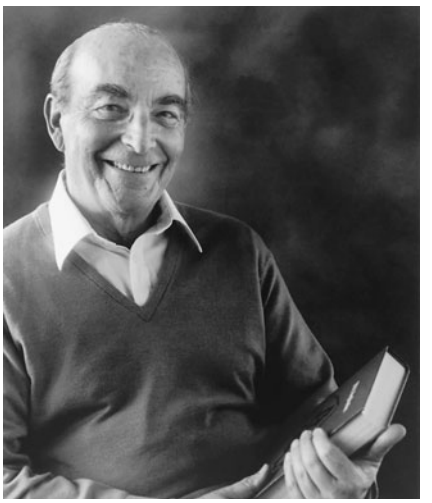
20.2 DNA Polymerase

The synthesis of a new strand of DNA is achieved by the successive addition of nucleotides to the end of a growing chain. This polymerization is catalyzed by enzymes known as DNA-directed DNA polymerases, or simply DNA polymerases. *E. coli* cells contain three different DNA polymerases; each protein is identified by a roman numeral according to the order of its discovery. DNA polymerase I repairs DNA and participates in the synthesis of one of the strands of DNA during replication. DNA polymerase II plays a role in DNA repair. DNA polymerase III is the major DNA replication enzyme responsible for chain elongation during DNA replication and is the essential part of the **replisome**.

DNA polymerase III contains ten different polypeptide subunits. It is by far the largest of the three DNA polymerases (Table 20.1). The purified holoenzyme is an asymmetric dimer consisting of two copies of each polypeptide as shown in Figure 20.5. The α, ε, and θ polypeptides combine to form two core complexes that are responsible for the polymerization reactions. The β subunits form a sliding clamp that surrounds each of the two DNA strands at the replication fork. Most of the remaining subunits make up the γ complex, or "clamp loader" that assists in assembly of the replisome and helps to keep the enzyme bound to parental DNA during successive polymerization reactions.



▲ **Figure 20.4**
Electron micrograph of replicating DNA from an embryo of the fruit fly *Drosophila melanogaster*. Note the large number of replication forks at opposite ends of "bubbles" of duplicated DNA.



▲ **Arthur Kornberg (1918–2007).** Kornberg received the Nobel Prize in 1959 for his discovery of DNA polymerase.

KEY CONCEPT

Two replication forks move in opposite directions from the origin of replication.

Table 20.1 Subunits of DNA polymerase III holoenzyme

Subunit	M_r		Gene	Activity
α	130 000	core	*polC/dnaE*	Polymerase
ε	27 000	core	*dnaQ/mutD*	3′ → 5′ exonuclease
θ	8846	core	*holE*	?
β	40 000		*dnaN*	Forms sliding clamp
τ	71 000		*dnaX*	Enhances dimerization of core; ATPase
γ	47 000	γ complex	*dnaX*	Enhance processivity; assist in replisome assembly
δ	38 700	γ complex	*holA*	
δ'	36 900	γ complex	*holB*	
χ	16 600	γ complex	*holC*	
ψ	15 174	γ complex	*holD*	

A. Chain Elongation Is a Nucleotidyl Group Transfer Reaction

All DNA polymerases, including DNA polymerase III, synthesize DNA by adding one nucleotide at a time to the 3′ end of the growing chain. The nucleotide substrate is a deoxyribonucleoside 5′-triphosphate (dNTP). The specific nucleotide is determined by Watson-Crick base pairing to the template strand; adenine (A) pairs with thymine (T) and guanine (G) pairs with cytosine (C). Since the pool of each dNTP in a cell is approximately equal, this means that on average the enzyme spends three quarters of its time discriminating against incorrect dNTPs that have diffused into the catalytic site where they try to base pair with the template strand.

DNA polymerase III catalyzes the formation of a phosphodiester linkage between the incoming dNTP and the growing chain. The incoming dNTP forms a base pair with a residue of the template strand (Figure 20.6). Once a correct base pair has formed, the free 3′-hydroxyl group of the nascent DNA chain carries out a nucleophilic attack on the α-phosphorus atom of the incoming dNTP. This reaction leads to the addition of a nucleoside monophosphate and displacement of pyrophosphate. Subsequent hydrolysis of the pyrophosphate by the abundant enzyme pyrophosphatase makes the polymerization reaction highly favorable in the direction of polymerization. The direction of polymerization (chain growth) is defined as 5′ → 3′, reading across the carbon atoms on the sugar ring of the newly added residue.

The convention for assigning the direction of DNA strands is described in Section 19.2A.

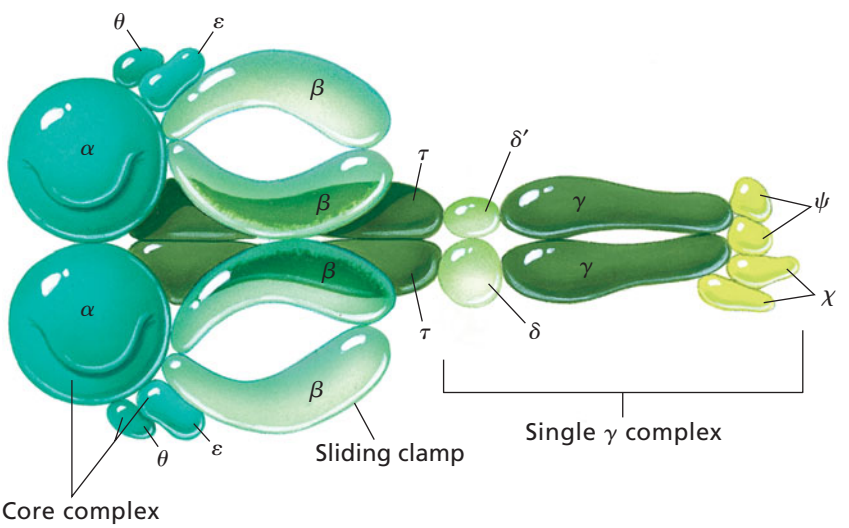
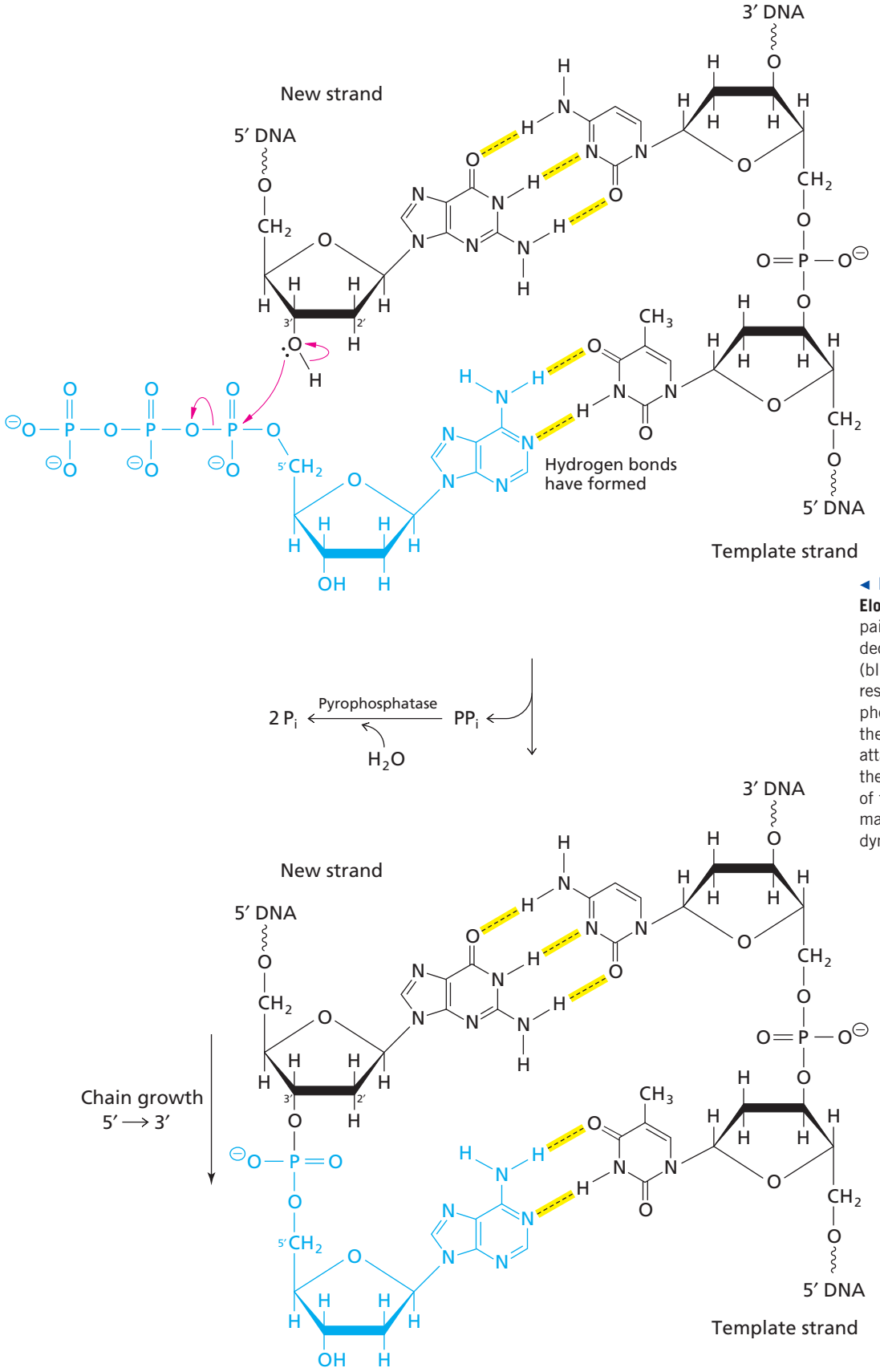


Figure 20.5 ▸
Diagram of the subunit composition of *E. coli* DNA polymerase III. The holoenzyme consists of two core complexes (containing α, ε and θ), paired copies of β and τ, and a single γ complex (γ, δ, δ', with two copies each of ψ, and χ). The structure is thus an asymmetric dimer. Other models of the holoenzyme structure have been proposed. [Adapted from O'Donnell, M. (1992). Accessory protein function in the DNA polymerase III holoenzyme from *E. coli*. *BioEssays* 14:105–111.]



◀ **Figure 20.6**
Elongation of a DNA chain. A base pair is created when an incoming deoxynucleoside 5′-triphosphate (blue) forms hydrogen bonds with a residue of the parental strand. A phosphodiester linkage forms when the terminal 3′-hydroxyl group attacks the α-phosphorus atom of the incoming nucleotide. Hydrolysis of the released pyrophosphate makes the overall reaction thermodynamically favorable.

DNA polymerase III advances by one residue, after each addition reaction, and another nucleotidyl group transfer reaction occurs. This mechanism ensures that the new chain is extended by the stepwise addition of single nucleotides that are properly aligned by base pairing with the template strand. As expected, DNA polymerase III cannot synthesize DNA in the absence of a template, nor can it add nucleotides in the absence of a 3′ end of a preexisting chain. In other words, DNA polymerase III requires both a template and a primer as substrates for synthesis to occur.

As noted earlier, DNA replication inside the cell proceeds at a rate of approximately 1000 nucleotides per second. This is the fastest known rate of any *in vivo* polymerization reaction. The rate of polymerization catalyzed by purified DNA polymerase III *in vitro*, however, is much slower, indicating that the isolated enzyme lacks some components necessary for full activity. Only when the complete replisome is assembled does polymerization *in vitro* occur at approximately the rate found inside the cell.

B. DNA Polymerase III Remains Bound to the Replication Fork

Once DNA synthesis has been initiated, the polymerase remains bound at the replication fork until replication is complete. The 3′ end of the growing chain remains associated with the active site of the enzyme while many nucleotides are added sequentially. As part of the replisome, the DNA polymerase III holoenzyme is highly *processive* (see Section 12.5A). This means that only a small number of DNA polymerase III molecules are needed to replicate the entire chromosome. Processivity also accounts for the rapid rate of DNA replication.

The processivity of the DNA polymerase III holoenzyme is due, in part, to the β subunits of the enzyme. These subunits have no activity on their own but when assembled into the holoenzyme they form a ring that can completely surround the DNA molecule. The ring is formed by two β subunits that form a head-to-tail dimer. Each of the subunits contains three similar domains consisting of a β sandwich fold with two α helices at the interior edge that interact with DNA (Figure 20.7). The β subunits thus act as a sliding clamp locking the polymerase onto the DNA substrate. Incorporating DNA polymerase III into an even larger protein machine at the replication fork further ensures that the enzyme remains associated with the nascent DNA chains during polymerization. Many other biochemically characterized DNA replication systems have evolved the same strategy to make DNA replication faster (more efficient). For example, two related bacteriophage, T_4 and RB69, both encode a replication accessory protein, gp45, that forms a circular clamp (Figure 20.7). This clamp structure locks the phage-encoded DNA polymerases onto their DNA substrates and enhances processivity. Figure 20.8 shows a model for how this is likely to work *in vivo* for bacteriophage DNA polymerase bound to DNA. The sliding clamp surrounds the double-stranded region of DNA and interacts with the subunits containing the polymerase activity that bind to the single-stranded region of the replication fork. Eukaryotic DNA polymerases use the same strategy to clamp onto their substrates (see Section 20.6).

The elongation reaction in fatty acid synthesis is another example of a processive polymerization reaction catalyzed by a large complex (Section 16.1C). The glycogen synthase reaction is an example of a distributive polymerization reaction (Section 12.5A).

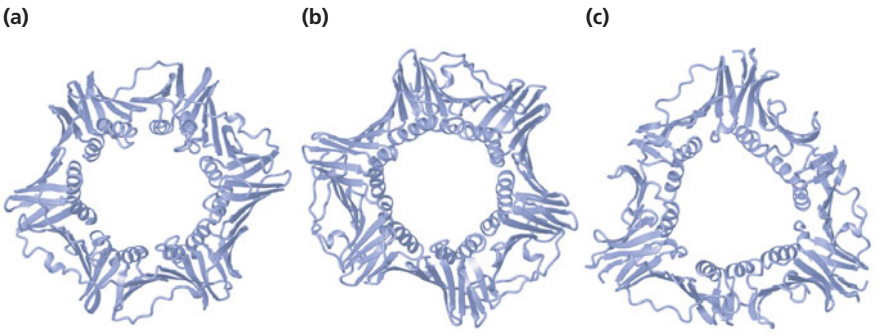
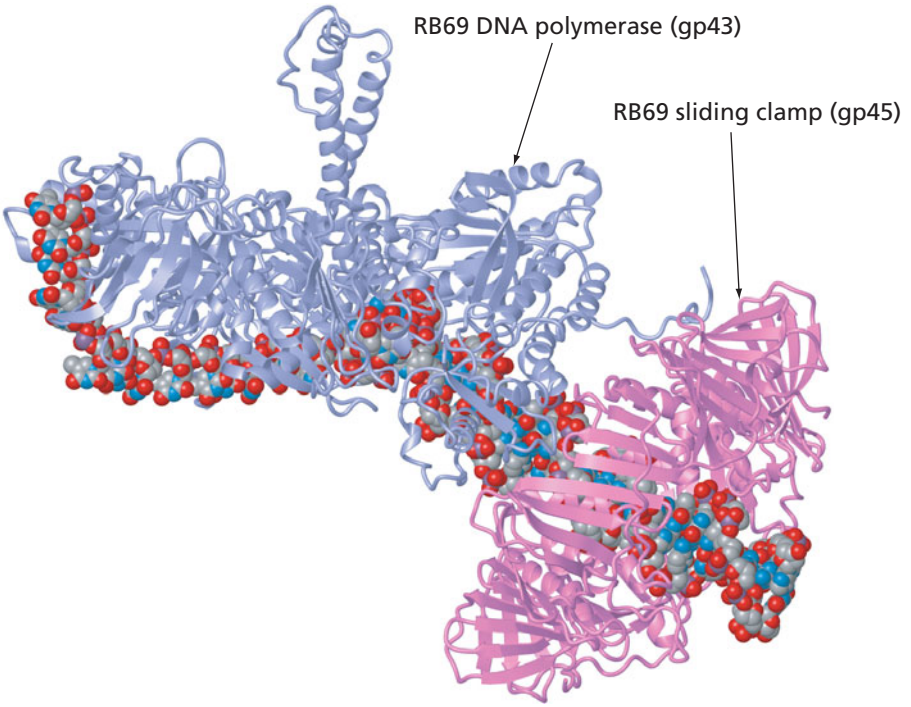


Figure 20.7 ▸
DNA polymerases can use sliding ring clamps to increase processivity. These three crystal structures show the convergent evolution of structure and function; **(a)** the β subunit of *E. coli* DNA polymerase III [PDB 1MMI]; **(b)** Proliferating Cell Nuclear Antigen (PCNA) that performs the same function in archaebacteria; [PDB 3LX1] **(c)** gp45 from bacteriophage T4 is also a sliding ring that clamps DNA polymerase to its DNA substrate. [PDB 1CZD]



◀ **Figure 20.8**
Model of bacteriophage DNA polymerase bound to DNA. The sliding clamp (purple) surrounds the newly synthesized double-stranded DNA. The subunit containing the active site is shown in blue. The 3′ end of the nascent strand is positioned at the active site and the single-stranded region of the template strand extends leftward. The DNA polymerase will move from right to left as the nascent strand is extended. [PDB 1WAI].

C. Proofreading Corrects Polymerization Errors

The DNA polymerase III holoenzyme also possesses a 3′ → 5′ exonuclease activity. This exonuclease, whose active site lies primarily within the ε subunit, can catalyze hydrolysis of the phosphodiester linkage that joins the 3′-terminal residue to the rest of the growing polynucleotide chain. Thus, the DNA polymerase III holoenzyme can catalyze both chain elongation and degradation. The exonuclease activity allows the holoenzyme to proofread, or edit, newly synthesized DNA in order to correct any mismatched base pairs. When DNA polymerase III recognizes a distortion in the DNA produced by an incorrectly paired base, the exonuclease activity of the enzyme catalyzes removal of the mispaired nucleotide before polymerization continues.

Proofreading is possible because the polymerization mechanism is head growth not tail growth (Box 12.3).

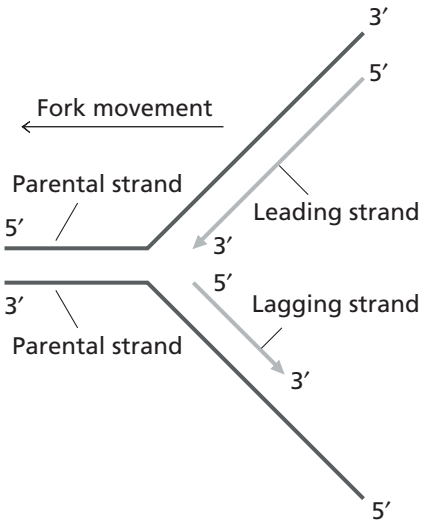
An incorrect base is incorporated about once every 10^5 polymerization steps for an error rate of about 10^{-5}. The 3′ → 5′ proofreading exonuclease activity will remove 99% of these incorrect nucleotides. It has an error rate of 10^{-2}. The combination of these two sequential reactions yields an error rate for polymerization of 10^{-7}. This is one of the lowest error rates of any enzyme. Most of these replication errors are subsequently repaired by separate DNA repair enzymes (Section 20.7) yielding an overall error rate for DNA replication of between 10^{-9} and 10^{-10}. Despite this impressive accuracy, replication errors are common when large genomes are duplicated. (Recall that the human genome contains 3.2×10^9 bp, which means that, on average, each time the genome is replicated an error gets transmitted to one of the two daughter cells.) Mistakes that occur during DNA replication are the most common source of mutation. What this means is that most of evolution is due to the inaccuracy of DNA replication!

KEY CONCEPT

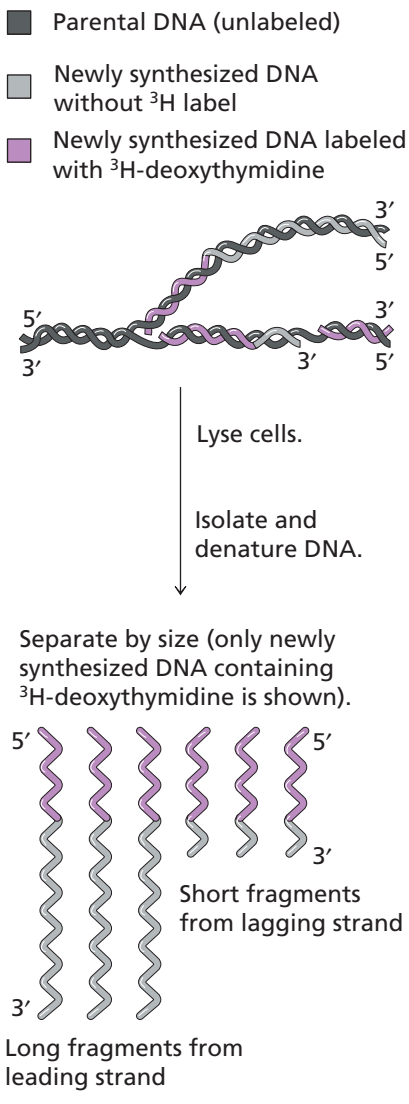
The accuracy of DNA polymerase combined with proofreading and DNA repair makes DNA replication the most accurate biochemical reaction known.

20.3 DNA Polymerase Synthesizes Two Strands Simultaneously

DNA polymerases catalyze chain elongation exclusively in the 5′ → 3′ direction, as shown in Figure 20.6. Since the two strands of DNA are antiparallel, 5′ → 3′ synthesis using one template strand occurs in the same direction as fork movement but 5′ → 3′ synthesis using the other template strand occurs in the direction opposite fork movement (Figure 20.9). The new strand formed by polymerization in the same direction as



▲ **Figure 20.9**
Diagram of the replication fork. The two newly synthesized strands have opposite polarity. On the leading strand, 5′ → 3′ synthesis moves in the same direction as the replication fork; on the lagging strand, 5′ → 3′ synthesis moves in the opposite direction.



fork movement is called the leading strand. The new strand formed by polymerization in the opposite direction is called the lagging strand. Recall that the DNA polymerase III holoenzyme dimer contains two core complexes that can catalyze polymerization. One of these is responsible for synthesis of the leading strand and the other is responsible for synthesis of the lagging strand.

A. Lagging Strand Synthesis Is Discontinuous

The leading strand is synthesized as one continuous polynucleotide beginning at the origin and ending at the termination site. In contrast, the lagging strand is synthesized discontinuously in short pieces in the direction opposite fork movement. These pieces of lagging strand are then joined by a separate reaction. In Section 20.4, we present a model of the replication fork that explains how one enzyme complex can synthesize both strands simultaneously.

An experiment that illustrates discontinuous DNA synthesis is shown in Figure 20.10. *E. coli* DNA is labeled with a short pulse of ^{3}H-deoxythymidine. The newly made DNA molecules are then isolated, denatured, and separated by size. The experiment detects two types of labeled DNA molecules: very large DNA molecules that collectively contain about half the radioactivity of the partially replicated DNA and shorter DNA fragments of about 1000 residues that collectively contain the other half of the radioactivity. The large DNA molecules arise from continuous synthesis of the leading strand while the shorter fragments arise from discontinuous synthesis of the lagging strand. The short pieces of lagging strand DNA are named **Okazaki fragments** in honor of their discoverer, Reiji Okazaki. The overall mechanism of DNA replication is called semidiscontinuous to emphasize the different mechanisms for replicating each strand.

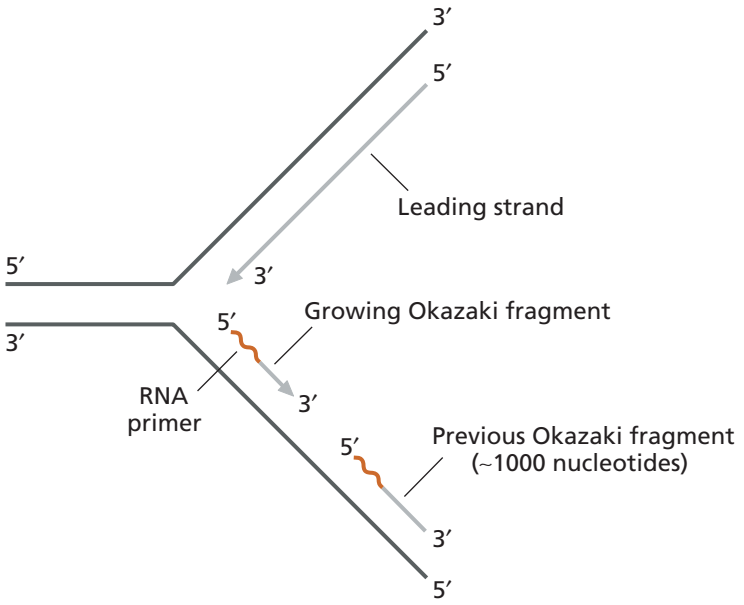
B. Each Okazaki Fragment Begins with an RNA Primer

It was clear that lagging strand synthesis is discontinuous but it was not obvious how synthesis of each Okazaki fragment is initiated. The problem is that no DNA polymerase can begin polymerization *de novo*; they can only add nucleotides to existing polymers. This limitation presents little difficulty for leading strand synthesis since once DNA synthesis is under way nucleotides are continuously added to a growing chain. However, on the lagging strand the synthesis of each Okazaki fragment requires a new initiation event. This is accomplished by making short pieces of RNA at the replication fork. These RNA primers are complementary to the lagging strand template. Each primer is extended from its 3′ end by DNA polymerase to form an Okazaki fragment as shown in Figure 20.11. (Synthesis of the leading strand also begins with an RNA primer but only one primer is required to initiate synthesis of the entire strand.)

The use of short RNA primers gets around the limitation imposed by the mechanism of DNA polymerase—namely, that it cannot initiate DNA synthesis *de novo*. The primers are synthesized by a DNA-dependent RNA polymerase enzyme called **primase**—the product of the *dnaG* gene in *E. coli*. The three-dimensional crystal structure of the DnaG catalytic domain revealed that its folding and active site are distinct from the well studied polymerases suggesting that it may employ a novel enzyme mechanism. Primase is part of a larger complex called the **primosome** that contains many other polypeptides in addition to primase. The primosome, along with DNA polymerase III, is part of the replisome.

As the replication fork progresses, the parental DNA is unwound and single-stranded DNA becomes exposed. Primase catalyzes the synthesis of a short RNA primer about once every second using this single-stranded DNA as a template. The primers are only a few nucleotides in length. Since the replication fork advances at a rate of about

◀ **Figure 20.10**
Discontinuous DNA synthesis demonstrated by analysis of newly synthesized DNA. Nascent DNA molecules are labeled in *E. coli* with a short pulse of ^{3}H-deoxythymidine. The cells are lysed, the DNA is isolated, and single strands are separated by size. The labeled DNA molecules fall into two classes: long molecules arising from continuous synthesis of the leading strand and short fragments arising from discontinuous synthesis of the lagging strand.



◂ **Figure 20.11**
Diagram of lagging strand synthesis. A short piece of RNA (brown) serves as a primer for the synthesis of each Okazaki fragment. The length of the Okazaki fragment is determined by the distance between successive RNA primers.

1000 nucleotides per second, one primer is synthesized for approximately every 1000 nucleotides that are incorporated. DNA polymerase III catalyzes synthesis of DNA in the 5′ → 3′ direction by extending each short RNA primer.

C. Okazaki Fragments Are Joined by the Action of DNA Polymerase I and DNA Ligase

Okazaki fragments are eventually joined to produce a continuous strand of DNA. The reaction proceeds in three steps: removal of the RNA primer, synthesis of replacement DNA, and sealing of the adjacent DNA fragments. The steps are carried out by the combined action of DNA polymerase I and DNA ligase.

DNA polymerase I of *E. coli* was the enzyme discovered by Arthur Kornberg. It was the first enzyme to be found that could catalyze DNA synthesis using a template strand. In a single polypeptide, DNA polymerase I contains the two activities found in the DNA polymerase III holoenzyme: 5′ → 3′ polymerase activity and 3′ → 5′ proofreading exonuclease activity. In addition, DNA polymerase I has 5′ → 3′ exonuclease activity, an activity not found in DNA polymerase III.

DNA polymerase I can be cleaved with certain proteolytic enzymes to generate a small fragment that contains the 5′ → 3′ exonuclease activity and a larger fragment that retains the polymerization and proofreading activities. The larger fragment consists of the C-terminal 605 amino acid residues, and the smaller fragment contains the remaining N-terminal 323 residues. The large fragment, known as the Klenow fragment, was widely used for DNA sequencing and is still used in many other techniques that require DNA synthesis without 5′ → 3′ degradation. In addition, many studies of the mechanisms of DNA synthesis and proofreading use the Klenow fragment as a model for more complicated DNA polymerases.

Figure 20.12 shows the structure of the Klenow fragment complexed with a fragment of DNA containing a mismatched terminal base pair. The 3′ end of the nascent strand is positioned at the 3′ → 5′ exonuclease site of the enzyme. During polymerization, the template strand occupies the groove at the top of the structure and at least 10 bp of double-stranded DNA are bound by the enzyme as shown in the figure. Many of the amino acid residues involved in binding DNA are similar in all DNA polymerases

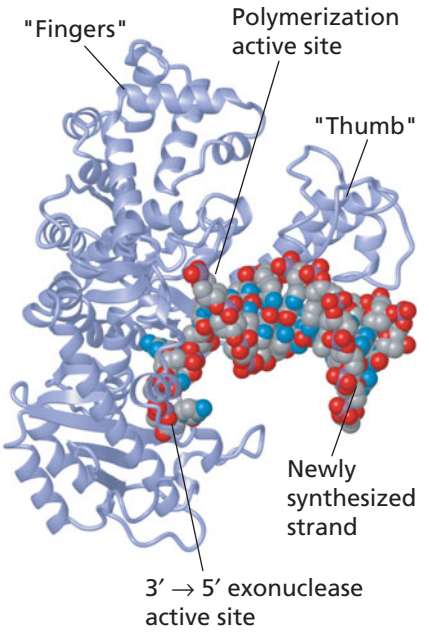
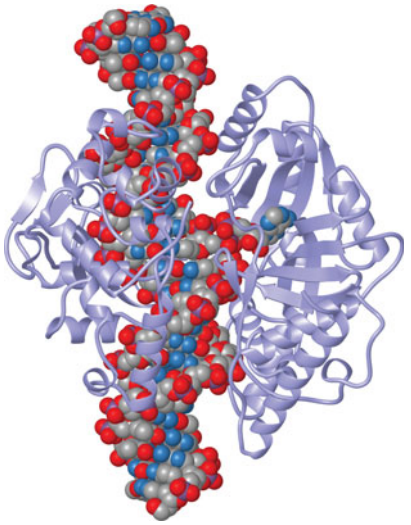
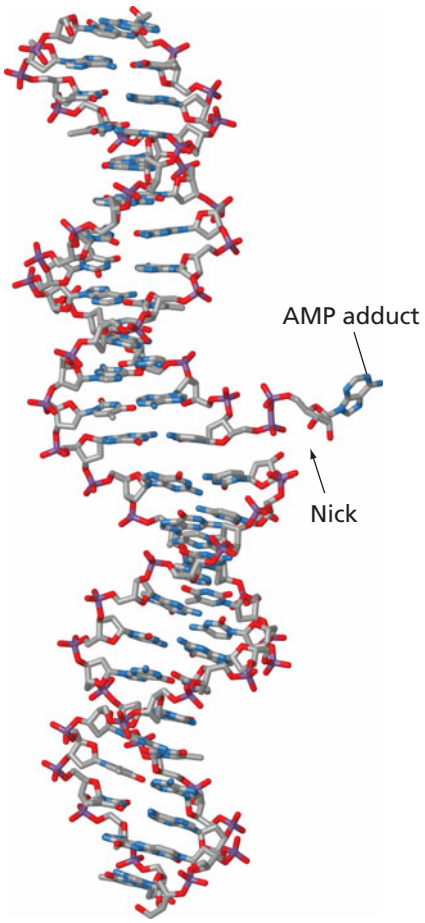


Figure 20.12 ▸
Structure of the Klenow fragment with a bound DNA fragment. The enzyme wraps around the DNA. The 3′ end of the nascent strand is positioned at the 3′ → 5′ exonuclease site (lower left). During DNA synthesis *in vivo* the template strand extends beyond the double-stranded region shown in the crystal structure. [PDB 1KLN].



▲ *E. coli* DNA ligase bound to nicked DNA. [PDB 20W0]



▲ Structure of nicked DNA substrate when bound by DNA ligase [PDB 20W0].

although the enzymes may be otherwise quite different in three-dimensional structure and amino acid sequence.

The unique 5′ → 3′ exonuclease activity of DNA polymerase I removes the RNA primer at the beginning of each Okazaki fragment. (Since it is not part of the Klenow fragment, the 5′ → 3′ exonuclease is not shown in Figure 20.12, but it would be located at the top of the structure next to the groove that accommodates the template strand.) As the primer is removed, the polymerase synthesizes DNA to fill in the region between Okazaki fragments, a process called **nick translation** (Figure 20.13). In nick translation, DNA polymerase I recognizes and binds to the nick between the 3′ end of an Okazaki fragment and the 5′ end of the next primer. The 5′ → 3′ exonuclease then catalyzes hydrolytic removal of the first RNA nucleotide while the 5′ → 3′ polymerase adds a deoxynucleotide to the 3′ end of the DNA chain. In this way, the enzyme moves the nick along the lagging strand. DNA polymerase I dissociates from the DNA after completing 10 or 12 cycles of hydrolysis and polymerization, leaving behind two Okazaki fragments that are separated by a nick in the phosphodiester backbone. The removal of RNA primers by DNA polymerase I is an essential part of DNA replication because the final product must consist entirely of double-stranded DNA.

The last step in the synthesis of the lagging strand of DNA is the formation of a phosphodiester linkage between the 3′-hydroxyl group at the end of one Okazaki fragment and the 5′-phosphate group of an adjacent Okazaki fragment. This step is catalyzed by DNA ligase. The DNA ligases in eukaryotic cells and in bacteriophage-infected cells require ATP as a cosubstrate. In contrast, *E. coli* DNA ligase uses $NAD^{\oplus}$ as a cosubstrate. $NAD^{\oplus}$ is the source of the nucleotidyl group that is transferred, first to the enzyme and then to the DNA, to create an ADP-DNA intermediate. The proposed mechanism of DNA ligase in *E. coli* is shown in Figure 20.14. The net reaction is

$$\text{DNA (nicked)} + NAD^{\oplus} \longrightarrow \text{DNA(sealed)} + NMN^{\oplus} + \text{AMP} \qquad \textbf{(20.1)}$$

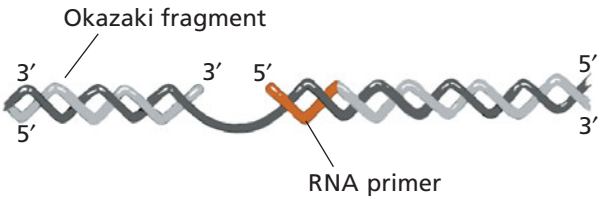
20.4 Model of the Replisome

The replisome contains a primosome, the DNA polymerase III holoenzyme, and additional proteins that are required for DNA replication. The assembly of many proteins into a single machine allows coordinated synthesis of the leading and lagging strands at the replication fork.

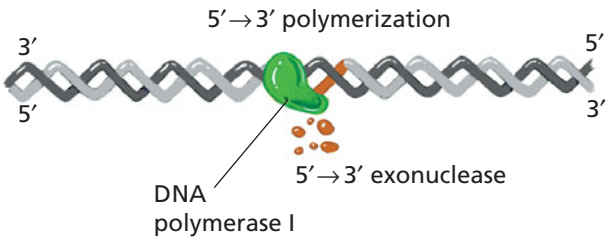
The template for DNA polymerase III is single-stranded DNA. This means that the two strands of the parental double helix must be unwound and separated during replication. This unwinding is accomplished primarily by a class of proteins called helicases. The helicase DnaB is required for DNA replication in *E. coli*. DnaB is one of the subunits of the primosome that, in turn, is part of the larger replisome. The rate of DNA unwinding is directly coupled to the rate of polymerization as the replisome moves along the chromosome. Unwinding is assisted by the actions of various topoisomerases (Section 19.3) that relieve supercoiling ahead of and behind the replication fork. These enzymes are not part of the replisome but they are required for replication. The most important topoisomerase in *E. coli* is topoisomerase II, or gyrase. Mutants lacking this enzyme cannot replicate their DNA. The end result is the production of two daughter molecules each containing one newly synthesized stand and one parental strand as shown in Figure 20.1. At no time during DNA replication is there a significant stretch of single-stranded DNA other than that found on the lagging strand template.

Another protein that is part of the replisome is single-strand binding protein (SSB), also known as helix-destabilizing protein. SSB binds to single-stranded DNA and prevents it from folding back on itself to form double-stranded regions. SSB is a tetramer of four identical small subunits. Each tetramer covers about 32 nucleotides of DNA. Binding of SSB to DNA is cooperative; that is, binding of the first tetramer facilitates binding of the second, and so on. The presence of several adjacent SSB molecules on single-stranded DNA produces an extended, relatively inflexible, DNA conformation. Single-stranded DNA coated with SSB is an ideal template for synthesis of the complementary strand during DNA replication because it is free of secondary structure.

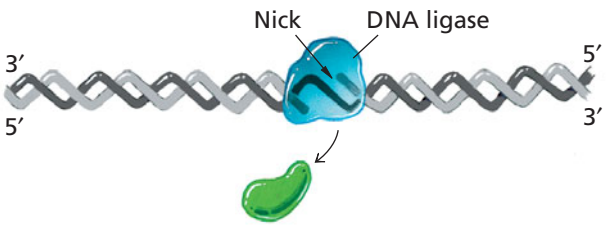
(a) Completion of Okazaki fragment synthesis leaves a nick between the Okazaki fragment and the preceding RNA primer on the lagging strand.



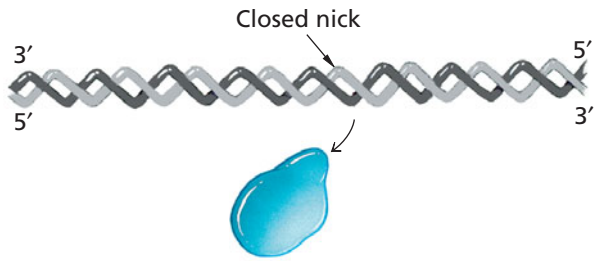
(b) DNA polymerase I extends the Okazaki fragment while its 5′→3′ exonuclease activity removes the RNA primer. This process, called nick translation, results in movement of the nick along the lagging strand.



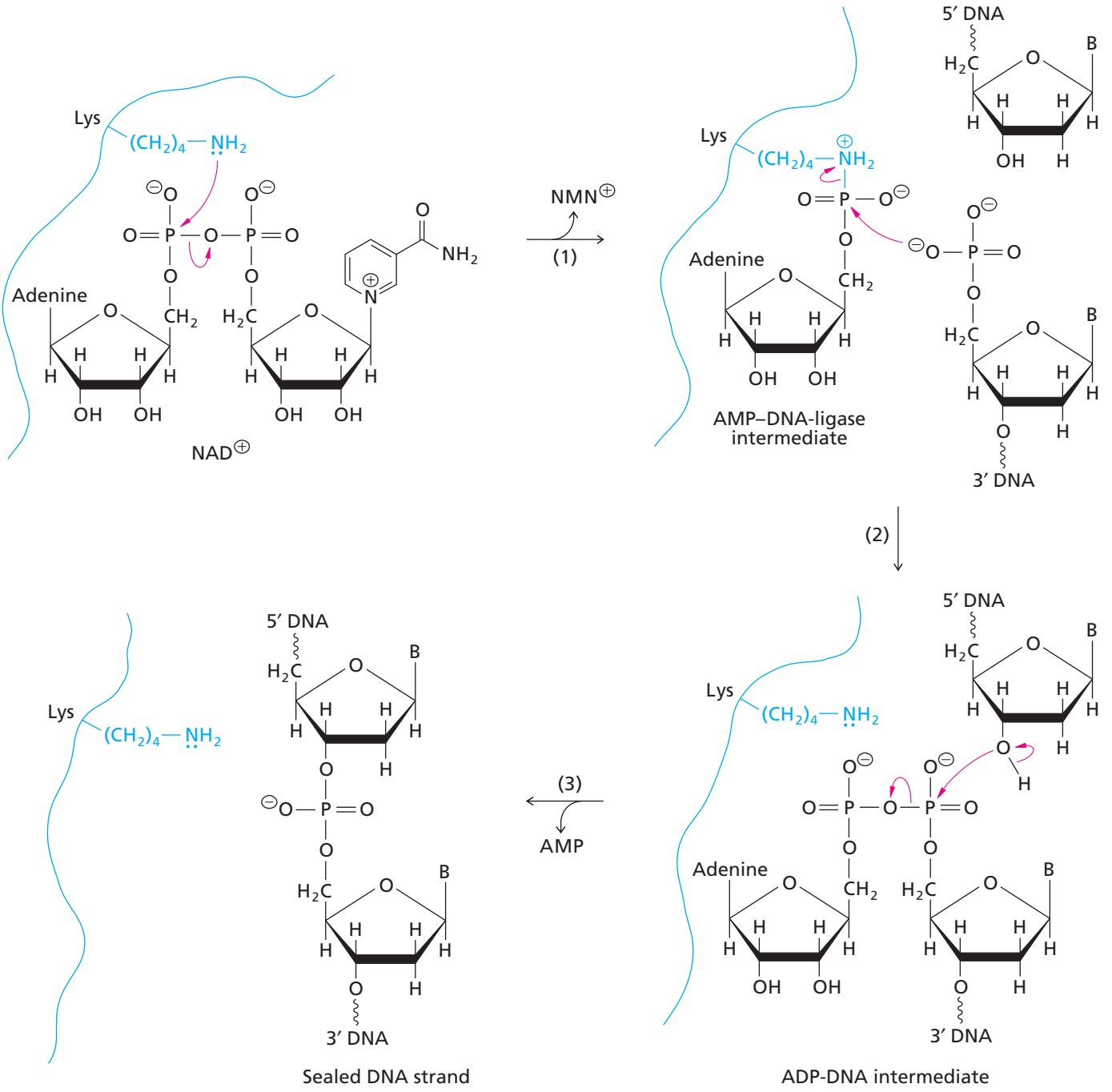
(c) DNA polymerase I dissociates after extending the Okazaki fragment 10–12 nucleotides. DNA ligase binds to the nick.



(d) DNA ligase catalyzes formation of a phosphodiester linkage, which seals the nick, creating a continuous lagging strand. The enzyme then dissociates from the DNA.

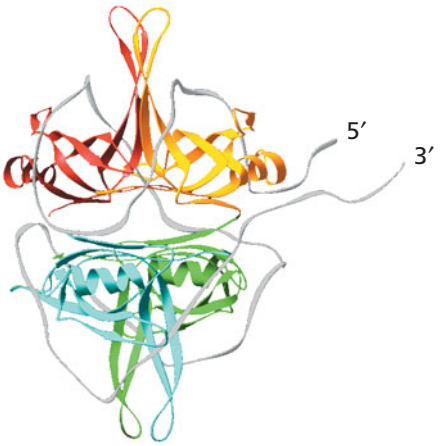


◀ **Figure 20.13**
Joining of Okazaki fragments by the combined action of DNA polymerase I and DNA ligase.

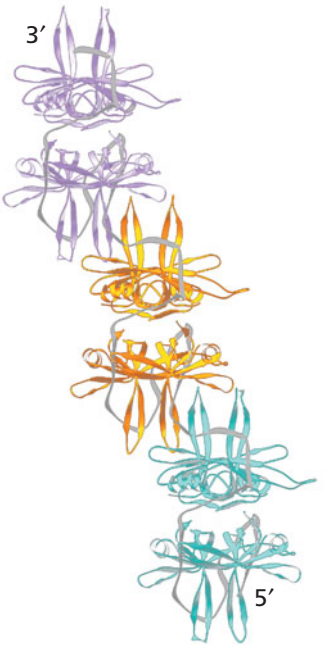


▲ Figure 20.14
Proposed mechanism of DNA ligase in *E. coli*. Using NAD⊕ as a cosubstrate, *E. coli* DNA ligase catalyzes the formation of a phosphodiester linkage at a nick in DNA. In Step 1, the ε-amino group of a lysine residue of DNA ligase attacks the phosphorus atom bonded to the 5′-oxygen atom of the adenosine moiety of NAD⊕. Nicotinamide mononucleotide (NMN⊕) is displaced, generating an AMP–DNA-ligase intermediate. (With DNA ligases that use ATP as the cosubstrate, pyrophosphate is displaced.) In Step 2, an oxygen atom of the free 5′-phosphate group of the DNA attacks the phosphate group of the AMP-enzyme complex, forming an ADP-DNA intermediate. In Step 3, the nucleophilic 3′-hydroxyl group on the terminal residue of the adjacent DNA strand attacks the activated 5′-phosphate group of ADP-DNA, releasing AMP and generating a phosphodiester linkage that seals the nick in the DNA strand. B represents any base.

A model of DNA synthesis by the replisome is shown in Figure 20.15. The primosome containing the primase and helicase is located at the head of the replication fork, followed by a DNA polymerase III holoenzyme. (In order to simplify the figure, only the core complexes of DNA polymerase III are shown.) Primase synthesizes an RNA primer approximately once every second as the helicase unwinds the DNA. One of the two core complexes in the holoenzyme dimer synthesizes the leading strand continuously



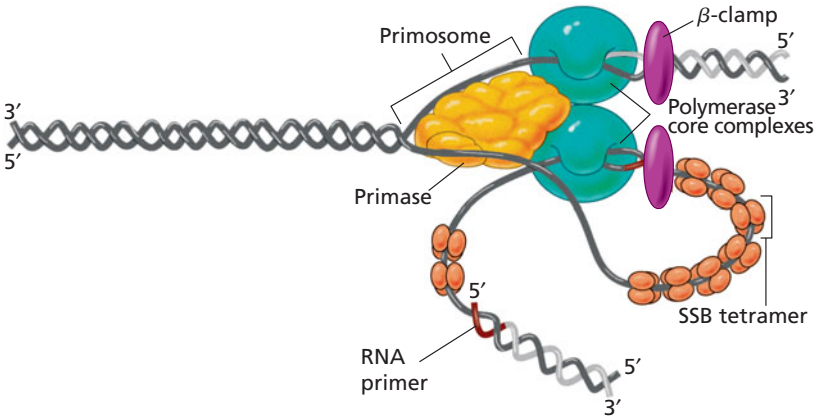
▲ Model for *E. coli* SSB tetramer bound to ssDNA [PDB 1EYG]



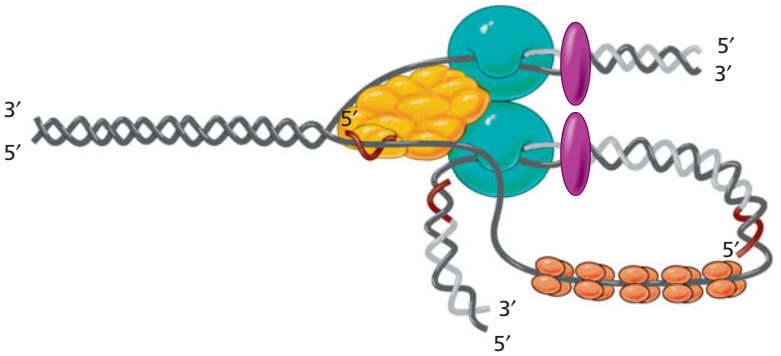
◄ DNA bound to SSB Model for the extended conformation of three SSB tetramers bound cooperatively to ssDNA. [PDB 1EYG]

Source: Nature Structural and Molecular Biology 7:648–652 (2000) Raghunathan et al.

(a) The lagging-strand template loops back through the replisome so that the leading and lagging strands are synthesized in the same direction. SSB binds to single-stranded DNA.



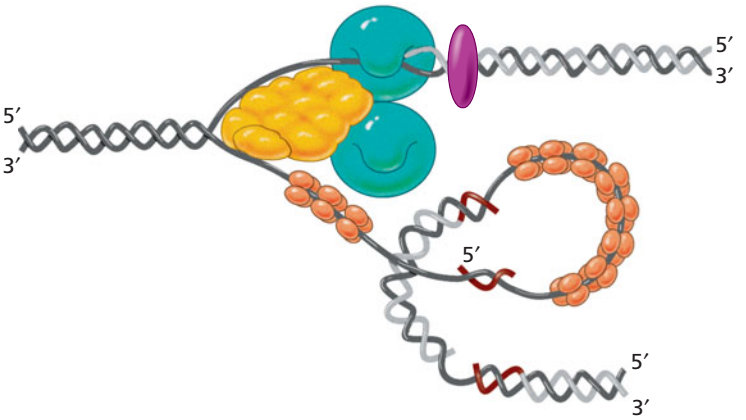
(b) As helicase unwinds the DNA template, primase synthesizes an RNA primer. The lagging-strand polymerase completes an Okazaki fragment.



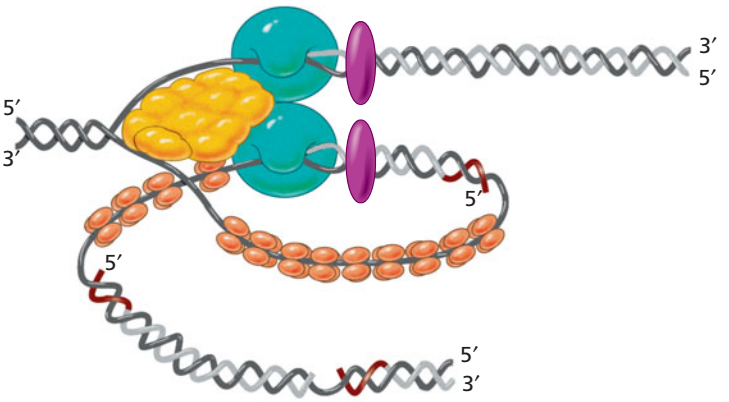
◄ **Figure 20.15**
Simultaneous synthesis of leading and lagging strands at a replication fork. The replisome contains the DNA polymerase III holoenzyme (only the core complexes are shown); a primosome containing primase, a helicase, and other subunits; and additional components including single-strand binding protein (SSB). One core complex of the holoenzyme synthesizes the leading strand while the other core complex synthesizes the lagging strand. The lagging-strand template is looped back through the replisome so that the leading and lagging strands can be synthesized in the same direction as fork movement. (c) and (d) continue on the next page.

Figure 20.15 (Continued) ►

(c) When the lagging-strand polymerase encounters the preceding Okazaki fragment, it releases the lagging strand.



(d) The lagging-strand polymerase binds to a newly synthesized primer and begins synthesizing another Okazaki fragment.



in the 5′ ⟶ 3′ direction while the other extends the RNA primers to form Okazaki fragments. The lagging-strand template is thought to fold back into a large loop. This configuration allows both the leading and lagging strands to be synthesized in the same direction as fork movement.

The two core complexes of the DNA polymerase III holoenzyme are drawn in the model as equivalent but their roles in DNA replication are not equivalent. One of them remains firmly bound to the leading-strand template whereas the other binds the lagging-strand template until it encounters the RNA primer of the previously synthesized Okazaki fragment. At this point the core complex releases the lagging-strand template. The lagging-strand template reassociates with the holoenzyme at the site of the next primer and synthesis continues (Figure 20.15d). The entire holoenzyme is extremely processive since half of it remains associated with the leading strand from the beginning of replication until termination while the other half processively synthesizes stretches of 1000 nucleotides in the lagging strand. The γ complex of the holoenzyme aids in binding and releasing the lagging-strand template by participating in the removal and reassembly of the sliding clamp formed by the β subunits.

The replisome model explains how synthesis of the leading and lagging strands is coordinated. The structure of the replisome also ensures that all the components necessary for

replication are available at the right time, in the right amount, and in the right place. Complexes of proteins that function together to carry out a biochemical task are frequently called protein machines. The replisome is an example of a protein machine, as are the bacterial flagellum (Chapter 4), the ATP synthase complex (Chapter 14), the photosynthetic reaction center (Chapter 15), and several others that are discussed in the following chapters.

20.5 Initiation and Termination of DNA Replication

As noted earlier, DNA replication begins at a specific DNA sequence called the origin. In *E. coli*, this site is called *oriC*, and it is located at about 10 o'clock on the genetic map of the chromosome (Figure 20.16). The initial assembly of replisomes at *oriC* depends on proteins that bind to this site causing local unwinding of the DNA. One of these proteins, DnaA, is encoded by the *dnaA* gene that is located very close to the origin. DnaA helps regulate DNA replication by controlling the frequency of initiation. The initial RNA primers required for leading-strand synthesis are probably made by the primosomes at the origin.

Termination of replication in *E. coli* occurs at the termination site (*ter*), a region opposite the origin on the circular chromosome. This region contains DNA sequences that are binding sites for a protein known as terminator utilization substance (Tus). The structure of Tus bound to a single termination site is shown in Figure 20.17. Regions of β strand lie in the major groove of DNA where the amino acid side chains make contact with the base pairs and recognize the *ter* sequence. Tus prevents replication forks from passing through the region by inhibiting the helicase activity of the replisome. The termination site also contains DNA sequences that play a role in the separation of daughter chromosomes when DNA replication is completed.

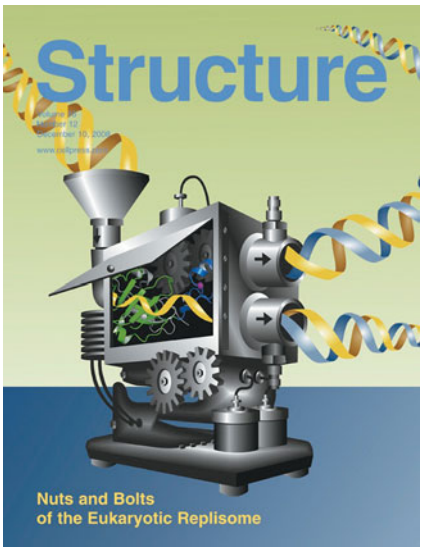
20.6 DNA Replication Technologies

Our understanding of the basic principles of DNA replication has led to the development of some amazing technologies that Watson and Crick could never have anticipated in 1953. We have already encountered site-directed matagenesis (Box 6.1). In this section we explore amplification and sequencing technologies that have transformed biochemistry and, indeed, all biology. These technologies have produced genome sequences of extinct species (e.g., *Homo neanderthalensis*) and to the discovery of the genetic basis of many human traits and diseases.

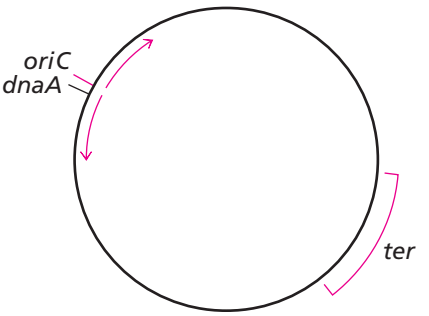
A. The Polymerase Chain Reaction Uses DNA Polymerase to Amplify Selected DNA Sequences

The **polymerase chain reaction** (PCR) is a valuable tool for amplifying a small amount of DNA or increasing the proportion of a particular DNA sequence in a population of mixed DNA molecules. The use of PCR technology avoids the need to take large samples of tissue in order to obtain enough DNA to manipulate for sequencing or cloning. The polymerase chain reaction also enables the production of a large number of copies of a gene that has not been isolated but whose sequence is known. It thus can serve as an alternative to cloning for gene amplification.

The PCR technique is illustrated in the figure on page 621. Sequence information from both sides of the desired locus is used to construct oligonucleotide primers that flank the DNA sequence to be amplified. The oligonucleotide primers are complementary to opposite strands and their 3′ ends are oriented toward each other. The DNA from the source (usually representing the entire DNA in a cell) is denatured by heating in the presence of excess oligonucleotides. On cooling, the primers preferentially anneal to their complementary sites, which border the DNA sequence of interest. The primers are then extended using a heat-stable DNA polymerase, such as *Taq* polymerase from the thermophilic bacterium *Thermus aquaticus*. After one cycle of synthesis, the reaction mixture



▲ **Protein machines.** Sometimes the machine metaphor can be taken too literally as in this humorous cover from the Journal *Structure*.



▲ **Figure 20.16**
Location of the origin (*ori*C) and terminus (*ter*) of DNA replication in *E. coli*. *dnaA* is the gene for the protein DnaA, which is required to initiate replication. The distance between *oriC* and *dnaA* is about 40 kb. The red arrows indicate the direction of movement of the replication forks.

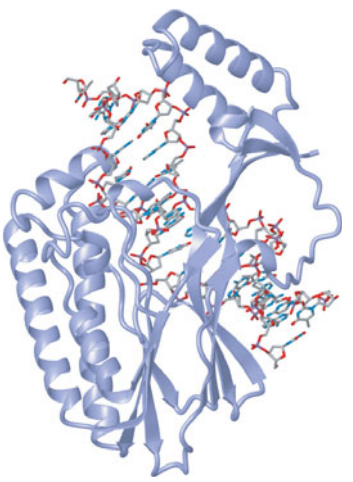


Figure 20.17 ►
Structure of *E. coli* Tus bound to DNA. Tus binds to specific sequences at the termination site of DNA replication. The bound protein blocks movement of the replisome. [PDB 1ECR].

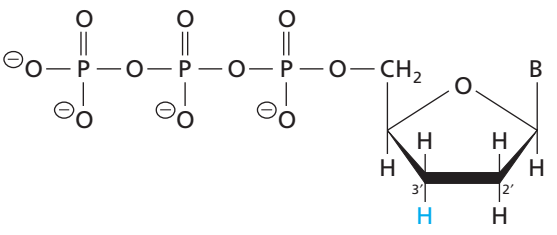
is again heated to dissociate the DNA strands and cooled to reanneal the DNA with the oligonucleotides. The primers are then extended again. In this second cycle, two of the newly synthesized, single-stranded chains are precisely the length of the DNA between the 5′ ends of the primers. The cycle is repeated many times, with reaction time and temperature carefully controlled. With each cycle, the number of DNA strands whose 5′ and 3′ ends are defined by the ends of the primers increases exponentially, whereas the number of DNA strands including sequences outside the region bordered by the primers increases arithmetically. As a result, the desired DNA is preferentially replicated until, after 20 to 30 cycles, it makes up most of the DNA in the test tube. The target DNA sequence can then be cloned, sequenced, or used as a probe for screening a recombinant DNA library.

B. Sequencing DNA Using Dideoxynucleotides

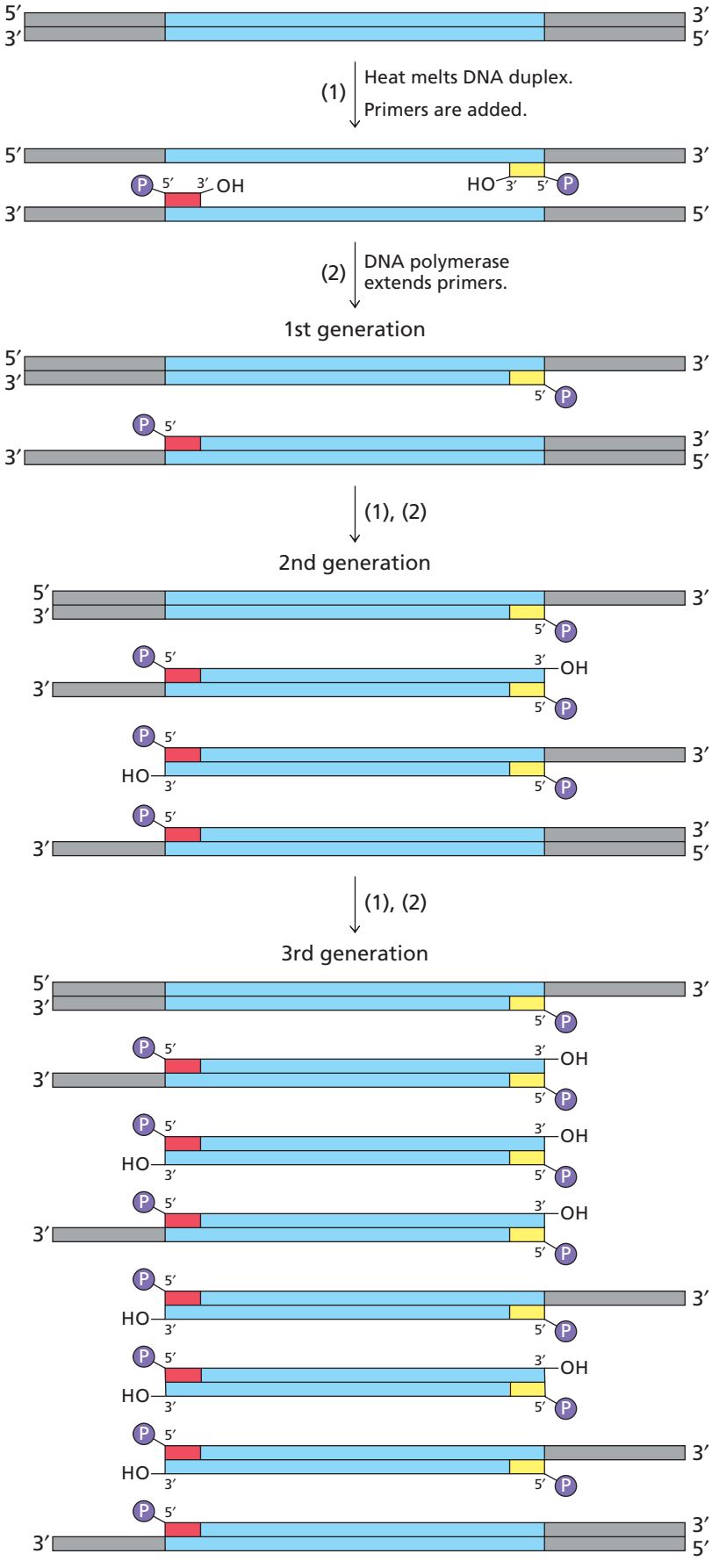
In 1976 Frederick Sanger developed a method for sequencing DNA enzymatically using the Klenow fragment of *E. coli* DNA polymerase I. Sanger was awarded his second Nobel Prize for this achievement (he received his first Nobel Prize for developing a method for sequencing proteins). The advantage of using the Klenow fragment for this type of reaction is that the enzyme lacks 5′ → 3′ exonuclease activity, which could degrade newly synthesized DNA. However, one of the disadvantages is that the Klenow fragment is not very processive and is easily inhibited by the presence of secondary structure in the single-stranded DNA template. This limitation can be overcome by adding SSB or analogous proteins, or more commonly, by using DNA polymerases from bacteria that grow at high temperatures. Such polymerases are active at 60° to 70°C, a temperature at which secondary structure in single-stranded DNA is unstable.

The Sanger sequencing method uses 2′, 3′-dideoxynucleoside triphosphates (ddNTPs), which differ from the deoxyribonucleotide substrates of DNA synthesis by lacking a 3′-hydroxyl group (see below). The dideoxyribonucleotides, which can serve as substrates for DNA polymerase, are added to the 3′ end of the growing chain. Because these nucleotides lack a 3′-hydroxyl group, subsequent nucleotide additions cannot take place and incorporation of a dideoxynucleotide terminates the growth of the DNA chain. When a small amount of a particular dideoxyribonucleotide is included in a DNA synthesis reaction, it is occasionally incorporated in place of the corresponding dNTP, immediately terminating replication. The length of the resulting fragment of DNA identifies the position of the nucleotide that should have been incorporated.

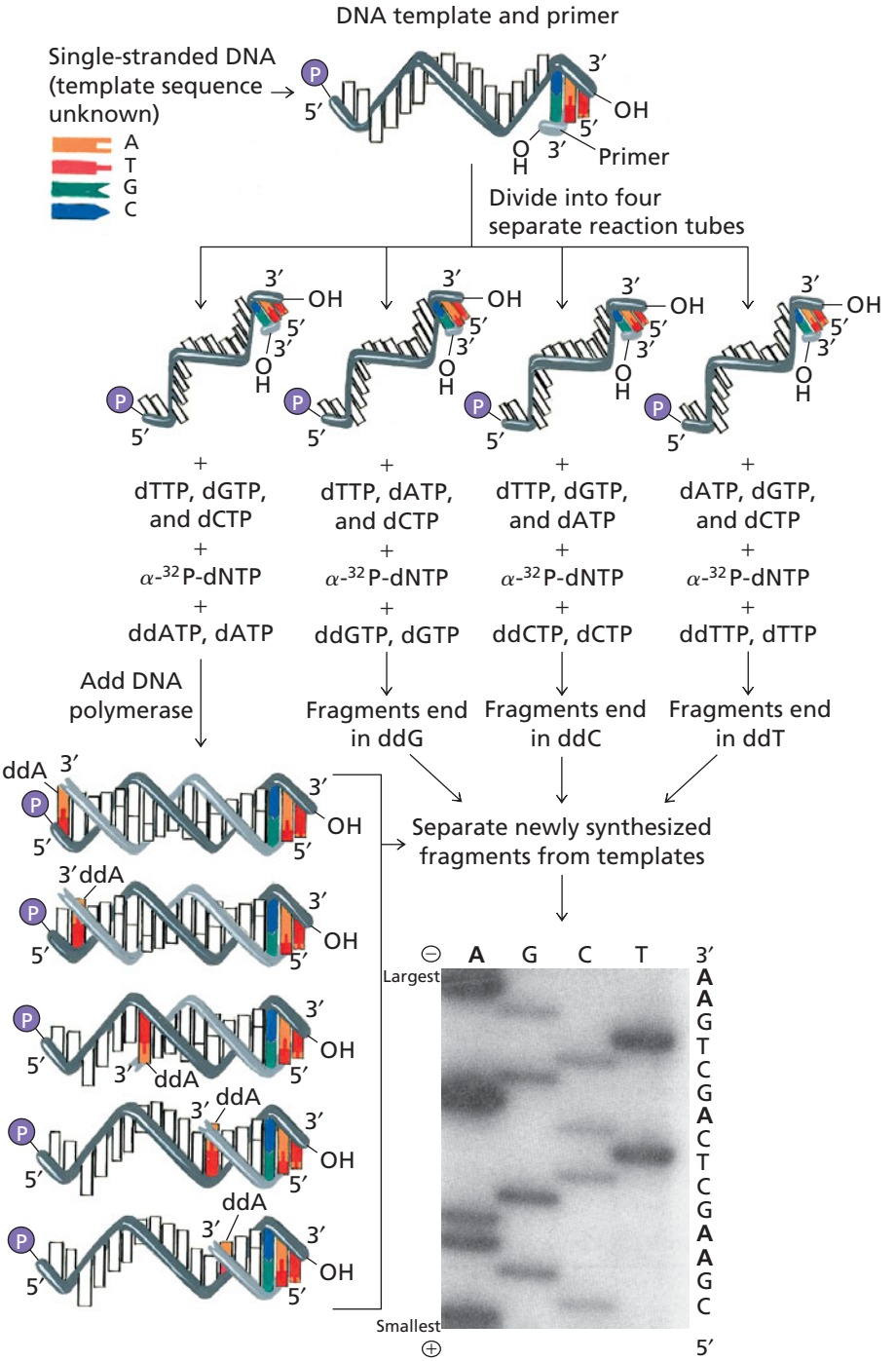
DNA sequencing using ddNTP molecules involves several steps (as shown on page 622). The DNA is prepared as single-stranded molecules and mixed with a short oligonucleotide complementary to the 3′ end of the DNA to be sequenced. This oligonucleotide acts as a primer for DNA synthesis catalyzed by DNA polymerase. The oligonucleotide-primed material is split into four separate reaction tubes. Each tube receives a small amount of an α-[^{32}P]-labeled dNTP, whose radioactivity allows the newly synthesized DNA to be visualized by autoradiography. Next, each tube receives an excess of the four nonradioactive dNTP molecules and a small amount of one of the four ddNTPs. For example, the A reaction tube receives an excess of nonradioactive dTTP, dGTP, dCTP, and dATP mixed with a small amount of ddATP. DNA polymerase is then added to the reaction mixture. As the polymerase replicates the DNA, it occasionally incorporates a ddATP residue instead of a dATP residue, and synthesis of the growing DNA chain is terminated. Random incorporation of ddATP results in the production of newly synthesized DNA fragments of different lengths, each ending with A (i.e., ddA). The length of each fragment corresponds to the distance from the 5′-end of the primer to one of the adenine residues in the sequence. Adding a different dideoxyribonucleotide



▶ Chemical structure of a 2′,3′-dideoxynucleoside triphosphate. *B* represents any base.



◀ **Three cycles of the polymerase chain reaction.** The sequence to be amplified is shown in blue. (1) The duplex DNA is melted by heating and cooled in the presence of a large excess of two primers (red and yellow) that flank the region of interest. (2) A heat-stable DNA polymerase catalyzes extension of these primers, copying each DNA strand. Successive cycles of heating and cooling in the presence of the primers allow the desired sequence to be repeatedly copied until, after 20 to 30 cycles, it represents most of the DNA in the reaction mixture.



Sanger method for sequencing DNA. ▶ Addition of a small amount of a particular dideoxynucleoside triphosphate (ddNTP) to each reaction mixture causes DNA synthesis to terminate when that dideoxynucleotide is incorporated in place of the normal nucleotide. The positions of incorporated dideoxynucleotides, determined by the lengths of the DNA fragments, indicate the positions of the corresponding nucleotide in the sequence. The fragments generated during synthesis with each ddNTP are separated by size using an electrophoretic sequencing gel, and the sequence of the DNA can be read from an autoradiograph of the gel (as shown by the column of letters to the right of the gel).

to each reaction tube produces a different set of fragments: ddTTP produces fragments that terminate with T, ddGTP with G, and ddCTP with C. The newly synthesized chains from each sequencing reaction are separated from the template DNA. Finally, the mixtures from each sequencing reaction are subjected to electrophoresis in adjacent lanes on a sequencing gel, where the fragments are resolved by size. The sequence of the DNA molecule can then be read from an autoradiograph of the gel.

This technique has also been modified to allow automation for high throughput applications like genomic sequencing. Instead of using radioactivity, automated sequencing relies on fluorescently labeled deoxynucleotides (four colors, one for each base) to detect the different chain lengths. In this system the gel is "read" by a fluorimeter and the data are stored in a computer file. Additionally, the sequencing machine can also provide a graphic chromatogram that shows the location and size of each fluorescent peak on the gel as they passed the detector.

C. Massively Parallel DNA Sequencing by Synthesis

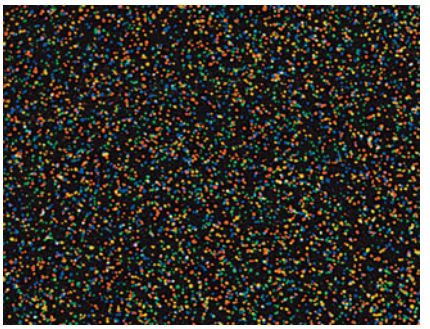
The automated DNA sequencing methods used to sequence the human genome have now been largely supplanted by a variety of so-called "next generation" sequencing technologies. While using slightly different experimental approaches, these devices can all rapidly generate millions (or even billions) of base pairs of sequence at a fraction of the cost of the automated Sanger technology described in the previous section. As an example of this novel approach, we describe the Illumina next-generation sequencing protocol.

In the first step, DNA (typically the entire genome) is randomly fragmented by shearing to yield short double-stranded fragments. The ends of the fragments are enzymatically repaired and a single-stranded oligonucleotide primer is ligated onto each end. Fragments of the desired length are purified from an agarose gel and then amplified using PCR. Oligonucleotides complementary to the PCR primers are covalently attached to the surface of a glass slide. The amplified genomic fragments are denatured into single strands, diluted, and hybridized to the oligonucleotides on the slide.

This creates a slide where millions of individual DNA fragments are bound to the surface. Each one is surrounded by a zone of free oligonucleotides bound to the surface. The individual DNA fragments on the slide's surface are then amplified *in situ* using a bridging technique to yield amplification clusters that are the substrate for the sequencing reaction.

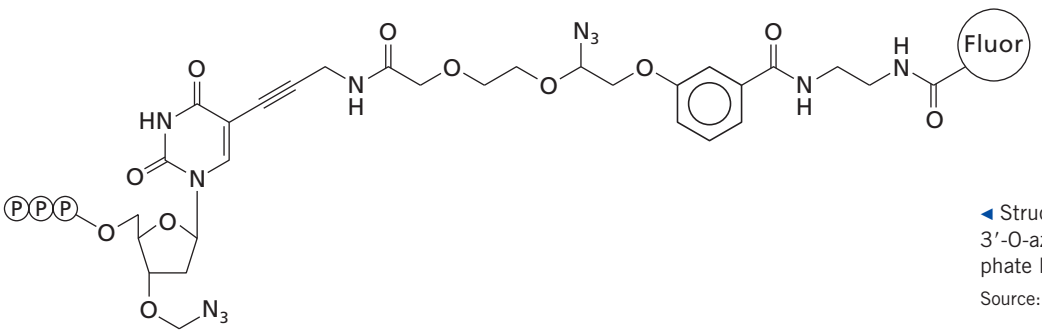
All of the clusters of amplified DNA fragments are sequenced at the same time, in parallel, using a mixture of the four dNTPs that have been labeled with a removable fluorophore (a different dye for each base) and a reversible terminator at the 3′position (see below). To increase the efficiency of this step, a genetically engineered mutant DNA polymerase from the deep hydrothermal vent archeon 9°N–7 that efficiently incorporates these bulky substrates is used. The DNA sequencing primer annealed to the template strands provides the 3′ hydroxyl group and the polymerase incorporates the next labeled nucleotide. The terminator at the 3′position of the incoming base prevents DNA synthesis beyond one single base. The slide is scanned using a laser-scanning confocal microscope to record the base that was incorporated into each growing cluster. The reducing agent TCEP is then added removing both the dye and the terminator to regenerate the 3′-OH. The whole cycle is then repeated. The growing DNA chains can only increase in length via a stepwise process: one base at a time.

The relatively short sequences (less than 100 nucleotides) are not suitable for assembling the genome sequence from a species that has never been sequenced before. However for resequencing a previously sequenced genome, fast computer algorithms can align these short "reads" with high accuracy and detect rare mutations or polymorphisms present in the sample.



▲ **Imaging clusters during the sequencing process.** Part of the image of a flow-cell with a low density of clusters is shown. Since each of the four deoxynucleotide bases is labeled with a different fluophore (each of which fluoresces at a different wavelength), the four separate images have been superimposed (after artificial coloring). After each cycle of DNA synthesis these images provide the raw data that reveal the last base that was incorporated into the growing polynucleotide chain.

Source: Bentley et al. (2008). *Nature* 456:53–59.



◄ Structure of the reversible terminator 3′-O-azidomethyl 2′-deoxythymine triphosphate labeled with a removable fluorophore.

Source: Bentley et al. (2008). *Nature* 456: 53–59.

20.7 DNA Replication in Eukaryotes

The mechanisms of DNA replication in prokaryotes and eukaryotes are fundamentally similar. In eukaryotes as in *E. coli*, synthesis of the leading strand is continuous and synthesis of the lagging strand is discontinuous. Furthermore, in both prokaryotes and eukaryotes, synthesis of the lagging strand is a stepwise process involving: primer synthesis, Okazaki fragment synthesis, primer hydrolysis, and gap filling by a polymerase. Eukaryotic primase, like prokaryotic primase, synthesizes a short primer once every second on the lagging-strand template. However, because the replication fork moves more slowly in eukaryotes, each Okazaki fragment is only about 100 to 200 nucleotide

Table 20.2 Eukaryotic DNA polymerases

DNA polymerase	Activities	Role
α	Polymerase Primase $3' \rightarrow 5'$ Exonuclease[a]	Primer synthesis Repair
β	Polymerase	Repair
γ	Polymerase $3' \rightarrow 5'$ Exonuclease	Mitochondrial DNA replication
δ	Polymerase $3' \rightarrow 5'$ Exonuclease	Leading- and lagging-strand synthesis Repair
ε	Polymerase $3' \rightarrow 5'$ Exonuclease $5' \rightarrow 3'$ Exonuclease	Repair Gap filling on lagging strand

[a]Polymerase $\alpha 3' \rightarrow 5'$ exonuclease activity is not detectable in all species.

residues long, considerably shorter than in prokaryotes. Interestingly, eukaryotic DNA primase does not share significant sequence similarity with the *E. coli* enzyme nor does eukaryotic primase contain some of the classical structural landmarks of DNA polymerases such as the "fingers" or "thumb" domains (Figure 20.12). This lack of homology suggests that the capacity to synthesize an RNA primer for DNA initiation may have evolved independently at least twice.

Most eukaryotic cells contain at least five different DNA polymerases: α, β, γ, δ, and ε (Table 20.2). DNA polymerases α, δ, and ε are responsible for the chain elongation reactions of DNA replication and for some repair reactions. DNA polymerase β is a DNA repair enzyme found in the nucleus and DNA polymerase γ plays a role in replicating mitochondrial DNA. A sixth DNA polymerase is responsible for replicating DNA in chloroplasts.

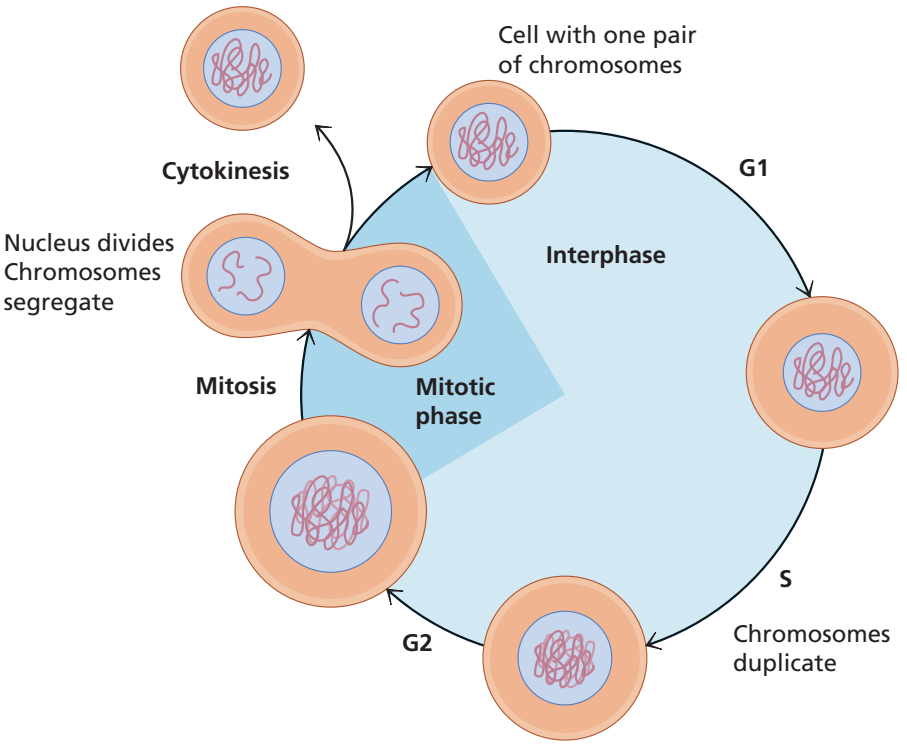
DNA polymerase δ catalyzes synthesis of the leading strand at the replication fork. This enzyme is composed of two subunits the larger of which contains the polymerase active site. The enzyme also has $3' \rightarrow 5'$ exonuclease activity. DNA replication in eukaryotic cells is extremely accurate. The low error rate indicates that DNA replication in eukaryotes includes an efficient proofreading step.

DNA polymerase α and DNA polymerase δ cooperate in lagging strand synthesis. DNA polymerase α is a multimeric protein that contains both DNA polymerase and RNA primase activity. The primer made by DNA polymerase α consists of a short stretch of RNA followed by DNA. This two part primer is extended by DNA polymerase δ to complete an Okazaki fragment.

DNA polymerase ε is a large, multimeric protein. The largest polypeptide chain includes polymerase activity and $3' \rightarrow 5'$ proofreading exonuclease activity. Like its functional counterpart in *E. coli* (DNA polymerase I), DNA polymerase ε probably acts as a repair enzyme and also fills gaps between Okazaki fragments.

Several accessory proteins are associated with the replication fork in eukaryotes. These proteins function like some of the proteins in the bacterial replisome. For example, PCNA (proliferating cell nuclear antigen) forms a structure that resembles the β-subunit sliding clamp of *E. coli* DNA polymerase III (Figure 20.7). The accessory protein RPC (replication factor C) is structurally, functionally, and evolutionarily related to the γ complex of DNA polymerase III. Another protein, called RPA (replication factor A), is the eukaryotic equivalent of prokaryotic SSB. In addition, the eukaryotic replication machine includes helicases that unwind DNA at the replication fork.

Each eukaryotic chromosome contains many origins of replication (Section 20.1). The largest chromosome of the fruit fly *Drosophila melanogaster*, for example, contains about 6000 replication forks implying that there are at least 3000 origins. As replication proceeds bidirectionally from each origin the forks move toward one another, merging to form bubbles of ever increasing size (Figure 20.4). Due to the large number of origins, the larger chromosomes of eukaryotes can still be replicated in less than one hour even though the rate of individual fork movement is much slower than in prokaryotes.



◀ **Figure 20.18**
The eukaryotic cell division cycle coordinates DNA replication and mitosis. DNA replication occurs exclusively during the synthesis, or S-phase of the cell cycle. There are two gap, or G, phases where a cell grows prior to dividing in the mitosis, or M-phase.

DNA replication in all cells occurs within the context of the cell's programmed cell division cycle. This cell cycle is a highly regulated progression through a series of dependent steps that at a minimum accomplishes two goals: (1) it faithfully duplicates all of the DNA in a cell to produce exactly two copies of each chromosome, and (2) it precisely segregates one copy of each replicated chromosome into one of the two daughter cells. In eukaryotic cells chromosomal segregation occurs at mitosis and this stage is called the mitotic phase, or M-phase (Figure 20.18). The step where DNA is synthesized is called S-phase. The interphase (resting) stage between mitosis and the next round of DNA replication is called G1. There may be a G2 stage between the end of DNA replication and the beginning of mitosis.

Eukaryotic DNA replication origins must be used once, and only once, during S-phase of each cell cycle. We are beginning to understand some of the key players that orchestrate this process. At the end of the previous M-phase and during the subsequent G1-phase, each functional *ori* becomes an assembly site for a conserved multiprotein complex named ORC (origin recognition complex). As the cell progresses through G1 each ORC stimulates the formation of a prereplication complex (pre-RC) that includes a helicase. The pre-RC remains poised until the activity of an S-phase protein kinase (SPK) drops to a critical threshold, whereby the initiation complex recruits waiting replisomes and the origin is said to "fire." The two replication forks are then launched along the chromosome in opposite directions. When SPK activity is high it prevents any new pre-RCs from loading onto the origins, thus preventing multiple rounds of initiation. SPK is proteolytically cleaved at the beginning of the mitotic phase allowing ORC proteins to bind to the origins waiting on each daughter chromosome beginning late in M-phase.

Eukaryotic replication origins do not all fire simultaneously at the beginning of S-phase. Instead, transcribed, or active, regions of a cell's genome tend to be replicated earlier during S-phase while the origins located in quiescent, or repressed, regions of the genome tend to be replicated later in S-phase. It remains to be determined whether this differential timing of replication actually depends on transcription or just reflects that "open" chromatin permits ORC to locate replication origins.

The differences between eukaryotic and prokaryotic DNA replication arise not only from the larger size of the eukaryotic genome but also from the packaging of eukaryotic DNA into chromatin. The binding of DNA to histones and its packaging into nucleosomes, (Section 19.5), is thought to be responsible in part for the slower movement of

Figure 20.19 ►
Photodimerization of adjacent deoxythymidylate residues. Ultraviolet light causes the bases to dimerize, thus distorting the structure of DNA. For clarity, only a single strand of DNA is shown.

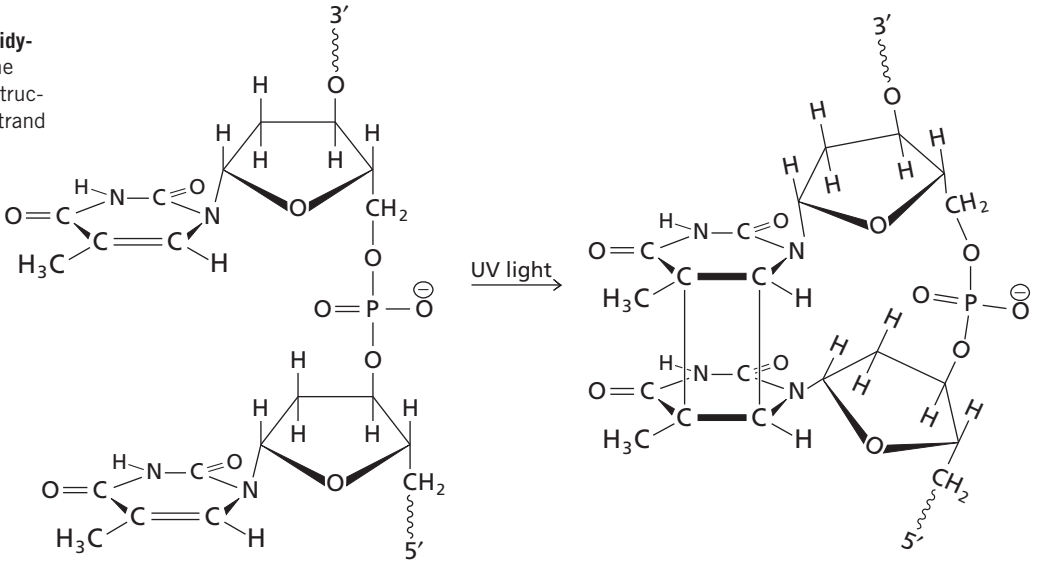
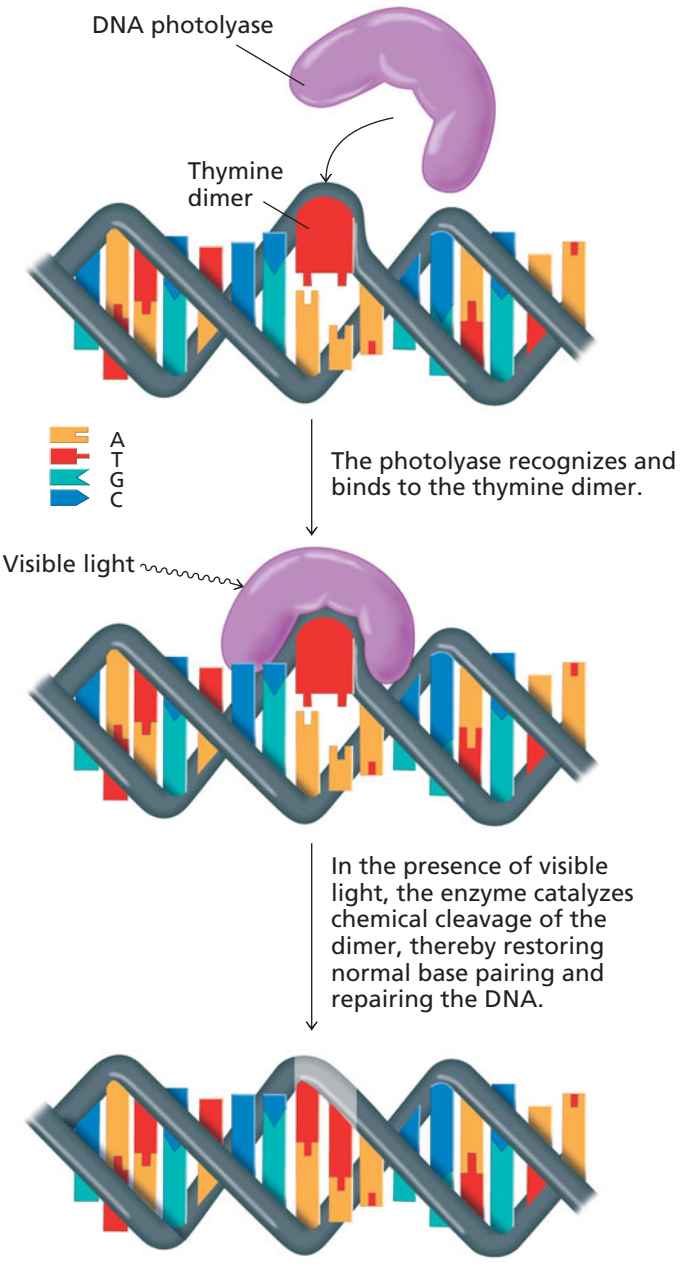


Figure 20.20 ►
Repair of thymine dimers by DNA photolyase.



the replication fork in eukaryotes. Eukaryotic DNA replication occurs with concomitant synthesis of histones; the number of histones doubles with each round of DNA replication. Histone duplication and DNA replication involve different enzymes acting in different parts of the cell yet both occur at about the same rate. It appears that existing histones remain bound to DNA during replication and that newly synthesized histones bind to DNA behind the replication fork shortly after synthesis of the new strands.

20.8 Repair of Damaged DNA

DNA is the only cellular macromolecule that can be repaired. This is probably because the cost to the organism of mutated or damaged DNA far outweighs the energy spent to repair the defect. Repairing other macromolecules is not profitable; for example, little is lost when a defective protein forms as a result of a translation error because the protein is simply replaced by a new, functional protein. When DNA is damaged, however, the entire organism may be in jeopardy if the instructions for synthesizing a critical molecule are altered. In single-celled organisms, damage to a gene encoding an essential protein may kill the organism. Even in multicellular organisms, the accumulation of defects in DNA over time can lead to progressive loss of cellular functions or to deregulated growth such as that seen in cancer cells.

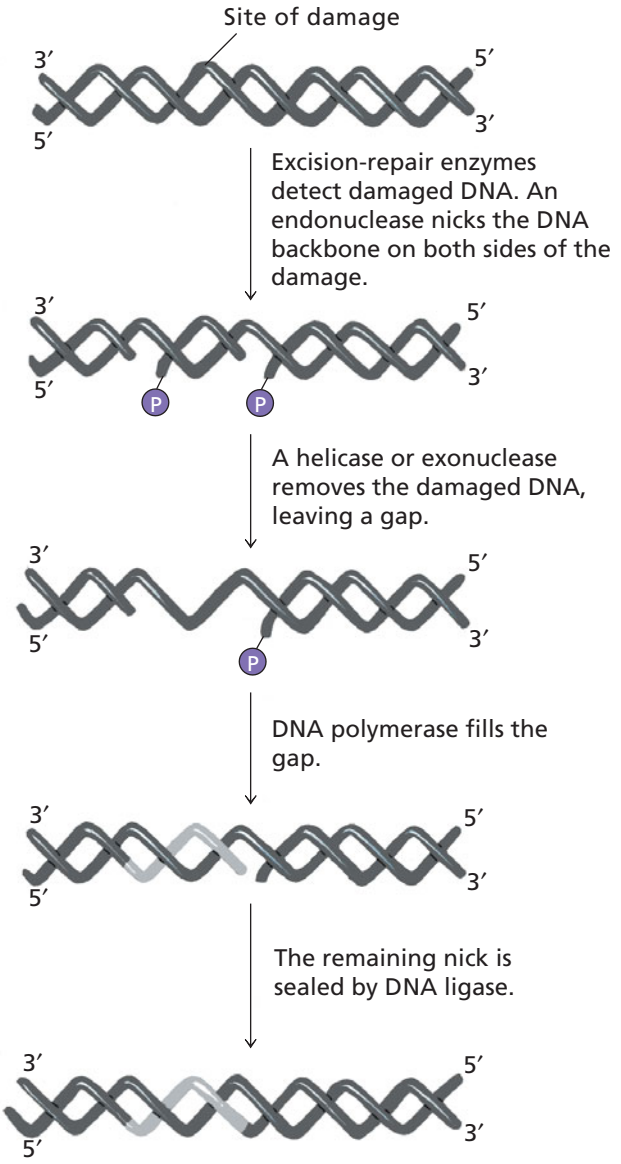
There are several types of DNA damage such as base modifications, nucleotide deletions or insertions, cross-linking of DNA strands, and breakage of the phosphodiester backbone. While some DNA damage is the result of environmental agents (e.g., chemicals or radiation) most DNA damage is the result of errors made during DNA replication. Severe damage may be lethal but much of the damage that occurs *in vivo* is repaired. Many modified nucleotides, as well as mismatched bases that escape the proofreading mechanism of DNA polymerase, are recognized by specific repair enzymes that continually scan DNA in order to detect alterations. Some of the lesions are fixed by **direct repair**, a process that does not require breaking the phosphodiester backbone of DNA. Other repairs require more extensive work.

DNA repair mechanisms protect individual cells as well as subsequent generations. In single-celled organisms, whether prokaryotes or eukaryotes, DNA damage that is not repaired may become a mutation that is passed directly to the daughter cells following DNA replication and cell division. In multicellular organisms, mutations can be passed on to the next generation only if they occur in the germ line. Germ line mutations may have no noticeable effect on the organism that contains them but may have profound effects on the progeny, especially if the mutated genes are important in development. When mutations occur in somatic cells however, while the defects are not transmissible, they can sometimes lead to unrestricted cell growth, or cancer. In spite of the accuracy of DNA replication and the efficiency of repair, the average human accumulates about 130 new mutations every generation. Most of these mutations are neutral and this leads to a huge amount of variation in human populations. It is this variation that makes possible the identification of individuals by DNA fingerprinting.

A. Repair after Photodimerization: An Example of Direct Repair

Double-helical DNA is susceptible to damage by ultraviolet (UV) light. The most common UV light-induced damage is dimerization of adjacent pyrimidines in a DNA strand. This process is an example of photodimerization. The most common dimers form between adjacent thymines (Figure 20.19). DNA replication cannot occur in the presence of pyrimidine dimers because they distort the template strand. Therefore, removal of pyrimidine dimers is essential for survival.

Many organisms can repair thymine dimer damage using direct repair (notably, humans and all placental mammals lack this repair mechanism—see below). The simplest repair process begins when an enzyme known as DNA photolyase binds to the distorted double helix at the site of the thymine dimer (Figure 20.20). As the DNA-enzyme complex absorbs visible light, the dimer is cleared. The photolyase then dissociates from the repaired DNA and normal A/T base pairs re-form. This process is called photo reactivation; it's an example of direct repair.



▲ **Figure 20.21**
General excision-repair pathway.

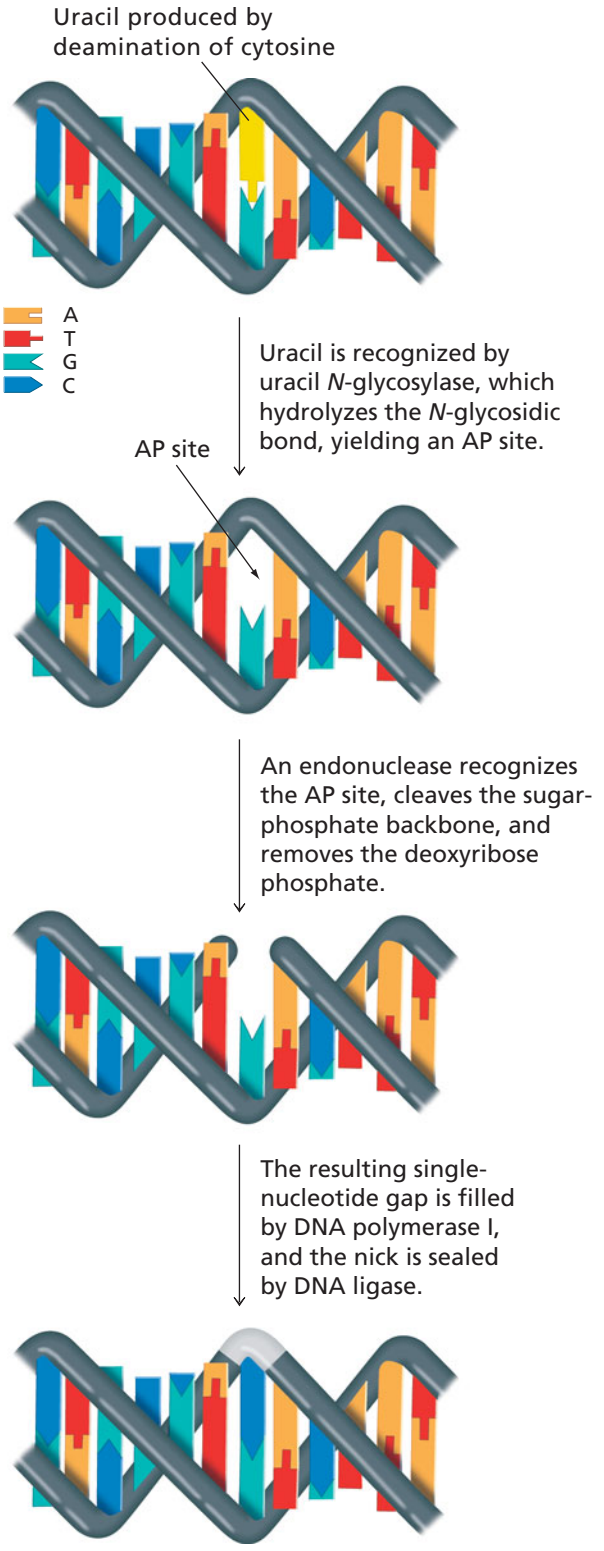
B. Excision Repair

Other forms of ionizing radiation and naturally occurring chemicals can damage DNA. Some compounds, including acids and oxidizing agents, can modify DNA by alkylation, methylation, or deamination. DNA is also susceptible to spontaneous loss of heterocyclic bases, a process known as depurination or depyrimidization. Many of these defects can be repaired by a general **excision repair pathway** whose overall features are similar in all organisms. The pathway begins when an endonuclease recognizes distorted, damaged DNA and cleaves on both sides of the lesion releasing an oligonucleotide containing 12 to 13 residues. This cleavage is catalyzed by the UvrABC enzyme in *E. coli*. Removal of the DNA oligonucleotide may require helicase activity that is often a component of the excision repair enzyme complex. The result is a single-stranded gap. The gap is then filled in by the action of DNA polymerase I in prokaryotes or repair DNA polymerases in eukaryotes. The nick is sealed by DNA ligase (Figure 20.21).

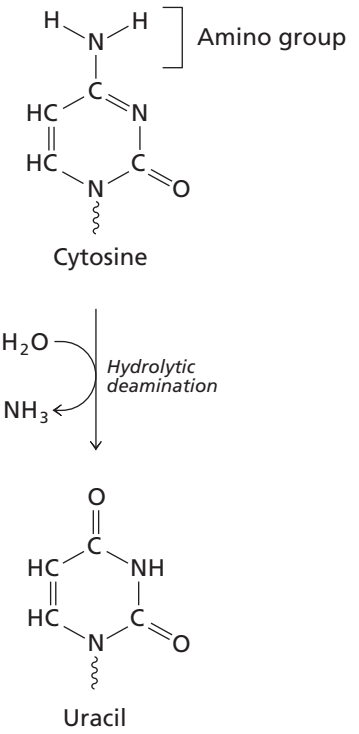
The UvrABC endonuclease also recognizes pyrimidine dimers and modified bases that distort the double helix (this is how thymine dimers are repaired in humans). Other excision-repair enzymes recognize DNA damaged by hydrolytic deamination of adenine, cytosine, or guanine. (Thymine is not subject to deamination because it does not have an amino group.) The deaminated bases can form incorrect base pairs resulting in the incorporation of incorrect bases during the next round of replication. Spontaneous deamination of cytosine is one of the most common types of DNA damage because the product of deamination is uracil that easily forms a base pair with adenine in the next round of replication (Figure 20.22).

Enzymes called DNA glycosylases remove deaminated bases and some other modified bases by catalyzing hydrolysis of the *N*-glycosidic bonds that link the modified bases to the sugars. Let's look at the repair of deaminated cytosine. Repair begins when the enzyme uracil *N*-glycosylase removes the uracil produced by deamination. The enzyme recognizes and binds to the incorrect U/G base pair and flips the uracil base outward, positioning the β-*N*-glycosidic bond in the active site of the enzyme where it is cleaved from the sugar residue (Figure 20.23). Next, an endonuclease recognizes the site where the base is missing and removes the deoxyribose phosphate, leaving a single-nucleotide gap in the duplex DNA. The endonuclease is called an AP-endonuclease because it recognizes apurinic and apyrimidinic sites (AP sites). Some specific DNA glycosylases are bifunctional enzymes with both glycosylase and AP-endonuclease activities in the same polypeptide chain. Excision repair enzymes with exonuclease activity often extend the gap produced by the endonuclease. In prokaryotes, DNA polymerase I binds to the exposed 3′ end of DNA and fills in the gap. Finally, the strand is sealed by DNA ligase. The steps of the excision repair pathway are summarized in Figure 20.24.

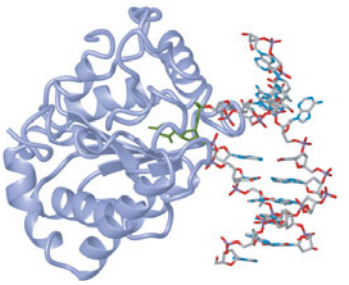
Whereas deamination of adenine or guanine is rare, deamination of cytosine is fairly common and would give rise to large numbers of mutations were it not for the replacement of uracil with thymine in DNA. (Recall that thymine is simply 5-methyluracil.) If uracil were normally found in DNA, as it is in RNA, it would be impossible to distinguish between a correct uridylate residue and one arising from the deamination of cytosine. However, since uracil is not one of the bases in DNA, damage arising from cytosine deamination can be recognized and repaired. Thus, the presence of thymine in DNA increases the stability of genetic information.



▲ **Figure 20.24**
Repair of damage resulting from the deamination of cytosine.



▲ **Figure 20.22**
Hydrolytic deamination of cytosine. Deamination of cytosine produces uracil, which pairs with adenine rather than guanine.



▲ **Figure 20.23**
Uracil *N*-glycosylase from human mitochondria. The enzyme is bound to a uracil-containing nucleotide (green) that has been flipped out of the stacked region of double-stranded DNA. [PDB 1EMH].

BOX 20.1 THE PROBLEM WITH METHYLCYTOSINE

5-Methylcytosine is common in eukaryotic DNA (Section 18.7). Deamination of 5-methylcytosine produces thymidine giving rise to a T opposite a G in damaged DNA. Repair enzymes cannot recognize which of these bases is incorrect, so the "repair" often results in a T:A base pair. This will also happen if the damaged DNA is replicated before it can be repaired. The cytosines at CG sites are preferentially methylated in mammalian genomes. Frequent loss of the cytosines by deamination of 5-methylcytosine has led to underrepresentation of CG sequences relative to TG, AG, and GG.

20.9 Homologous Recombination

Recombination is any event that results in the exchange or transfer of pieces of DNA from one chromosome to another or within a chromosome. Most recombinations are examples of **homologous recombination** because they occur between pieces of DNA that have closely related sequences. Exchanges between paired chromosomes during meiosis are examples of homologous recombination. Recombination between unrelated sequences is called **nonhomologous recombination**. **Transposons** are mobile genetic elements that jump from chromosome to chromosome by taking advantage of nonhomologous recombination mechanisms. Recombination between DNA molecules also occurs when bacteriophages integrate into host chromosomes. When recombination occurs at a specific location it is called **site specific recombination**.

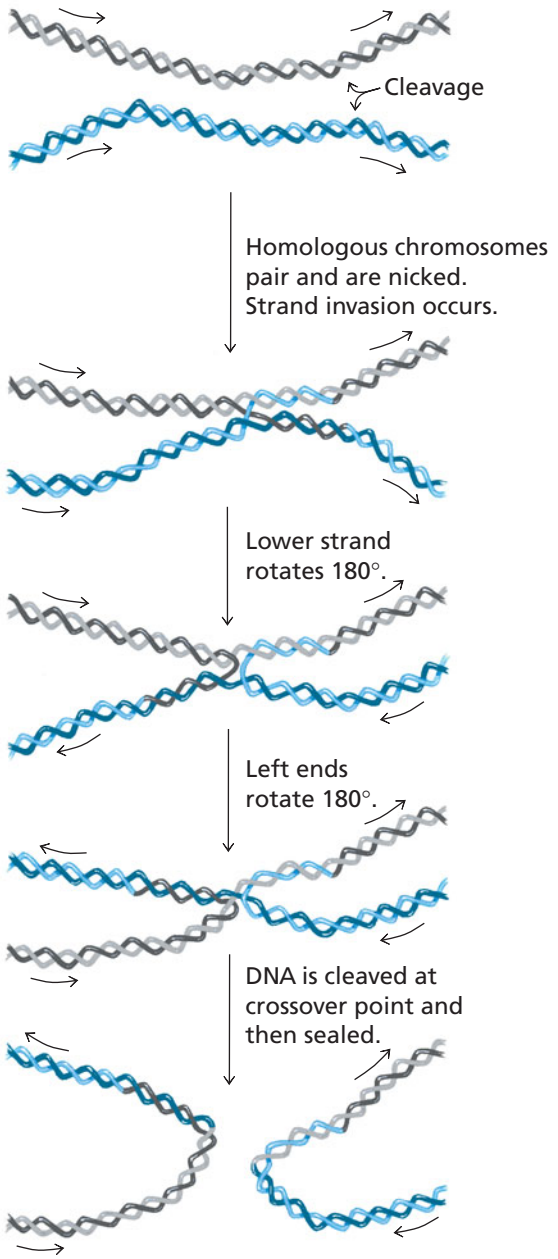
Mutation creates new genetic variation in a population and recombination is a mechanism that creates different combinations of mutations in a genome. Most species have some mechanism for exchanging information between individual organisms. Prokaryotes usually contain only a single copy of their genome (i.e., they are haploid), so this exchange requires recombination. Some eukaryotes are also haploid but most are diploid, having two sets of chromosomes, one contributed by each parent. Genetic recombination in diploids mixes the genes on the chromosomes contributed by each parent so that subsequent generations receive very different combinations of genes. None of your children's chromosomes, for example, will be the same as yours and none of yours are the same as those of your parents. (Although this mixing of alleles is an important consequence of recombination, it is not likely to be the reason why recombination mechanisms evolved in the first place. The problem of why sex evolved is one of the most difficult problems in biology.)

Recombination occurs by many different mechanisms. Many of the proteins and enzymes that participate in recombination reactions are also involved in DNA repair reactions illustrating the close connection between repair and recombination. In this section, we briefly describe the Holliday model of general recombination—a type of recombination that seems to occur in many species.

A. The Holliday Model of General Recombination

Homologous recombination begins with the introduction of either single-stranded or double-stranded breaks into DNA molecules. Recombination involving single-stranded breaks is often called general recombination. Recombination involving double-stranded breaks is not discussed here, although it is an important mechanism of recombination in some species.

Consider general recombination between two linear chromosomes as an example of recombination in prokaryotes. The exchange of information between the molecules begins with the alignment of homologous DNA sequences. Next, single-stranded nicks are introduced in the homologous regions and single strands exchange in a process called strand invasion. The resulting structure contains a region of strand crossover and



◂ **Figure 20.25**
Holliday model of general recombination. Nicks are introduced into a homologous region of each molecule. Subsequent strand invasion, DNA cleavage at the crossover junction, and sealing of nicked strands result in exchange of the ends of the chromosomes.



▲ **Asexual Daphnia**

is known as a Holliday junction after Robin Holliday who first proposed it in 1964 (Figure 20.25).

The chromosomes can be separated at this stage by cleaving the two invading strands at the crossover point. It is important to realize that the ends of the homologous DNA molecules can rotate generating different conformations of the Holliday junction. Rotation followed by cleavage produces two chromosomes that have exchanged ends as shown in Figure 20.25. Recombination in many different organisms probably occurs by a mechanism similar to the one shown in Figure 20.25.



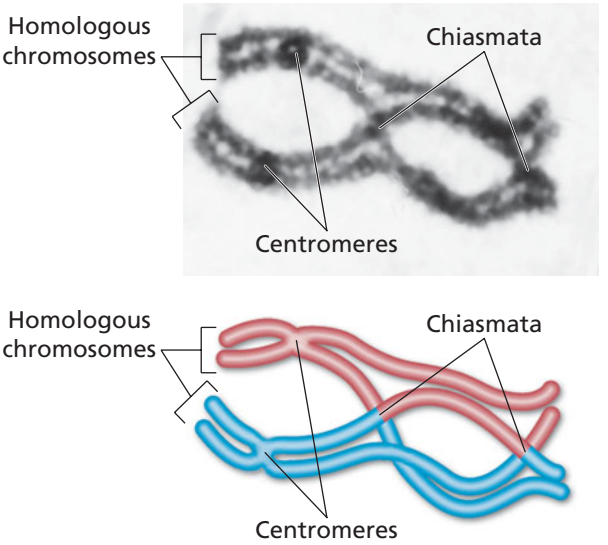
▲ **Male Drosophila melanogaster (no meiotic recombination)**

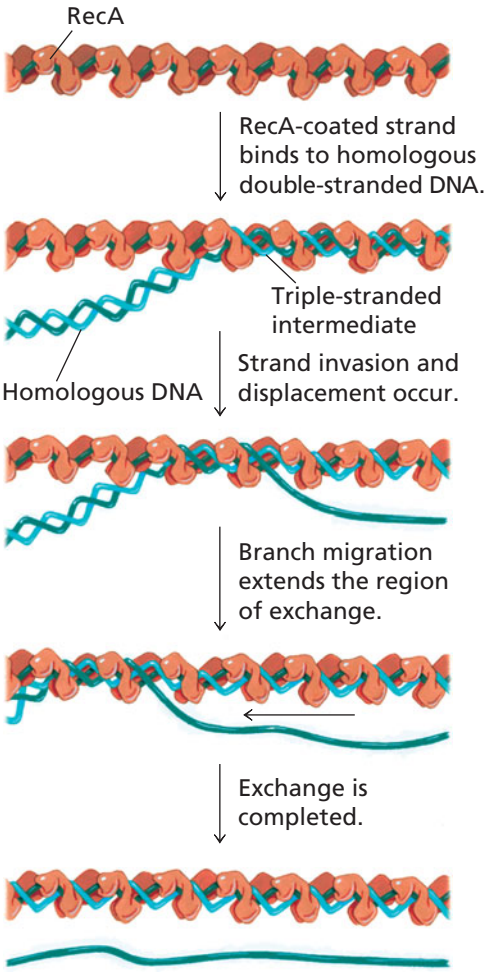
B. Recombination in *E. coli*

One of the first steps in recombination is the generation of single-stranded DNA with a free 3′ end. In *E. coli*, this step is carried out by RecBCD endonuclease, an enzyme with subunits that are encoded by three genes (*rec*B, *rec*C, and *rec*D) whose products have long been known to play a role in recombination. RecBCD binds to DNA and cleaves

Meiotic chisasmata ►

Source: © 2008 Sinauer Associates Sadava, D. et al. *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates and W. H. Freeman & Company), 198





▲ **Figure 20.26**
Strand exchange catalyzed by RecA.

one of the strands. It then unwinds the DNA in a process coupled to ATP hydrolysis generating single-stranded DNA with a 3′ terminus.

Strand exchange during recombination begins when the single-stranded DNA invades the double helix of a neighboring DNA molecule. Strand exchange is not a thermodynamically favorable event—the invasion must be assisted by proteins that promote recombination and repair. RecA is the prototypical strand exchange protein. It is essential for homologous recombination and for some forms of repair. The protein functions as a monomer that binds cooperatively to single-stranded DNA such as the single-stranded tails produced by the action of RecBCD. Each RecA monomer covers about five nucleotide residues and each successive monomer binds to the opposite side of the DNA strand.

One of the key roles of RecA in recombination is to recognize regions of sequence similarity. RecA promotes the formation of a triple-stranded intermediate between the RecA-coated single strand and a highly similar region of double-stranded DNA. RecA then catalyzes strand exchange in which the single strand displaces the corresponding strand from the double helix.

Strand exchange takes place in two steps: strand invasion, followed by branch migration (Figure 20.26). Both the single-stranded and the double-stranded DNA are in an extended conformation during the exchange reaction. The strands must rotate around each other, a process that is presumably aided by topoisomerases. Strand exchange is a slow process despite the fact that no covalent bonds are broken. (A "slow" process in biochemistry is one that takes several minutes.)

RecA can also promote strand invasion between two aligned, double-stranded DNA molecules. Both molecules must contain single-stranded tails bound to RecA. The tails wind around the corresponding complementary strands in the homologue. This exchange gives rise to a Holliday junction such as the one shown in Figure 20.25. Subsequent branch migration can extend the region of strand exchange. Branch migration can continue even after RecA dissociates from the recombination intermediate.

Branch migration at the double-stranded version of a Holliday junction is driven by a remarkable protein machine found in all species. The bacterial version is made up of RuvA and RuvB subunits. These proteins bind to the junction and



◄ **Bacterial conjugation (or sex).**

promote branch migration as shown in the schematic diagram (Figure 20.27). The two DNA helices are separated when RuvC binds to the Holliday junction and cleaves the crossover strands.

RuvA and RuvB form a complex consisting of four RuvA subunits bound to the Holliday junction and two hexameric rings of RuvB subunits that surround two of the DNA strands (Figure 20.28). The RuvB component is similar to the sliding clamps discussed in the section on DNA replication (Section 20.2B) and it drives branch migration by pulling the strands through the RuvA/Holliday junction complex in a reaction coupled to ATP hydrolysis (Figure 20.29). The rate of RuvAB-mediated branch migration is about 100, 000 bp per second—significantly faster than strand invasion.

RuvC catalyzes cleavage of the crossover strands to resolve Holliday junctions. Two types of recombinant molecules are produced as a result of this cleavage: those in which only single strands are exchanged and those in which the ends of the chromosome have been swapped (Figure 20.25).

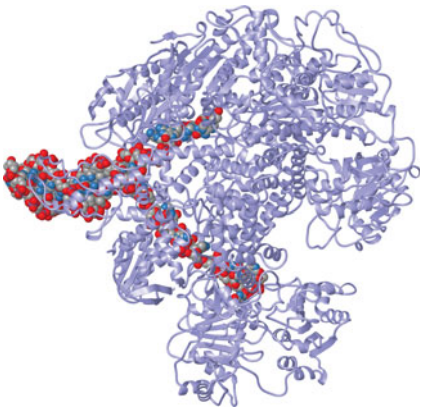
C. Recombination Can Be a Form of Repair

Since natural selection works predominantly at the level of individual organisms it is difficult to see why recombination would have evolved unless it affected survival of the individual. Recombination enzymes probably evolved because they play a role in DNA repair, which confers a selective advantage. For example, severe lesions in DNA are bypassed during DNA replication, leaving a daughter strand with a single-stranded region. RecA-mediated strand exchange between the homologous daughter chromosomes allows the intact strand from one daughter molecule to act as a template for repairing the broken strand of the other daughter molecule.

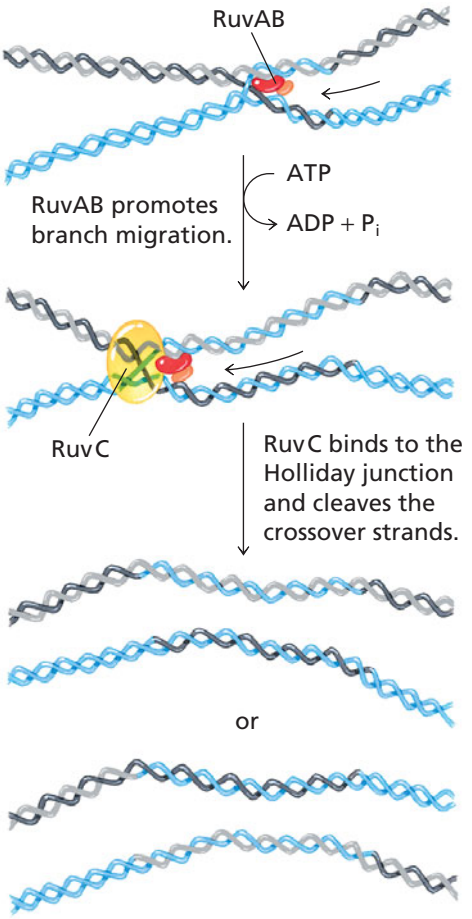
Recombination also creates new combinations of genes on a chromosome and this may be an added bonus for the population and its chances for evolutionary survival. More than 100 *E. coli* genes are required for recombination and repair, and there are twice as many in most eukaryotes.

Most, if not all, of the genes used in recombination play some role in repair as well. Mutations in several human genes give rise to rare genetic defects that result from deficiencies in DNA repair and/or recombination. For example, xeroderma pigmentosum is a hereditary disease associated with extreme sensitivity to ultraviolet light and increased frequency of skin cancer. Excision repair is defective in patients with this disease but the phenotype can be due to mutations in at least eight different genes. One of these genes encodes a DNA glycosylase with AP-endonuclease activity. Other affected genes include some that encode helicases that are required for both repair and recombination.

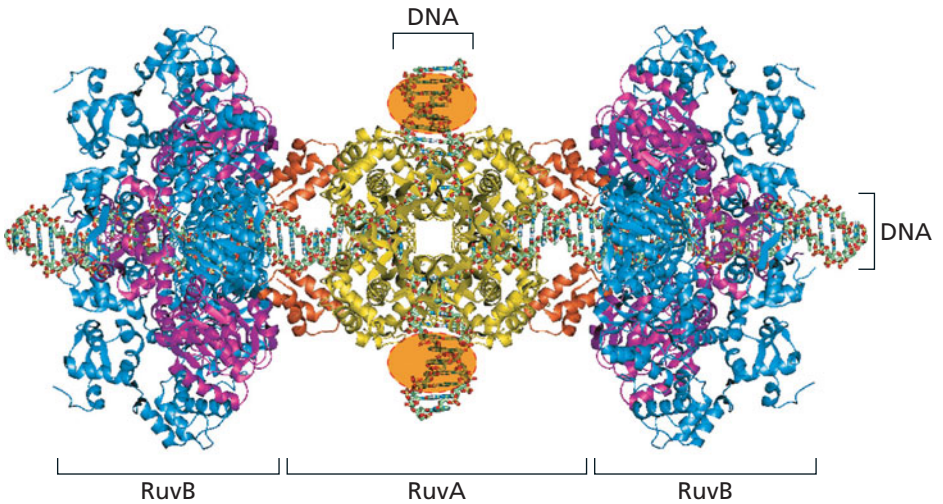
Many other genetic defects related to deficiencies in repair and recombination have not been well characterized. Some of them are responsible for increased incidences of cancer in affected patients.



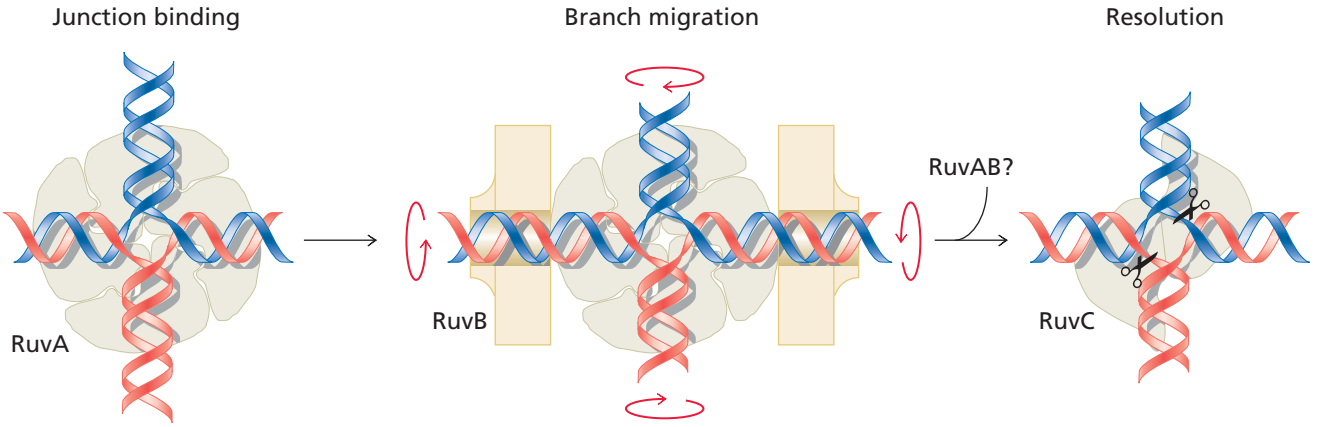
▲ **RecBCD bound to DNA showing separation of strands.** [PDB 3K70]



▲ **Figure 20.27**
Action of Ruv proteins at Holliday junctions. RuvAB promotes branch migration in a reaction coupled to ATP hydrolysis. RuvC cleaves Holliday junctions. Two types of recombinant molecules can be generated in this reaction.



◀ **Figure 20.28**
Model of RuvA and RuvB bound to a Holliday junction.

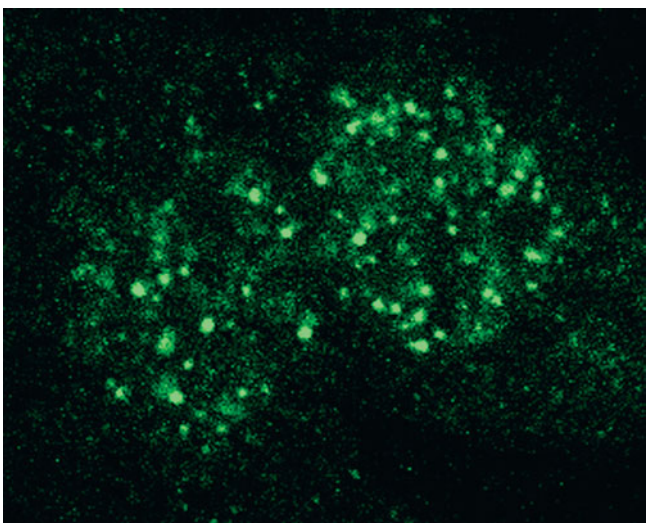


▲ **Figure 20.29**
Branch migration and resolution. [Adapted from Rafferty, J. B., et al. (1996). Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science* 274:415–421.]

BOX 20.2 MOLECULAR LINKS BETWEEN DNA REPAIR AND BREAST CANCER

About 180,000 women are diagnosed with breast cancer every year in North America. Approximately one-fifth of these new cases have a familial or genetic component and one-third of these, or 12,000 cases, are due to mutations in one of the two genes named *BRCA1* or *BRCA2* that encode proteins by the same name.

Both of these proteins are required for normal recombinational repair of double strand breaks (DSB). BRCA2 forms a complex with the eukaryotic RecA homologue RAD51. BRCA2 also binds specifically to BRCA1 to form a heterotrimer. Following exposure to ionizing radiation, these three DNA repair proteins are found localized to discrete sites, or foci, inside the interphase nuclei (see figure). These foci are the sites where the proteins are repairing double strand breaks. The BRCA proteins are so vital that cells become susceptible to damage if either copy of the gene is damaged. When one or both copies of the *BRCA1* or *BRCA2* genes are defective, the capacity to repair DSBs is compromised leading eventually to a higher frequency of mutations. Some of these new mutations may allow the cell to escape from the rigorous constraints imposed by the eukaryotic cell cycle, eventually leading to cancer. The BRCA proteins function as sentinels by continually monitoring the genome to identify and correct potential mutagenic lesions. In fact, some humans with a rare autosomal recessive disease called Fanconi's Anemia (FA) have an increased sensitivity to several mutagenic compounds and a genetic predisposition to many different types of cancers. It has been shown that FA patients are affected in one of seven different genes that are presumably important for DNA repair. One of these genes is *BRCA2*, underscoring its essential role in the repair process.



▲ Ionizing radiation induces nuclear foci of the DNA repair protein BRCA1. Energetic γ-rays can induce double-stranded breaks in DNA and trigger DNA repair. This tissue culture cell nucleus was exposed to IR and then treated with antibodies that recognize BRCA1 (stained green).

Summary

1. DNA replication is semiconservative; each strand of DNA serves as the template for synthesis of a complementary strand. The products of replication are two double-stranded daughter molecules consisting of one parental strand and one newly synthesized strand. DNA replication is bidirectional, proceeding in both directions from an origin in replication.
2. DNA polymerases add nucleotides to a growing DNA chain by catalyzing nucleotidyl-group–transfer reactions. DNA synthesis proceeds in the 5′ → 3′ direction. Errors in DNA synthesis are removed by the 3′ → 5′ exonuclease activity of the polymerase. Some DNA polymerases contain an additional 5′ → 3′ exonuclease activity.
3. The leading strand of DNA is synthesized continuously but the lagging strand is synthesized discontinuously producing Okazaki fragments. Synthesis of the leading strand and of each Okazaki fragment begins with an RNA primer. In *E. coli*, the primer is removed and replaced with DNA by the action of DNA polymerase I. The action of DNA ligase joins the separate fragments of the lagging strand.
4. The replisome is a complex protein machine that is assembled at the replication fork. The replisome contains two DNA polymerase molecules plus additional proteins such as helicase and primase.
5. Assembly of the replisome ensures simultaneous synthesis of two strands of DNA. In *E. coli*, a helicase unwinds the parental DNA and SSB binds to the single strands. The lagging-strand template is looped through the replisome so that the synthesis of both strands proceeds in the same direction as replication fork movement. Because it is part of the replisome, DNA polymerase is highly processive.
6. Initiation of DNA replication occurs at specific DNA sequences (e.g., *oriC* in *E. coli*) and depends on the presence of additional proteins. In bacteria, termination of DNA replication also occurs at specific sites and requires additional proteins.
7. Several new technologies such as PCR and DNA sequencing are based on an understanding of DNA replication.
8. Eukaryotic DNA replication resembles prokaryotic DNA replication except that eukaryotic chromosomes contain multiple origins of replication and eukaryotic Okazaki fragments are smaller. The slower movement of the replication fork in eukaryotes than in prokaryotes is due to the presence of nucleosomes.
9. DNA damaged by radiation or chemical agents can be repaired by direct-repair mechanisms or by a general excision-repair pathway. Excision-repair mechanisms also remove misincorporated nucleotides. Specific enzymes recognize damaged or misincorporated nucleotides.
10. Recombination can occur when a single strand of DNA exchanges with a homologous strand in double-stranded DNA producing a Holliday junction. Strand invasion is promoted by RecA in *E. coli*. Branch migration and resolution of Holliday junctions are catalyzed by RuvABC in *E. coli*.
11. Repair and recombination are similar processes and use many of the same enzymes. Defects in human genes required for repair and recombination cause sensitivity to ultraviolet light and increased risks of cancer.

Problems

1. The chromosome of a certain bacterium is a circular, double-stranded DNA molecule of 5.2×10^6 base pairs. The chromosome contains one origin of replication and the rate of replication-fork movement is 1000 nucleotides per second.
   (a) Calculate the time required to replicate the chromosome.
   (b) Explain how the bacterial generation time can be as short as 25 minutes under extremely favorable conditions.
2. In many DNA viruses the viral genes can be divided into two nonoverlapping groups: early genes, whose products can be detected prior to replication of the viral genome; and late genes, whose products accumulate in the infected cell after replication of the viral genome. Some viruses, like bacteriophage T4 and T7, encode their own DNA polymerase enzymes. Would you expect the gene for T4 DNA polymerase to be in the early or late class? Why?
3. (a) Why does the addition of SSB to sequencing reactions often increase the yield of DNA?
   (b) What is the advantage of carrying out sequencing reactions at 65°C using a DNA polymerase isolated from bacteria that grow at high temperatures?
4. How does the use of an RNA primer rather than a DNA primer affect the fidelity of DNA replication in *E. coli*?
5. Both strands of DNA are synthesized in the 5′ → 3′ direction.
   (a) Draw a hypothetical reaction mechanism for synthesis of DNA in the 3′ → 5′ direction using a 5′-dNTP and a growing chain with a 5′-triphosphate group.
   (b) How would DNA synthesis be affected if the hypothetical enzyme had proofreading activity?
6. Ciprofloxacin is an antimicrobial used in the treatment of a wide variety of bacterial infections. One of the targets of ciprofloxacin in *E. coli* is topoisomerase II. Explain why the inhibition of topoisomerase II is an effective target to treat infections by *E. coli*.
7. The entire genome of the fruit fly *D. melanogaster* consists of 1.65×10^8 bp. If replication at a single replication fork occurs at the rate of 30 bp per second, calculate the minimum time required to replicate the entire genome if replication were initiated
   (a) at a single bidirectional origin
   (b) at 2000 bidirectional origins
   (c) In the early embryo, replication can require as few as 5 minutes. What is the minimum number of origins necessary to account for this replication time?
8. Ethyl methane sulfonate (EMS) is a reactive alkylating agent that ethylates the O-6 residue of guanine in DNA. If this modified G is not excised and replaced with a normal G, what would be the outcome of one round of DNA replication?
9. Why do cells exposed to visible light following irradiation with ultraviolet light have a greater survival rate than cells kept in the dark after irradiation with ultraviolet light?
10. *E. coli* uses several mechanisms to prevent the incorporation of the base uracil into DNA. First, the enzyme dUTPase, encoded by the *dut* gene, degrades dUTP. Second, the enzyme uracil *N*-glycosylase,

encoded by the *ung* gene, removes uracils that have found their way into DNA. The resulting apyrimidinic sites have to be repaired.

(a) If we examine the DNA from a strain carrying a mutation in the *dut* gene, what will we find?

(b) What if we examine the DNA from a strain in which both the *dut* and *ung* genes are mutated?

11. Explain why uracil *N*-glycosylase cannot repair the damage when 5-methylcytosine is deaminated to thymine.

12. Why are high rates of mutation observed in regions of DNA that contain methylcytosine?

13. Explain why the overall error rate for DNA replication in *E. coli* is approximately 10^{-9} although the rate of misincorporation by the replisome is about 10^{-5}.

14. Will DNA repair in *E. coli* be dependent on the enzymatic cofactor $NAD^{\oplus}$?

15. Describe two methods that can be used to repair pyrimidine dimers in *E. coli*.

16. Damage to a single strand of DNA is readily repaired through a variety of mechanisms while damage to bases on both strands of DNA is more difficult for the cell to repair. Explain.

17. Why does homologous recombination occur only between DNAs with identical, or almost identical, sequences?

18. Why are two different DNA polymerase enzymes required to replicate the *E. coli* chromosome?

Selected Readings

General

Adams, R. L. P., Knowler, J. T., and Leader, D. P. (1992). *The Biochemistry of the Nucleic Acids*, 11th ed. (New York: Chapman and Hall).

Aladjem, M. I. (2007). Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat. Rev. Genet.* 8:588–600.

Bentley, D. R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456: 53–59.

Kornberg, A., and Baker, T. (1992). *DNA Replication*, 2nd ed. (New York: W. H. Freeman).

DNA Replication

Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993). Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* 260: 352–355.

Bell, S. P. (2002). The origin recognition complex: from simple origins to complex functions. *Genes & Devel.* 16:659–672.

Davey, M. J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002). Motors and switches: AAA + machines within the replisome. *Nat. Rev. Mol. Cell Biol.* 3:1–10.

Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. *Science* 294:96–100.

Keck, J. L., and Berger, J. M. (2001). Primus inter pares (First among equals). *Nat. Struct. Biol.* 8:2–4.

Kong, X.-P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992). Three-dimensional structure of the β subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* 69:425–437.

Kunkel, T. A., and Bebenek, K. (2000). DNA replication fidelity. *Annu. Rev. Biochem.* 69:497–529.

Marians, K. J. (1992). Prokaryotic DNA replication. *Annu. Rev. Biochem.* 61:673–719.

McHenry, C. S. (1991). DNA polymerase III holoenzyme. *J. Biol. Chem.* 266:19127–19130.

Meselson, M., and Stahl, F. W. (1958). The replication of DNA in *Escherichia coli. Proc. Natl. Acad. Sci. USA* 44:671–682.

Radman, M. (1998). DNA replication: one strand may be more equal. *Proc. Natl. Acad. Sci. USA* 95:9718–9719.

Waga, S., and Stillman, B. (1998). The DNA replication fork in eukaryotic cells. *Annu. Rev. Biochem.* 67:721–751.

Wake, R. G., and King, G. F. (1997). A tale of two terminators: crystal structures sharpen the debate on DNA replication fork arrest mechanisms. *Structure* 5:1–5.

Wyman, C., and Botchan, M. (1995). A familiar ring to DNA polymerase processivity. *Curr. Biol.* 5:334–337.

DNA Repair

Echols, H., and Goodman, M. F. (1991). Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.* 60:477–511.

Hanawalt, P. C. and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell. Biol.* 9:958–970.

Kogoma, T. (1997). Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol. Mol. Biol. Rev.* 61:212–238.

McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annu. Rev. Biochem.* 68:255–285.

Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P., and Taylor, J. A. (1999). DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.* 28:101–128.

Tainer, J. A., Thayer, M. M., and Cunningham, R. P. (1995). DNA repair proteins. *Curr. Opin. Struct. Biol.* 5:20–26.

Yang, W. (2000). Structure and function of mismatch repair proteins. *Mutat. Res.* 460:245–256.

Recombination

Ortiz-Lombardia, M., González, A., Ertja, R., Aymami, J., Azorin, F., and Coll, M. (1999). Crystal structure of a Holliday junction. *Nat. Struct. Biol.* 6:913–917.

Rafferty, J. B., Sedelnikove, S. E., Hargreaves, D., Artmiuk, P. J., Baker, P. J., Sharples, G. J., Mahdi, A. A., Lloyd, R. G., and Rice, D. W. (1996). Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science* 274:415–421.

Rao, B. J., Chiu, S. K., Bazemore, L. R., Reddy, G., and Radding, C. M. (1995). How specific is the first recognition step of homologous recombination? *Trends Biochem. Sci.* 20:109–113.

West, S. C. (1996). The RuvABC proteins and Holliday junction processing in *Escherichia coli. J. Bacteriol.* 178:1237–1241.

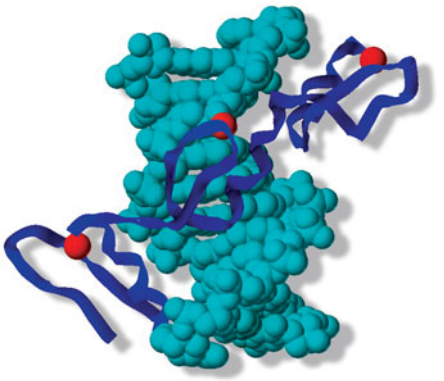
West, S. C. (1997). Processing of recombination intermediates by the RuvABC proteins. *Annu. Rev. Genet.* 31:213–244.

West, S. C. (2003). Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.* 4:1–11.

White, M. F., Giraud-Panis, M.-J. E., Pöhler, J. R. G., and Lilley, D. M. J. (1997). Recognition and manipulation of branched DNA structure by junction-resolving enzymes. *J. Mol. Biol.* 269:647–664.

Wuethrich, B. (1998). Why sex? *Science* 281:1980–1982.





Transcription and RNA Processing

As we have seen, the structure of DNA proposed by Watson and Crick in 1953 immediately suggested a means of replicating DNA to transfer genetic information from one generation to the next but it did not reveal how an organism makes use of the information stored in its genetic material.

Based on studies of the bread mold *Neurospora crassa*, George Beadle and Edward Tatum proposed that a single unit of heredity, or gene, directed the production of a single enzyme. A full demonstration of the relationship between genes and proteins came in 1956 when Vernon Ingram showed that hemoglobin from patients with the heritable disease sickle-cell anemia differed from normal hemoglobin by the replacement of a single amino acid. Ingram's results indicated that genetic changes can manifest themselves as changes in the amino acid sequence of a protein. By extension, the information contained in the genome must specify the primary structure of each protein in an organism.

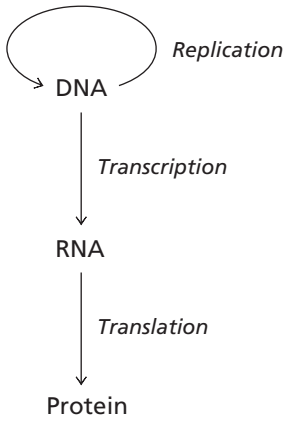
We define a **gene** as a DNA sequence that is transcribed. This definition includes genes that do not encode proteins (not all transcripts are messenger RNA). The definition normally excludes regions of the genome that control transcription but are not themselves transcribed. We will encounter some exceptions to our definition of a gene—surprisingly, there is no definition that is entirely satisfactory.

Many prokaryotic genomes contain several thousand genes, although some simple bacteria have only 500 to 600 genes. Most of these are "housekeeping genes" that encode proteins or RNA molecules that are essential for the normal activities of all living cells. For example, the enzymes involved in the basic metabolic processes of glycolysis and the synthesis of amino acids and DNA are encoded by such housekeeping genes, as are transfer RNAs and ribosomal RNAs. The number of housekeeping genes in unicellular eukaryotes, such as yeast and some algae, is similar to the number in complex prokaryotes.

"This fraction (which we shall designate "messenger RNA" or M-RNA) amounts to only about 3% of the total RNA. . . . The property attributed to the structural messenger of being an unstable intermediate is one of the most specific and novel implications of this scheme. . . . This leads to a new concept of the mechanism of information transfer, where the protein synthesizing centers (ribosomes) play the role of non-specific constituents which can synthesize different proteins, according to specific instructions which they receive from the genes through M-RNA."

—François Jacob and Jacques Monod, 1961

Top: A portion of the mouse transcription factor Zif268 (dark blue) bound to DNA (light blue). Side chains from three zinc-containing domains interact with base pairs in DNA.



▲ **Figure 21.1**
Biological information flow. The normal flow of biological information is from DNA to RNA to protein.

KEY CONCEPT

Before a cell can access the genetic information stored in its DNA, the DNA must be transcribed into RNA.

In addition to housekeeping genes, all cells contain genes that are expressed only in special circumstances, such as during cell division. Multicellular organisms also contain genes that are expressed only in certain types of cells. For example, all cells in a maple tree contain the genes for the enzymes that synthesize chlorophyll but these genes are expressed only in cells that are exposed to light, such as cells on the surface of a leaf. Similarly, all cells in mammals contain insulin genes, but only certain pancreatic cells produce insulin. The total number of genes in multicellular eukaryotes ranges from as few as 15,000 in *Drosophila melanogaster* to more than 50,000 in some other animals.

In this chapter and the next, we will examine how the information stored in DNA directs the synthesis of proteins. A general outline of this flow of information is summarized in Figure 21.1. In this chapter, we describe transcription (the process where information stored in DNA is copied into RNA thereby making it available for either protein synthesis or other cellular functions) and RNA processing (the post-transcriptional modification of RNA molecules). We also briefly examine how gene expression is regulated by factors that affect the initiation of transcription. In Chapter 22, we will examine translation (the process where information coded in mRNA molecules directs the synthesis of individual proteins).

One feature of the complete pathway outlined in Figure 21.1 is that it is irreversible. In particular, the information contained in the amino acid sequence of a protein cannot be translated back into nucleic acid. This irreversibility of information flow is known as the "Central Dogma" of molecular biology and was predicted by Francis Crick in 1958, many years before the mechanisms of transcription and translation were worked out (see Section 1.1). The original version of the Central Dogma did not rule out information flow from RNA to DNA. Such a pathway was eventually discovered in retrovirus-infected cells; it is known as reverse transcription.

21.1 Types of RNA

Several classes of RNA molecules have been discovered. *Transfer RNA* (tRNA) carries amino acids to the translation machinery. *Ribosomal RNA* (rRNA) makes up much of the ribosome. A third class of RNA is *messenger RNA* (mRNA), whose discovery was due largely to the work of François Jacob, Jacques Monod, and their collaborators at the Pasteur Institute in Paris. In the early 1960s, these researchers showed that ribosomes participate in protein synthesis by translating unstable RNA molecules (mRNA). Jacob and Monod also discovered that the sequence of an mRNA molecule is complementary to a segment of one of the strands of DNA. A fourth class of RNA consists of small RNA molecules that participate in various metabolic events, including RNA processing. Many of these small RNA molecules have catalytic activity. Some of these small RNAs are regulatory molecules that can bind specifically to mRNAs and down-regulate that messenger and the protein it encodes.



▲ **François Jacob (1920–).** Jacob and Monod received the Nobel Prize in Physiology or Medicine in 1965 for their work on the genetic control of enzyme synthesis.

A large percentage of the total RNA in a cell is ribosomal RNA, and only a small percentage is mRNA. But if we compare the rates at which the cell synthesizes RNA rather than the steady state levels of RNA, we see a different picture (Table 21.1). Even though mRNA accounts for only 3% of the total RNA in *Escherichia coli*, the bacterium devotes almost one-third of its capacity for RNA synthesis to the production of mRNA. This value may increase to about 60% when the cell is growing slowly and does not need to replace ribosomes and tRNA. The discrepancy between steady state levels of various RNA molecules and the rates at which they are synthesized can be explained by the differing stabilities of the RNA molecules: rRNA and tRNA molecules are extremely stable, whereas mRNA is rapidly degraded after translation. Half of all newly synthesized mRNA is degraded by nucleases within three minutes in bacterial cells. In eukaryotes, the average half-life of mRNA is about ten times longer. The relatively high stability of eukaryotic mRNA results from processing events that prevent eukaryotic mRNA from being degraded during transport from the nucleus, where transcription occurs, to the cytoplasm, where translation occurs.

Table 21.1 The RNA content of an *E. coli* cell

Type	Steady state level	Synthetic type capacity[a]
rRNA	83%	58%
tRNA	14%	10%
mRNA	03%	32%
RNA primers[b]	<1%	<1%
Other RNA molecules[c]	<1%	<1%

[a]Relative amount of each type of RNA being synthesized at any instant.
[b]RNA primers are those used in DNA replication; they are not synthesized by RNA polymerase.
[c]Other RNA molecules include several RNA enzymes, such as the RNA component of RNase P.

[Adapted from Bremer, H., and Dennis, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Vol. 2, F. C. Neidhardt, ed. (Washington, DC: American Society for Microbiology), pp. 1527–1542.]

21.2 RNA Polymerase

About the time that mRNA was identified, researchers in several laboratories independently discovered an enzyme that catalyzes the synthesis of RNA when provided with ATP, UTP, GTP, CTP, and a template DNA molecule. The newly discovered enzyme was RNA polymerase. This enzyme catalyzes DNA-directed RNA synthesis, or **transcription**.

RNA polymerase was initially identified by its ability to catalyze polymerization of ribonucleotides but further study of the enzyme revealed that it does much more. RNA polymerase is the core of a larger transcription complex just as DNA polymerase is the core of a larger replication complex (Section 20.4). This complex assembles at one end of a gene when transcription is initiated. During initiation, the template DNA partially unwinds and a short piece of RNA is synthesized. In the elongation phase of transcription, RNA polymerase catalyzes the processive elongation of the RNA chain while the DNA is continuously unwound and rewound. Finally, the transcription complex responds to specific transcription termination signals and disassembles.

Although the composition of the transcription complex varies considerably among different organisms, all transcription complexes catalyze essentially the same types of reactions. We introduce the general process of transcription by discussing the reactions catalyzed by the well-characterized transcription complex in *E. coli*. The more complicated eukaryotic transcription complexes are presented in Section 21.5.

A. RNA Polymerase Is an Oligomeric Protein

Core RNA polymerase is isolated from *E. coli* cells as a multimeric protein with four different types of subunits (Table 21.2). Five of these subunits combine with a stoichiometry of $\alpha_2\beta\beta'\omega$ to form the core enzyme that participates in many of the transcription reactions. The large β and β' subunits make up the active site of the enzyme; the β' subunit contributes to DNA binding, whereas the β subunit contains part of the polymerase active site. The α subunits are the scaffold for assembly of the other subunits and they also interact with many proteins that regulate transcription. The role of the small ω subunit is not well characterized.

The structure of RNA polymerase holoenzyme from the bacterium *Thermus aquaticus* complexed with DNA is shown in Figure 21.2. The β and β' subunits form a large groove at one end. This is where DNA binds and polymerization takes place. The groove is large enough to accommodate about 16 base pairs of double-stranded B-DNA and is shaped like the DNA-binding sites of DNA polymerases (such as DNA polymerase I; Figure 20.12). The pair of α subunits is located at the "back end" of the molecule. This region also contacts DNA when the polymerase is actively transcribing a gene. The ω subunit is bound to the outer surface of the β' subunit. We will see later that various transcription factors interact with RNA polymerase by binding to the α subunits.

Table 21.2 Subunits of *E. coli* RNA polymerase holoenzyme

Subunit	M_r
β'[a]	155,600
β	150,600
σ[b]	70,300[c]
α	36,500
ω	11,000

[a]The β and β' subunits are unrelated despite the similarity of their names.
[b]This subunit is not part of the core RNA polymerase.
[c]The molecular weight given is for the σ subunit found in the most common form of the holoenzyme.

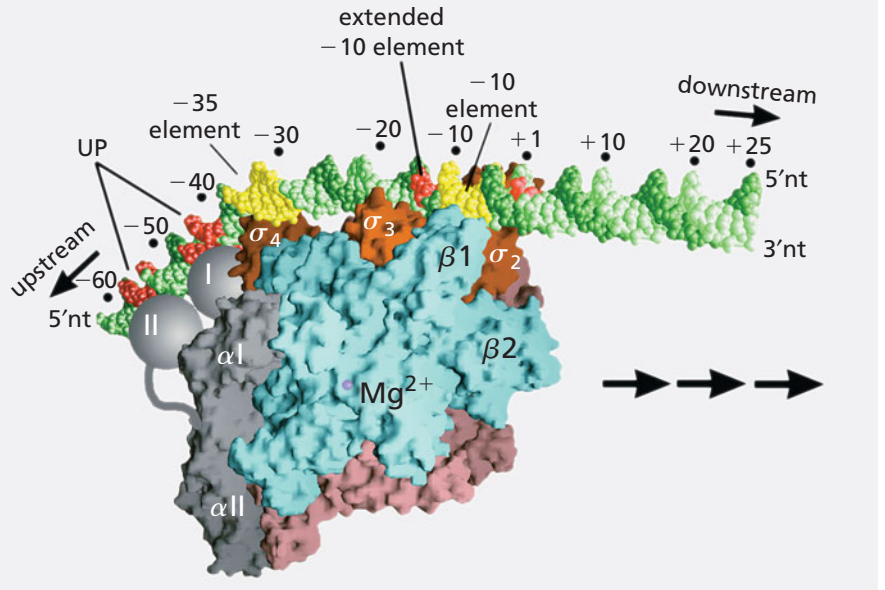


Figure 21.2 ►
***Thermus aquaticus (taq)* RNA polymerase holoenzyme/promoter DNA closed complex.** The template strand is dark green and the coding strand is light green; both the –10 and –35 elements are yellow. The transcription start site is shown in red and labeled +1. Once the open complex forms, then transcription will proceed downstream, to the right as shown by the arrows. The α and ω subunits are shown in gray; the β subunit is cyan, while the β'' subunit is pink. The σ subunit is orange.

The σ subunit of the holoenzyme plays an important role in transcription initiation. Bacteria contain several different types of σ subunits. The major form of the holoenzyme in *E. coli* contains the subunit σ^{70} (M_r 70,300). The σ subunits contact DNA during transcription initiation and bind to the core enzyme in the region of the ω subunit. The overall dimensions of RNA polymerase are 10 × 10 × 16 nm. This makes it considerably larger than a nucleosome but smaller than a ribosome or a replisome.

B. The Chain Elongation Reaction

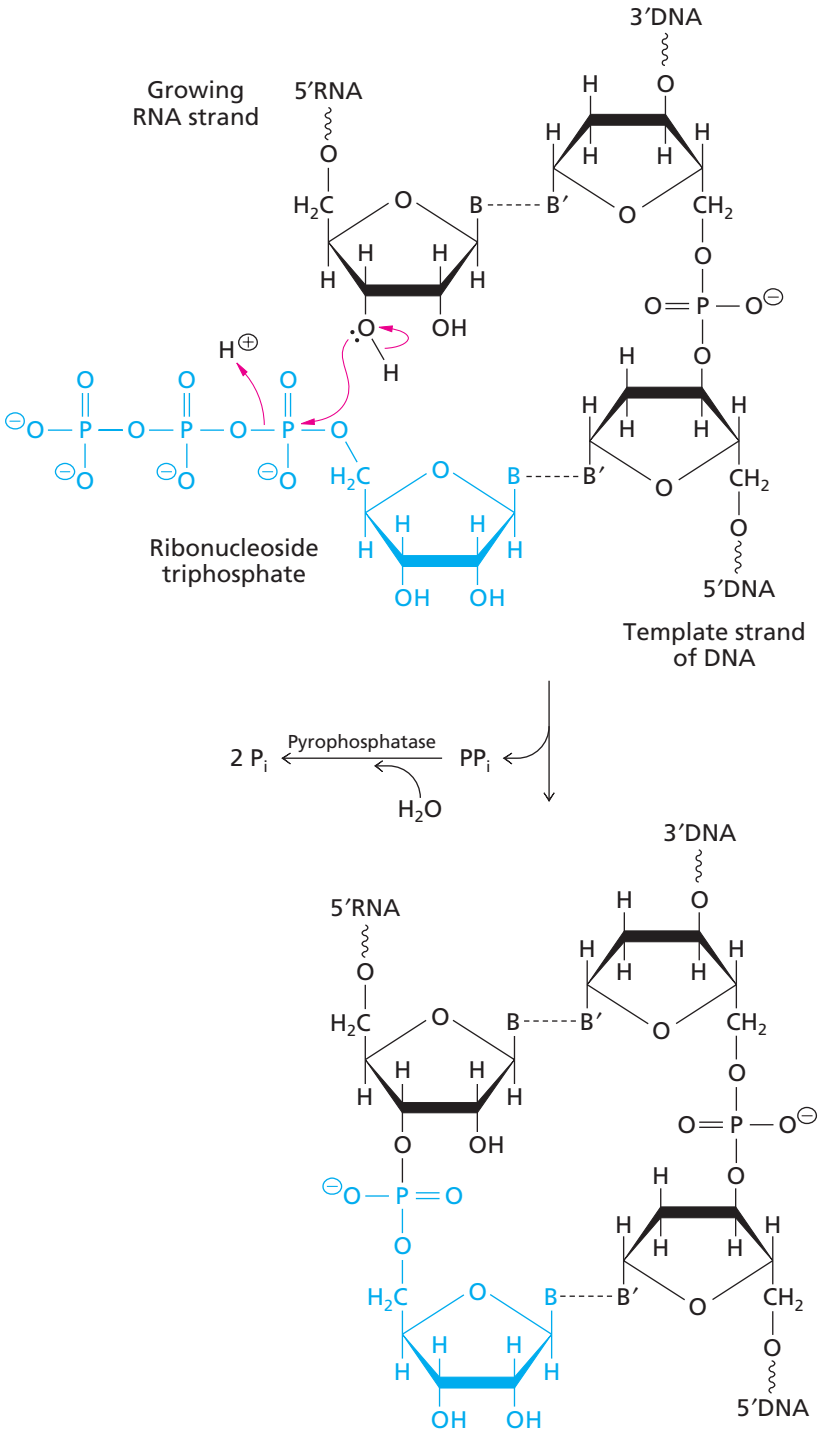
RNA polymerase catalyzes chain elongation by a mechanism almost identical to that used by DNA polymerase (Figure 20.6). Part of the growing RNA chain is base-paired to the DNA template strand, and incoming ribonucleoside triphosphates are tested in the active site of the polymerase for correct hydrogen bonding to the next unpaired nucleotide of the template strand. When the incoming nucleotide forms correct hydrogen bonds, RNA polymerase catalyzes a nucleotidyl-group–transfer reaction, resulting in formation of a new phosphodiester linkage and the release of pyrophosphate (Figure 21.3).

Like DNA polymerase III, RNA polymerase catalyzes polymerization in the 5′ ⟶ 3′ direction and is highly processive when it is bound to DNA as part of a transcription complex. The overall reaction of RNA synthesis can be summarized as

$$\text{RNA}_n - \text{OH} + \text{NTP} \longrightarrow \text{RNA}_{n+1} - \text{OH} + \text{PP}_i \qquad \textbf{(21.1)}$$

The Gibbs free energy change for this reaction is highly favorable because of the high concentration of NTPs relative to RNA. In addition, the RNA polymerase reaction like the DNA polymerase reaction is thermodynamically assisted by the subsequent hydrolysis of pyrophosphate inside the cell. Thus, two phosphoanhydride linkages are expended for every nucleotide added to the growing chain.

RNA polymerase differs from DNA polymerase in using ribonucleoside triphosphates (UTP, GTP, ATP, and CTP) as substrates rather than deoxyribonucleoside triphosphates (dTTP, dGTP, dATP, and dCTP). Another difference is that the growing RNA strand only interacts with the template strand over a short distance (see below). The final product of transcription is single-stranded RNA, not an RNA-DNA duplex. Transcription is much slower than DNA replication. In *E. coli*, the rate of transcription ranges from 30 to 85 nucleotides per second, or less than one-tenth the rate of DNA replication.



◂ **Figure 21.3**
Reaction catalyzed by RNA polymerase. When an incoming ribonucleoside triphosphate correctly pairs with the next unpaired nucleotide on the DNA template strand, RNA polymerase catalyzes a nucleophilic attack by the 3′-hydroxyl group of the growing RNA strand on the α-phosphorus atom of the incoming ribonucleoside triphosphate. As a result, a phosphodiester forms and pyrophosphate is released. The subsequent hydrolysis of pyrophosphate catalyzed by pyrophosphatase provides additional thermodynamic driving force for the reaction. (B and B′ represent complementary bases, and hydrogen bonding between bases is indicated by a single dashed line.)

RNA polymerase catalyzes the formation of a new phosphodiester linkage only when the incoming ribonucleoside triphosphate fits the active site of the enzyme precisely. A precise fit requires base stacking and appropriate hydrogen bonding between the incoming ribonucleoside triphosphate and the template nucleotide.

Despite the requirement for an accurate fit, RNA polymerase does make mistakes. The error rate of RNA synthesis is 10^{-6} (one mistake for every 1 million nucleotides incorporated). This rate is higher than the overall error rate of DNA synthesis because, in contrast to most DNA polymerases, RNA polymerase does not possess an exonuclease proofreading activity. Extreme precision in DNA replication is necessary to minimize mutations that could be passed on to progeny but accuracy in RNA synthesis is not as crucial to survival.

21.3 Transcription Initiation

The elongation reactions of RNA synthesis are preceded by a distinct initiation step in which a transcription complex assembles at an initiation site and a short stretch of RNA is synthesized. The regions of DNA that serve as sites of transcription initiation are called **promoters**. In bacteria, several genes are often co-transcribed from a single promoter; such a transcription unit is called an **operon**. In eukaryotic cells, each gene usually has its own promoter. There are hundreds of promoters in bacterial cells and thousands in eukaryotic cells.

The frequency of transcription initiation at any given promoter is usually related to the need for that gene's particular product. For example, in cells that are dividing rapidly, the genes for ribosomal RNA are usually transcribed frequently. Every few seconds a new transcription complex begins transcribing at the promoter. This process gives rise to structures such as those seen in Figure 21.4 showing multiple transcription complexes on one *E. coli* ribosomal RNA operon. Transcripts of increasing length are arrayed along the genes because many RNA polymerases transcribe the genes at the same time. In contrast, some bacterial genes are transcribed only once every two generations. In these cases initiation may occur only once every few hours. (Outside of the laboratory, the average generation time of most bacteria is many hours.)

A. Genes Have a 5′ → 3′ Orientation

In Section 19.2A, we introduced the convention that single-strand nucleic acid sequences are written from left to right in the 5′ → 3′ direction. When a sequence of double-stranded DNA is displayed, the sequence of the top strand is written 5′ → 3′ and the sequence of the bottom, antiparallel, strand is written 3′ → 5′ (left to right).

Since our operational definition of a gene is a DNA sequence that is transcribed, a gene begins at the point where transcription starts (designated +1) and ends at the point where transcription terminates. The beginning of a gene is called the 5′ end, corresponding to the convention for writing sequences. Moving along a gene in the 5′ → 3′ direction is described as moving "downstream" and moving in the 3′ → 5′ direction is moving "upstream." RNA polymerization proceeds in the 5′ → 3′ direction. Consequently, in accordance with the convention for writing DNA sequences, the transcription start site of a gene is shown on the left of a diagram of double-stranded DNA and the termination site is on the right. The top strand is often called the **coding strand** because its sequence corresponds to the DNA version of the mRNA that encodes the amino acid sequence of a protein. The bottom strand is called the **template strand** because it is the strand used as a template for RNA synthesis (Figure 21.5). Alternatively, the top strand may be called the **sense strand** to indicate that translating ribosomes attempting to "read" the codons in an mRNA with this sequence will make the correct protein. Therefore the bottom strand becomes the **antisense strand** because an mRNA with this sequence will not make the correct protein. Note that RNA is synthesized in

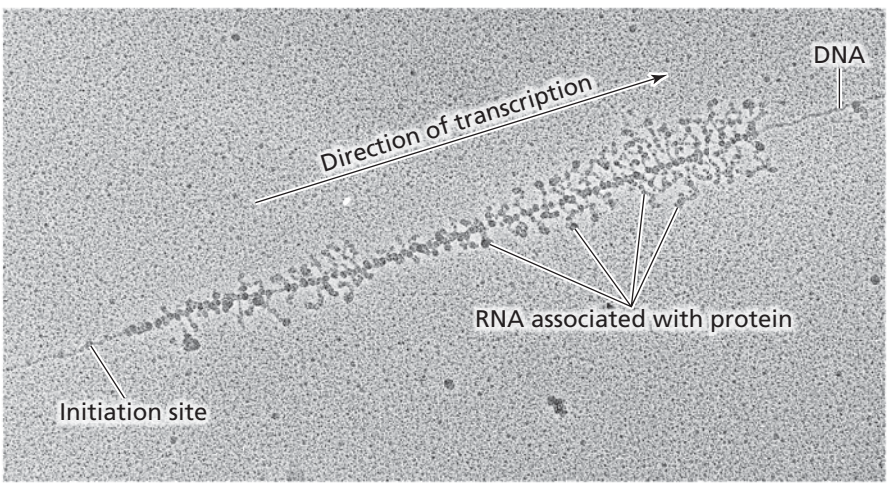
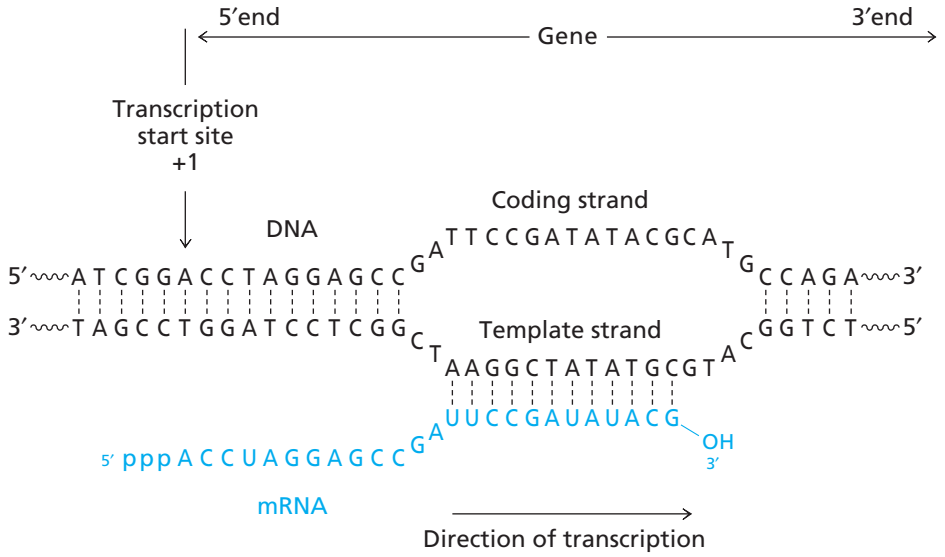


Figure 21.4 ▸
Transcription of *E. coli* ribosomal RNA genes. The genes are being transcribed from left to right. The nascent rRNA product associates with proteins and is processed by nucleolytic cleavage before transcription is complete.



◀ **Figure 21.5**
Orientation of a gene. The sequence of a hypothetical gene and the RNA transcribed from it are shown. By convention, the gene is said to be transcribed from the 5′ end to the 3′ end but the template strand of DNA is copied from the 3′ end to the 5′ end. Growth of the ribonucleotide chain proceeds 5′ → 3′.

the 5′ → 3′ direction but the template strand is copied from its 3′ end to its 5′ end. Also note that the RNA product is identical in sequence to the coding strand except that U replaces T.

B. The Transcription Complex Assembles at a Promoter

A transcription complex forms when one or more proteins bind to the promoter sequence and also to RNA polymerase. These DNA-binding proteins direct RNA polymerase to the promoter site. In bacteria, the σ subunit of RNA polymerase is required for promoter recognition and formation of the transcription complex.

The nucleotide sequence of a promoter is one of the most important factors affecting the frequency of transcription of a gene. Soon after the development of DNA-sequencing technology, many different promoters were examined. The start sites, the points at which transcription actually begins, were identified, and the regions upstream of these sites were sequenced to learn whether the promoter sequences of different genes were similar. This analysis revealed a common pattern called a **consensus sequence**—a hypothetical sequence made up of the nucleotides found most often in each position.

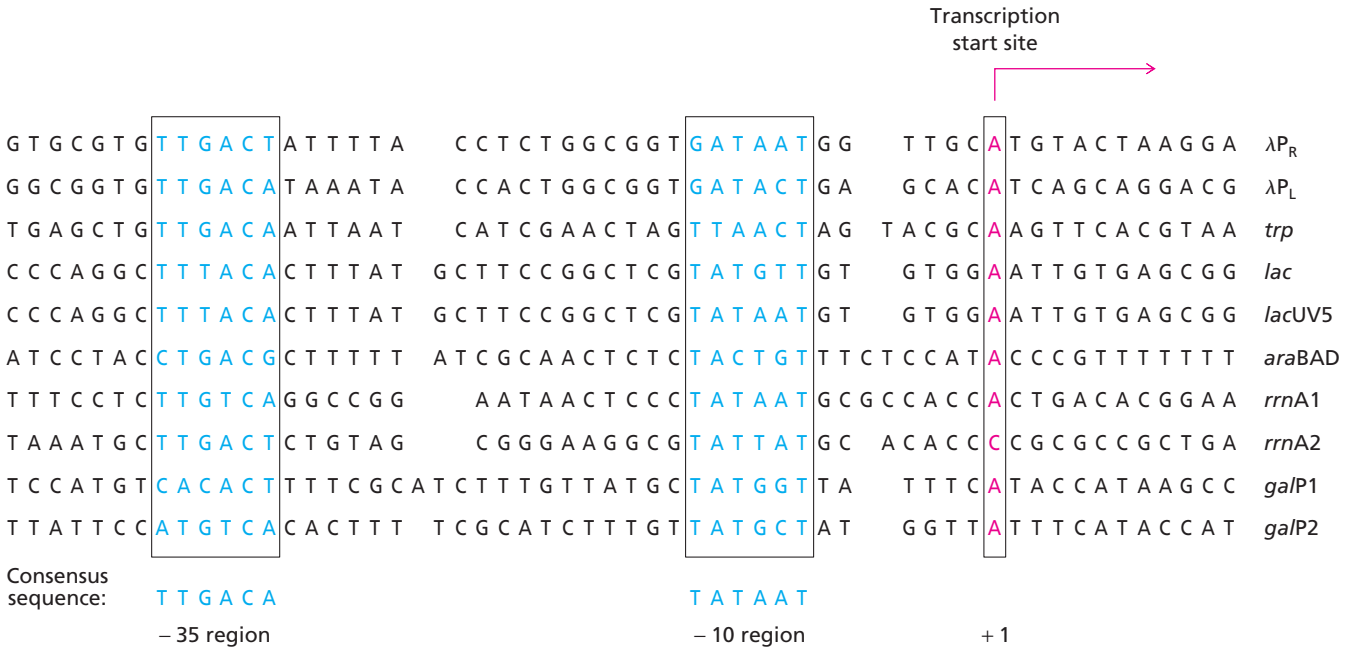
The consensus sequence of the most common type of promoter in *E. coli* is shown in Figure 21.6. This promoter is bipartite, which means that there are two separate regions of sequence similarity. The first region is 10 bp upstream of the transcription start site and is rich in A/T base pairs. The consensus sequence is TATAAT. The second part of the promoter sequence is centered approximately 35 bp upstream of the start site. The consensus sequence in this region is TTGACA. The average distance between the two parts of the promoter is 17 bp.

The −10 region is known as a **TATA box**, and the −35 region is simply referred to as the **–35 region**. Together, the two regions define the promoter for the *E. coli* holoenzyme containing σ^{70}, the most common σ subunit in *E. coli* cells. The σ^{70}-containing holoenzyme binds specifically to sequences that resemble the consensus sequence. Other *E. coli* σ subunits recognize and bind to promoters with quite different consensus sequences (Table 21.3). Orthologous σ subunits from other prokaryotic species may recognize different promoter consensus sequences.

A consensus sequence is not an exact sequence but indicates the nucleotides most commonly found at each position. Very few promoters match their consensus sequence exactly. In some cases, the match is quite poor, with G or C found at positions normally occupied by A or T. Such promoters are known as weak promoters and are usually associated with genes that are transcribed infrequently. Strong promoters, such as the promoters for ribosomal RNA operons, resemble the consensus sequence quite closely. These operons are transcribed very efficiently. Observations such as these suggest that the consensus sequence describes the most efficient promoter sequence for the RNA polymerase holoenzyme.

KEY CONCEPT

Promoter sequences contain the information that instructs transcription complexes: "Initiate a transcript here."



▲ **Figure 21.6**

Promoter sequences from ten bacteriophage and bacterial genes. All these promoter sequences are recognized by the σ^{70} subunit in *E. coli*. The nucleotide sequences are aligned so that their +1, –10, and –35 regions are in register. Note the degree of sequence variation at each position. The consensus sequence was derived from a much larger database of more than 300 well-characterized promoters.

The promoter sequence of each gene has likely been optimized by natural selection to fit the requirements of the cell. An inefficient promoter is ideal for a gene whose product is not needed in large quantities whereas an efficient promoter is necessary for producing large amounts of a gene product.

C. The σ Subunit Recognizes the Promoter

The effect of σ subunits, also called σ factors, on promoter recognition can best be explained by comparing the DNA-binding properties of core polymerase versus the holoenzyme containing σ^{70}. The core polymerase, which lacks a σ subunit, binds to DNA nonspecifically; it has no greater affinity for promoters than for any other DNA sequence (the association constant, K_a, is approximately $10^{10}\ M^{-1}$). Once formed, this DNA-protein complex dissociates slowly ($t_{1/2} \approx 60$ minutes). In contrast, the holoenzyme, which contains the σ^{70} subunit, binds more tightly to promoter sequences ($K_a \approx 2 \times 10^{11}\ M^{-1}$) than the core polymerase and forms more stable complexes ($t_{1/2} \approx 2$ to 3 hours). Although the holoenzyme binds preferentially to promoter sequences, it also has appreciable affinity for the rest of the DNA in a cell

Table 21.3 *E. coli* σ subunits

Subunit	Gene	Genes transcribed	Consensus	
			–35	–10
σ^{70}	*rpoD*	Many	TTGACA	TATAAT
σ^{54}	*rpoN*	Nitrogen metabolism	None	CTGGCACNNNNNTTGCA[a]
σ^{38}	*rpoS*	Stationary phase	?	TATAAT
σ^{28}	*flaI*	Flagellar synthesis and chemotaxis	TAAA	GCCGATAA
σ^{32}	*rpoH*	Heat shock	CTTGAA	CCCATNTA[a]
σ^{gp55}	gene 55	Bacteriophage T4	None	TATAAATA

[a]N represents any nucleotide.

($K_a \approx 5 \times 10^6\ M^{-1}$). The complex formed by nonspecific binding of the holoenzyme to DNA dissociates rapidly ($t_{1/2} \approx 3$ seconds). These binding parameters reveal the functions of the σ^{70} subunit. One of the roles of σ^{70} is to decrease the affinity of the core polymerase for nonpromoter sequences. Another equally important role is to increase the affinity of the core polymerase for specific promoter sequences.

The association constants do not tell us how the RNA polymerase holoenzyme finds the promoter. We might expect the holoenzyme to search for the promoter by continuously binding and dissociating until it encounters a promoter sequence. Such binding would be a second-order reaction, and its rate would be limited by the rate at which the holoenzyme diffuses in three dimensions. However, promoter binding is 100 times faster than the maximum theoretical value for a diffusion-limited second-order reaction. This remarkable rate is achieved by one-dimensional diffusion of RNA polymerase along the length of the DNA molecule. During the short period of time that the enzyme is bound nonspecifically, it can scan 2000 bp in its search for a promoter sequence. Several other sequence-specific DNA-binding proteins, such as restriction enzymes (Section 19.6C), locate their binding sites in a similar manner.

The binding properties of RNA polymerase tell us that many RNA polymerase molecules will be located on random stretches of DNA that may, or may not, resemble a promoter sequence.

D. RNA Polymerase Changes Conformation

Initiation of transcription is slow, even though the holoenzyme searches for and binds to the promoter very quickly. In fact, initiation is often the rate limiting step in transcription because it requires unwinding of the DNA helix and synthesis of a short stretch of RNA that serves as a primer for subsequent chain elongation. During DNA replication these steps are carried out by a helicase and a primase but in transcription these steps are carried out by the RNA polymerase holoenzyme itself. Unlike DNA polymerases, RNA polymerases can initiate polynucleotide synthesis on their own in the presence of initiation factors such as σ^{70} (when a DNA template and rNTPs are available as substrates).

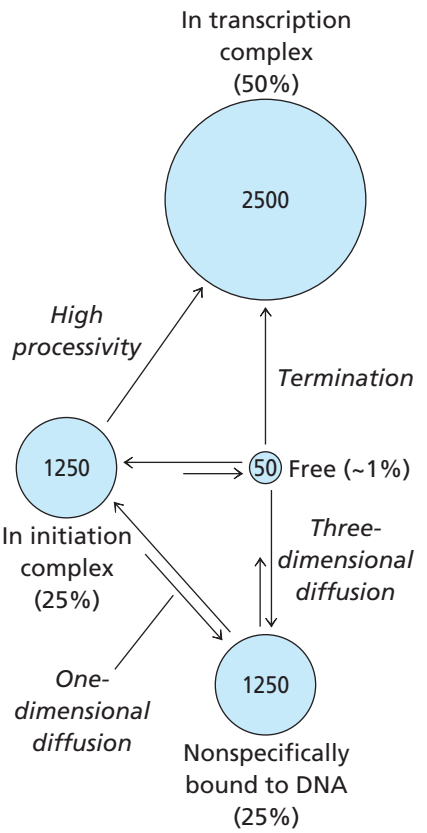
The unwinding of DNA at the initiation site is an example of a conformational change in which RNA polymerase (R) and the promoter (P) shift from a closed complex (RP_c) to an open complex (RP_o). In the closed complex, the DNA is double-stranded. In the open complex, 18 bp of DNA are unwound, forming a transcription bubble. Formation of the open complex is usually the slowest step of the initiation events.

Once the open complex forms, the template strand is positioned at the polymerization site of the enzyme. In the next step, a phosphodiester linkage forms between two ribonucleoside triphosphates that have diffused into the active site and formed hydrogen bonds with the +1 and +2 nucleotides of the template strand. This initiation reaction is slower than the analogous polymerization reaction during chain elongation where one of the substrates (the growing RNA chain) is held in place by the formation of a short RNA-DNA helix.

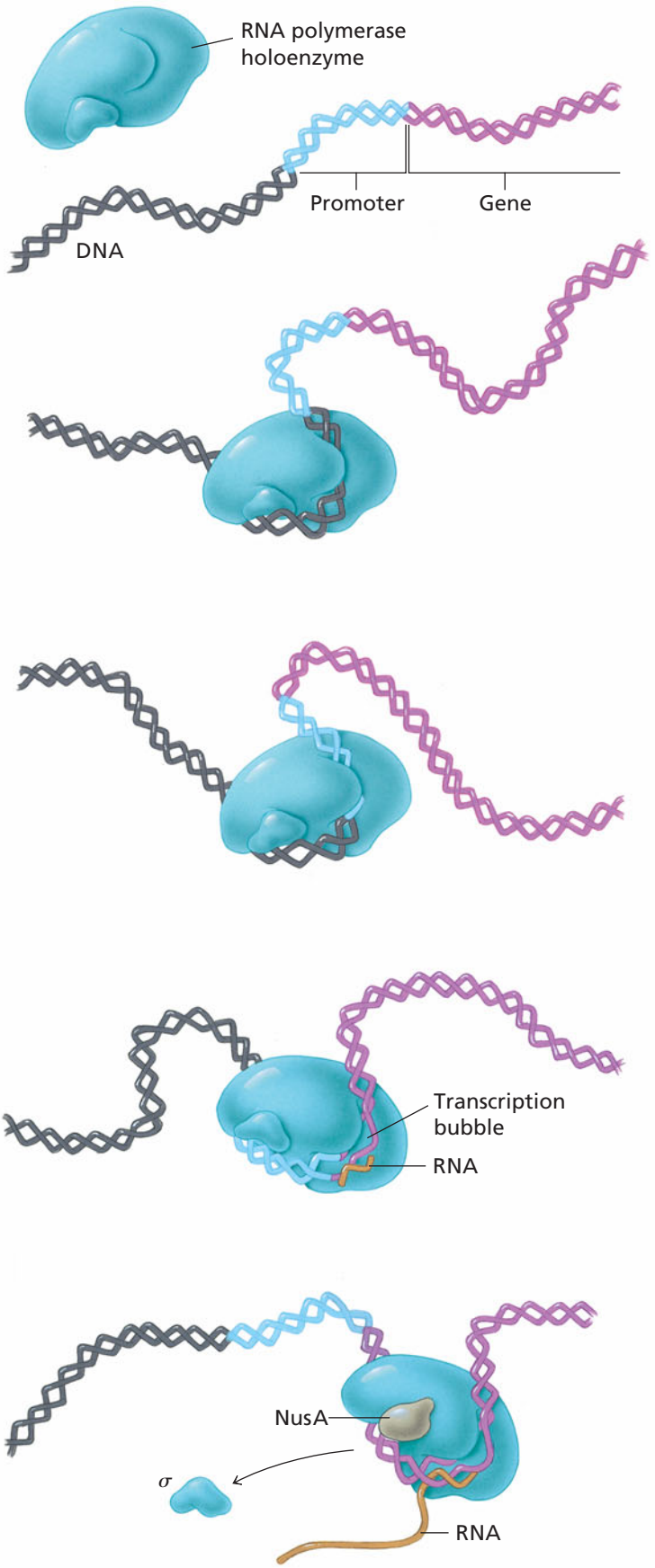
Additional nucleotides are then added to the dinucleotide to create a short RNA that is paired with the template strand. When this RNA is approximately ten nucleotides long, the RNA polymerase holoenzyme undergoes a transition from the initiation to the elongation mode, and the transcription complex moves away from the promoter along the DNA template. This step is called promoter clearance. The initiation reactions can be summarized as

$$R + P \overset{K_{assoc}}{\longleftrightarrow} RP_c \underset{\text{conformational change}}{\longrightarrow} RP_o \underset{\text{promoter clearance}}{\longrightarrow} \qquad (21.2)$$

As noted earlier, the holoenzyme containing σ factor has a much greater affinity for the promoter sequence than for any other DNA sequence. Because of this tight binding, it resists moving away from the initiation site. However, during elongation, the core polymerase binds nonspecifically to all DNA sequences to form a highly processive complex. The transition from initiation to chain elongation is associated with a conformational change in the holoenzyme that causes release of the σ subunit. Without σ, the enzyme no longer binds specifically to the promoter and is able to leave, or exit, the initiation site. At this time, several accessory proteins bind to the



▲ **RNA polymerase distribution.** Estimate of the distribution of the approximately 5000 RNA polymerase molecules typically found in an *E. coli* cell. Very few molecules are free in the cytosol, yet only half of all RNA polymerases are actively transcribing.



◀ **Figure 21.7**
Initiation of transcription in *E. coli*.

(**a**) RNA polymerase holoenzyme binds nonspecifically to DNA.

(**b**) The holoenzyme conducts a one-dimensional search for a promoter.

(**c**) When a promoter is found, the holoenzyme and the promoter form a closed complex.

(**d**) A conformational change from the closed complex to an open complex produces a transcription bubble at the initiation site. A short stretch of RNA is then synthesized.

(**e**) The σ subunit dissociates from the core enzyme, and RNA polymerase clears the promoter. Accessory proteins, including NusA, bind to the polymerase.

core polymerase to create the complete protein machine required for RNA chain elongation. The binding of one of these accessory proteins, NusA, helps convert RNA polymerase to the elongation form. The elongation complex is responsible for most of the synthesis of RNA. NusA also interacts with other accessory proteins and plays a role in termination. Transcription initiation in *E. coli* is summarized in Figure 21.7.

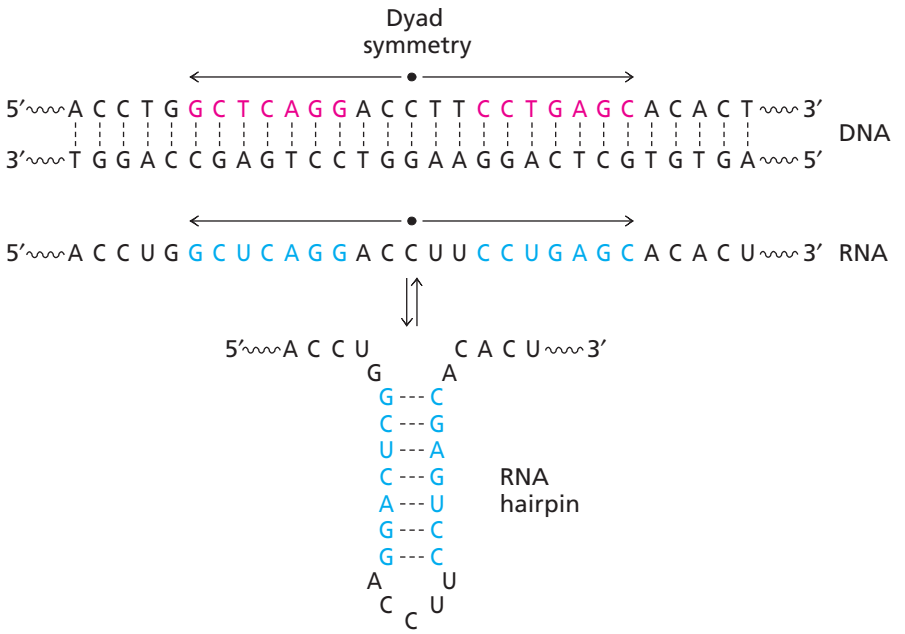
21.4 Transcription Termination

Only certain regions of DNA are specifically transcribed. Transcription complexes assemble at promoters and, in bacteria, disassemble at the 3′ end of genes at specific sequences called **termination sequences**. There are two types of transcription termination sequences. The simplest form of termination occurs at certain DNA sequences where the elongation complex is unstable, and the transcription complex spontaneously disassembles. The other type of termination requires a specific protein named *rho* that facilitates disassembly of the transcription complex, template, and mRNA.

Transcription termination often occurs near **pause sites**. These are regions of the gene where the rate of elongation slows down or stops temporarily. For example, because it is more difficult to melt G/C base pairs than it is to melt A/T base pairs, a transcription complex pauses when it encounters a GC-rich region.

Pausing is exaggerated at sites where the DNA sequence is palindromic, or has dyad symmetry (Section 19.6C). When the DNA is transcribed, the newly synthesized RNA can form a hairpin (Figure 21.8). (A three-dimensional representation of such a structure is shown in Figure 19.21.) Formation of an RNA hairpin may destabilize the RNA-DNA hybrid in the elongation complex by prematurely stripping off part of the newly transcribed RNA. This partial disruption of the transcription bubble probably causes the transcription complex to cease elongation until the hybrid re-forms. NusA increases pausing at palindromic sites, perhaps by stabilizing the hairpin. The transcription complex may pause for 10 seconds to 30 minutes, depending on the structure of the hairpin.

Some of the strong pause sites in *E. coli* are termination sequences. Such termination sites are found at the 3′ end of a gene beyond the region that encodes the polypeptide chain (for protein-encoding genes) or the complete functional RNA (for other genes). These sites specify an RNA hairpin structure that is weakly bound to the template



◀ **Figure 21.8**
Formation of an RNA hairpin. The transcribed DNA sequence contains a region of dyad symmetry. Complementary sequences in RNA can base-pair to form a hairpin.

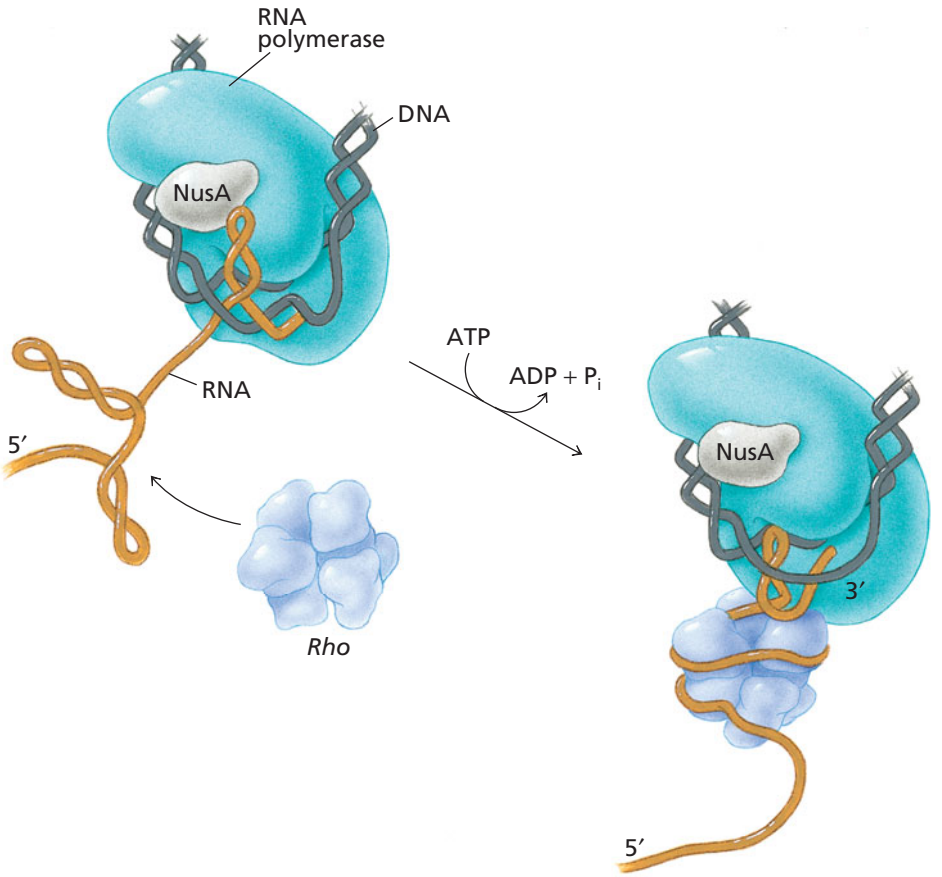


Figure 21.9 ►
***Rho*-dependent termination of transcription in *E. coli*.** RNA polymerase is stalled at a pause site where *rho* binds to newly synthesized RNA. This binding is accompanied by ATP hydrolysis. *Rho* probably wraps the nascent RNA chain around itself, thereby destabilizing the RNA-DNA hybrid and terminating transcription.

[Adapted from Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* 55:339–372.]

strand by a short stretch of A/U base pairs. These are the weakest possible base pairs (Table 19.3) and they are easily disrupted during pausing. Disruption leads to release of RNA from the transcription complex.

The other type of bacterial termination sequences are said to be *rho*-dependent. *Rho* also triggers disassembly of transcription complexes at some pause sites. It is a hexameric protein with a potent ATPase activity and an affinity for single-stranded RNA. *Rho* may also act as an RNA-DNA helicase. It binds to single-stranded RNA that is exposed behind a paused transcription complex in a reaction coupled to hydrolysis of ATP. Approximately 80 nucleotides of RNA wrap around the protein, causing the transcript to dissociate from the transcription complex (Figure 21.9). *Rho*-dependent termination results from both destabilization of the RNA-DNA hybrid and direct contact between the transcription complex and *rho* as *rho* binds RNA. *Rho* can also bind to accessory proteins, such as NusA. This interaction may cause the RNA polymerase to change conformation and dissociate from the template DNA.

Rho-dependent termination requires exposure of single-stranded RNA. In bacteria, RNA transcribed from protein-encoding genes is typically bound by translating ribosomes that interfere with *rho* binding. Single-stranded RNA only becomes exposed to *rho* when transcription passes beyond the point where protein synthesis terminates. Transcription terminates at the next available pause site. In other words, *rho*-dependent termination does not occur at pause sites within the coding region but can occur at pause sites past the translation termination codon. The net effect is to couple transcription termination to translation. The advantages of such a coupling mechanism are that synthesis of an mRNA coding region is not interrupted (which would prevent protein synthesis) and that there is minimal wasteful transcription downstream of the coding region.

21.5 Transcription in Eukaryotes

KEY CONCEPT

Eukaryotic transcription complexes tend to have more factors than the analogous bacterial complexes.

The same processes carried out by a single RNA polymerase in *E. coli* are carried out in eukaryotes by several similar enzymes. The activities of eukaryotic transcription complexes also require many more accessory proteins than those seen in bacteria.

A. Eukaryotic RNA Polymerases

Three different RNA polymerases transcribe nuclear genes in eukaryotes. Other RNA polymerases are found in mitochondria and chloroplasts. Each nuclear enzyme transcribes a different class of genes (Table 21.4). RNA polymerase I transcribes genes that encode large ribosomal RNA molecules (class I genes). RNA polymerase II transcribes genes that encode proteins and a few that encode small RNA molecules (class II genes). RNA polymerase III transcribes genes that encode a number of small RNA molecules, including tRNA and 5S rRNA (class III genes). (Some of the RNA molecules listed in the table are discussed in subsequent sections.)

The mitochondrial version of RNA polymerase is a monomeric enzyme encoded by the nuclear genome. It is substantially similar in amino acid sequence to the RNA polymerases of T3 and T7 bacteriophages. This similarity suggests that these enzymes share a common ancestor. It is likely that the gene for mitochondrial RNA polymerase was transferred to the nucleus from the primitive mitochondrial genome.

Chloroplast genomes often contain genes that encode their own RNA polymerase. The genes encoding the chloroplast RNA polymerase are similar in sequence to those of RNA polymerase in cyanobacteria. This is further evidence that chloroplasts, like mitochondria, originated from bacterial endosymbionts in ancestral eukaryotic cells.

The three nuclear RNA polymerases are complex multisubunit enzymes. They differ in subunit composition, although they share several small polypeptides in common. The exact number of subunits in each polymerase varies among organisms but there are always 2 large subunits and 7 to 12 smaller ones (Figure 21.10). RNA polymerase II transcribes all protein-coding genes as well as some genes that encode small RNA molecules. The protein-coding RNA synthesized by this enzyme was originally called heterogeneous nuclear RNA (hnRNA) but it is now more commonly referred to as mRNA precursor, or pre-mRNA. The processing of this precursor into mature mRNA is described in Section 21.9.

About 40,000 molecules of RNA polymerase II are found in large eukaryotic cells; the activity of this enzyme accounts for roughly 20% to 40% of all cellular RNA synthesis. The two largest subunits of each nuclear eukaryotic RNA polymerase are similar in sequence to the β and β' subunits of *E. coli* RNA polymerase indicating that they share a common ancestor. Like their prokaryotic counterparts, the core eukaryotic RNA

Table 21.4 Eukaryotic RNA polymerases

Polymerase	Location	Copies per cell	Products	Polymerase activity of cell
RNA polymerase I	Nucleolus	40,000	35–47S pre-rRNA	50%–70%
RNA polymerase II	Nucleoplasm	40,000	mRNA precursors U1, U2, U4, and U5 snRNA	20%–40%
RNA polymerase III	Nucleoplasm	20,000	5S rRNA tRNA U6 snRNA 7S RNA Other small RNA molecules	10%
Mitochondrial RNA polymerase	Mitochondrion	?	Products of all mitochondrial genes	<1%
Chloroplast RNA polymerase	Chloroplast	?	Products of all chloroplast genes	<1%

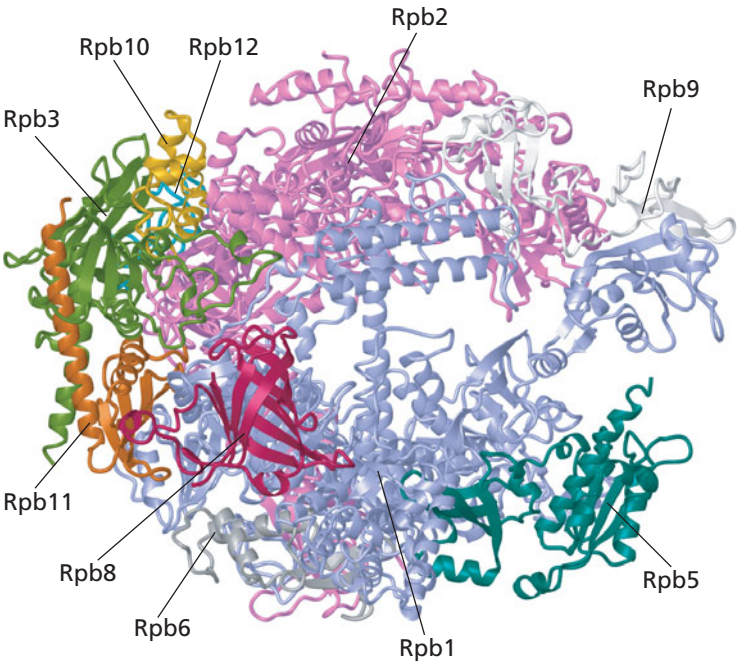


Figure 21.10 ▸
RNA polymerase II from the yeast *Saccharomyces cerevisiae*. The large subunit colored purple (Rpb2) is the homolog of the β subunit of the prokaryotic enzyme shown in Figure 21.2. [PDB 1EN0].

polymerases do not bind on their own to promoters. RNA polymerase II requires five different biochemical activities, or factors, to form a basal transcription complex capable of initiating transcription on a minimal eukaryotic promoter (Figure 21.11). These general transcription factors (GTFs) are: TFIIB, TFIID, TFIIE, TFIIF and TFIIH (Table 21.5).

Many class II genes contain an A/T-rich region, also called a TATA box, that is functionally similar to the prokaryotic TATA box discussed above (recall that A/T-rich regions are more easily unwound to create an open complex, especially if the DNA is negatively supercoiled (Section 19.3)). This eukaryotic A/T-rich region is located 19 to 27 bp upstream of the transcription start site and serves to recruit RNA polymerase II to the DNA during assembly of the initiation complex.

The general transcription factor TFIID is a multisubunit factor and one of its subunits, TATA-binding protein (TBP), binds to the region containing the TATA box. The structure of TBP from the plant *Arabidopsis thaliana* is shown in Figure 21.12. TBP forms a saddle-shaped molecular clamp that almost surrounds the DNA at the TATA box. The main contacts between TBP and DNA are due to interactions between acidic amino acid side chains in β strands and the edges of base pairs in the minor groove. When TBP binds to DNA, the promoter DNA is bent so that it no longer resembles the standard B-DNA conformation. This is an unusual interaction for DNA-binding proteins. The TBP subunit of TFIID is also required to initiate transcription of class I and class III genes by RNA polymerases I and III, respectively.

The eukaryotic RNA polymerase II subunit homologous to the prokaryotic RNA polymerase β' subunit has an unusual carboxy-terminal domain (CTD) or "tail" that

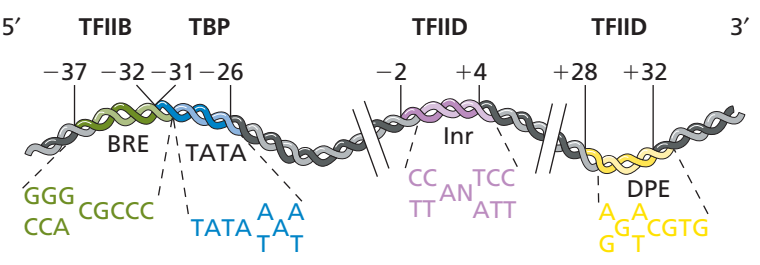
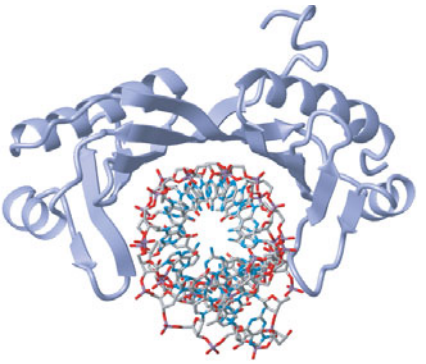


Figure 21.11 ▸
A generic eukaryotic promoter showing the basal or "core" promoter elements. The TATA box is described in the text. The BRE is the TFIIB recognition element, while Inr stands for the initiator element. The DPE is the downstream promoter element. The names of the factors that bind to each site are shown above the promoter, and the consensus recognition sequences for each site are shown below the schematic promoter fragment.

Table 21.5 Some representative RNA polymerase II transcription factors

Factor	Characteristics
TFIIA	Binds to TFIID; can interact with TFIID in the absence of DNA
TFIIB	Interacts with RNA polymerase II
TFIID	RNA polymerase II initiation factor
TBP	TATA-binding protein; subunit of TFIID
TAFs	TBP-associated factors; many subunits
TFIIE	Interacts with RNA polymerase II
TFIIH	Required for initiation; helicase activity; couples transcription to DNA repair
TFIIS[a]	Binds to RNA polymerase II; elongation factor
TFIIF	Binds to RNA polymerase II; two subunits—RAP30 and RAP74
SP1	Binds to GC-rich sequence
CTF[b]	Family of different proteins that recognize the core sequence CCAAT

[a]Also known as sII or RAP38
[b]Also known as NP1.



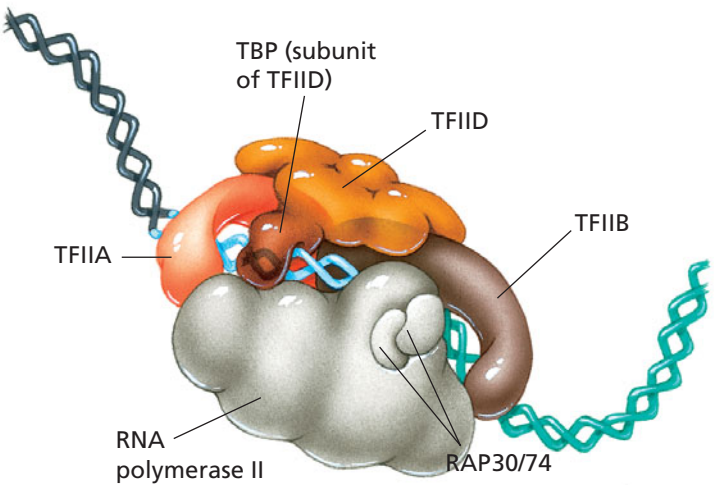
▲ **Figure 21.12**
***Arabidopsis thaliana* TATA-binding protein (TBP) bound to DNA.** TBP (blue) is bound to a double-stranded DNA fragment with a sequence corresponding to a TATA box (5′-TATAAAG-3′) DNA is shown as a wire-frame model. Note that the β sheet of TBP lies in the minor groove of the DNA fragment. [PDB 1VOL].

consists of multiple repeats of the amino acid heptamer PTSPSYS. The Ser and Thr residues in the tail are phosphorylation targets for nuclear protein kinases. RNA polymerase II molecules with a hyperphosphorylated CTD are typically transcriptionally active, or engaged, while the cellular pol II with hypophosphorylated CTDs are usually quiescent.

Although it has proven possible to purify RNA polymerase II and each GTF and use them to reconstitute accurate transcription initiation *in vitro*, these basal transcription complexes are not competent to recognize the many different types of *trans*-acting factors and *cis*-acting sequences that are known to play important roles *in vivo*. Searching for cellular constituents that could respond to transcriptional activators *in vitro* led to the discovery of a large preformed RNA pol II holoenzyme that contains not only the five GTFs but also many other polypeptides that mediate interactions between pol II and sequence-specific DNA-binding proteins. This eukaryotic holoenzyme is analogous to the core + σ holoenzyme in *E. coli.*

B. Eukaryotic Transcription Factors

TFIIA and TFIIB are essential components of the RNA polymerase II holoenzyme complex. Neither TFIIA nor TFIIB can bind to DNA in the absence of TFIID. TFIIF (also known as Factor 5 or RAP30/74) binds to RNA polymerase II during initiation (Figure 21.13). TFIIF plays no direct role in recognizing the promoter but it is analogous to bacterial σ factors in two ways: it decreases the affinity of RNA polymerase II for nonpromoter



◀ **Figure 21.13**
RNA polymerase II holoenzyme complex bound to a promoter. This model shows various transcription factors bound to RNA polymerase II at a promoter. The transcription factors are often larger and more complex than those shown in this diagram.

DNA, and it helps form the open complex. TFIIH, TFIIE, and other, less well-characterized factors, are also part of the transcription initiation complex.

Once the initiation complex assembles at the site of the promoter, the next steps are similar to those in bacteria. An open complex is formed, a short stretch of RNA is synthesized, and the transcription complex clears the promoter. Most transcription factors dissociate from DNA and RNA polymerase II once elongation begins. However, TFIIF may remain bound and a specific elongation factor, TFIIS (also called sII or RAP38), associates with the transcribing polymerase. TFIIS may play a role in pausing and transcription termination that is similar to the role of NusA in bacteria.

With the exception of TBP, the transcription factors that interact with the other two eukaryotic RNA polymerases are not the same as those required by RNA polymerase II.

C. The Role of Chromatin in Eukaryotic Transcription

As described in Chapter 19, the eukaryotic genome is packaged using small, ubiquitous building blocks, called nucleosomes, that contain an octamer of the four core histone proteins. It is estimated that approximately 35% of the mammalian genome is transcribed into protein-coding genes (including the introns) and so most of a cell's DNA is relatively inert. But even within that 35%, which contains about 20,000 protein-coding genes, the majority of the sequences are quiescent. In any single cell, the primary determinant of whether a gene is competent to be transcribed resides in the state of its chromatin. This status is modulated by two mechanisms. The first involves implementing or removing post-translational modifications on the flexible amino-terminal arms of the four core histones (Section 19.5B). Specific Lys residues are targeted for methylation or acetylation, specific Arg residues may also be methylated, while Ser and Thr side chains can be phosphorylated. Different modifications serve as signals to recruit either activators or repressors to the chromatin. The second mechanism for specifying the transcriptional status of a eukaryotic gene involves nucleosome positioning and remodeling.

Nontranscribed genes are relatively inaccessible in the nucleus while transcribed genes are relatively accessible to transcription factors, pol II holoenzyme, and other nuclear proteins. How does a gene move between these two conflicting states? The answer lies with large multiprotein complexes that use the energy from hydrolyzing ATP to physically remodel a gene's nucleosomes and allow proteins to have access to the DNA. Some of the remodeling complexes actually contain histone-modifying enzymes like histone acetylase (HAT) or histone deacetylase (HDAC).

21.6 Transcription of Genes Is Regulated

As noted at the beginning of this chapter, many genes are expressed in every cell. The expression of these housekeeping genes is said to be ***constitutive***. In general, such genes have strong promoters and are transcribed efficiently and continuously. Genes whose products are required at low levels usually have weak promoters and are transcribed infrequently. In addition to constitutively expressed genes, cells contain genes that are expressed at high levels in some circumstances and not at all in others. Such genes are said to be regulated.

Regulation of gene expression can occur at any point in the flow of biological information but occurs most often at the level of transcription. Various mechanisms have evolved that allow cells to program gene expression during differentiation and development and to respond to environmental stimuli.

The initiation of transcription of regulated genes is controlled by regulatory proteins that bind to specific DNA sequences. Transcriptional regulation can be negative or positive. Transcription of a negatively regulated gene is prevented by a regulatory protein called a **repressor**. A negatively regulated gene can be transcribed only in the absence of an active repressor. Transcription of a positively regulated gene can be activated

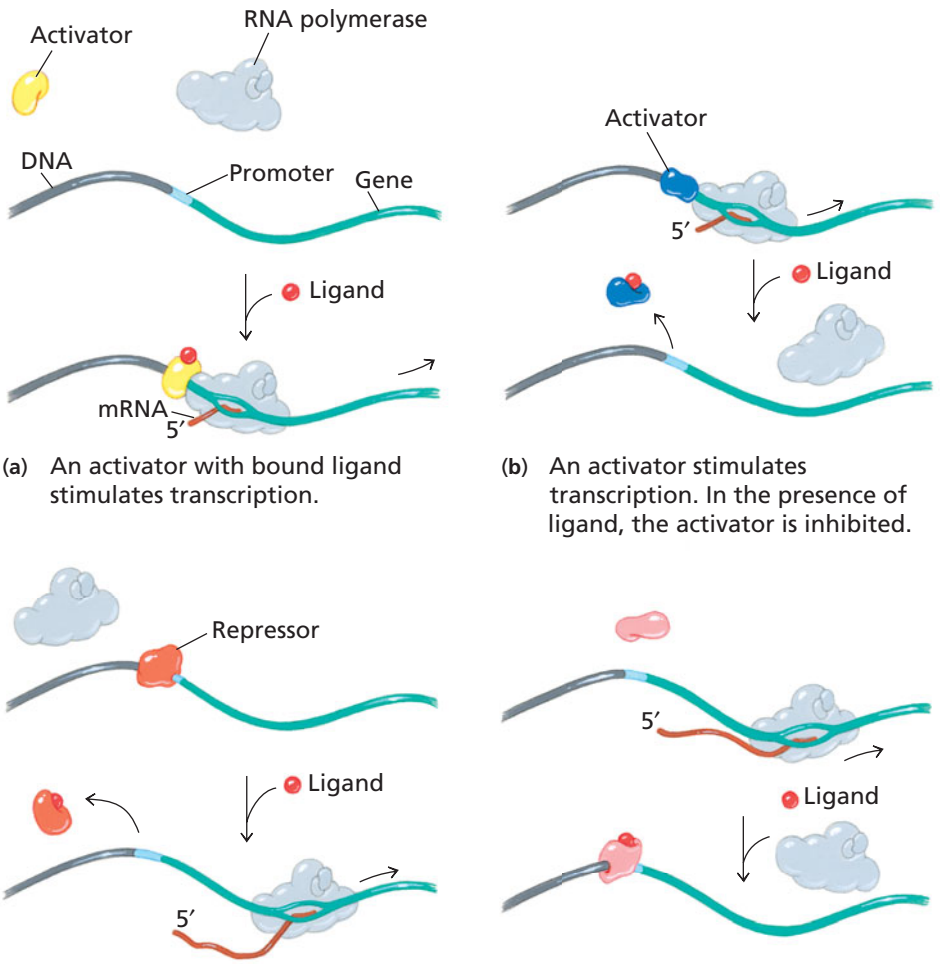
by a regulatory protein called an **activator**. A positively regulated gene is transcribed poorly or not at all in the absence of the activator.

Repressors and activators are often allosteric proteins whose function is modified by ligand binding. In general, a ligand alters the conformation of the protein and affects its ability to bind to specific DNA sequences. For example, some repressors control the synthesis of enzymes for a catabolic pathway. In the absence of substrate for these enzymes, the genes are repressed. When substrate is present, it binds to the repressor, causing the repressor to dissociate from the DNA and allowing the genes to be transcribed. Ligands that bind to and inactivate repressors are called **inducers** because they induce transcription of the genes controlled by the repressors. In contrast, some repressors that control the synthesis of enzymes for a biosynthetic pathway bind to DNA only when associated with a ligand. The ligand is often the end product of the biosynthetic pathway. This regulatory mechanism ensures that the genes in the pathway are turned off as product of the pathway accumulates. Ligands that bind to and activate repressors are called **corepressors**. The DNA-binding activity of allosteric activators can also be affected in two ways by ligand binding. Four general strategies for regulating transcription are illustrated in Figure 21.14. Examples of all four strategies have been identified.

Few regulatory systems are as simple as those described above. For example, the transcription of many genes is regulated by a combination of repressors and activators or by multiple activators. Elaborate mechanisms for regulating transcription

KEY CONCEPT

Cells don't synthesize a specific protein until it is required (e.g., the *lac* operon is not transcribed until the intracellular concentration of lactose inactivates the *lac* repressors).



(a) An activator with bound ligand stimulates transcription.

(b) An activator stimulates transcription. In the presence of ligand, the activator is inhibited.

(c) A repressor prevents transcription. Binding of ligand (inducer) to the repressor inactivates the repressor and allows transcription.

(d) In the absence of ligand, the repressor does not bind to DNA. Repression occurs only when ligand (corepressor) is present.

◂ **Figure 21.14**
Strategies for regulating transcription initiation by regulatory proteins.

have evolved to meet the specific requirements of individual organisms. A greater range of cellular responses is possible when transcription is regulated by a host of mechanisms acting together. By examining how the transcription of a few particular genes is controlled, we can begin to understand how positive and negative mechanisms can be combined to produce the remarkably sensitive regulation seen in bacterial cells.

21.7 The *lac* Operon, an Example of Negative and Positive Regulation

Some bacteria obtain the carbon they need for growth by metabolizing five- or six-carbon sugars via glycolysis. For example, *E. coli* preferentially uses glucose as a carbon source but can also use other sugars, including β-galactosides such as lactose. The enzymes required for β-galactoside uptake and catabolism are not synthesized unless a β-galactoside substrate is available. Even in the presence of their substrate, these enzymes are synthesized in limited amounts when the preferred carbon source (glucose) is also present. Synthesis of the enzymes required for β-galactoside utilization is regulated at the level of transcription initiation by a repressor and an activator.

The uptake and catabolism of β-galactosides requires three proteins. The product of the *lacY* gene is lactose permease, a symport transporter that is responsible for the uptake of β-galactosides. Most β-galactosides are subsequently hydrolyzed to metabolizable hexoses by the activity of β-galactosidase, a large enzyme with four identical subunits encoded by the *lacZ* gene. β-Galactosides that cannot be hydrolyzed are acetylated by the activity of thiogalactoside transacetylase, the product of the *lacA* gene. Acetylation helps to eliminate toxic compounds from the cell.

The three genes—*lacZ*, *lacY*, and *lacA*—form an operon that is transcribed from a single promoter to produce a large mRNA molecule containing three separate protein-coding regions. In this case, we refer to a protein-coding region as a gene, a definition that differs from our standard use of the term. The arrangement of genes with related functions in an operon is efficient because the concentrations of a set of proteins can be controlled by transcribing from a single promoter. Operons composed of protein-coding genes are common in *E. coli* and other prokaryotes but were thought to be extremely rare in eukaryotes. We now realize that operons are also quite common in the model organism *C. elegans*, a nematode or round worm, and are likely widespread in this large phylum. Operons are also common in mitochondrial and chloroplast genomes.

A. *lac* Repressor Blocks Transcription

Expression of the three genes of the *lac* operon is controlled by a regulatory protein called *lac* repressor, a tetramer of identical subunits. The repressor is encoded by a fourth gene, *lacI*, which is located just upstream of the *lac* operon but is transcribed from a separate promoter (Figure 21.15).

lac repressor binds simultaneously to two sites near the promoter of the *lac* operon. Repressor-binding sites are called **operators**. One operator (O_1) is adjacent to the promoter, and the other (O_2) is within the coding region of *lacZ*. When bound to both operators, the repressor causes the DNA to form a stable loop that can be seen

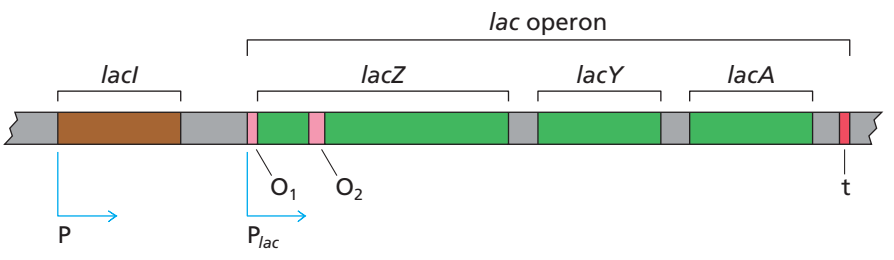
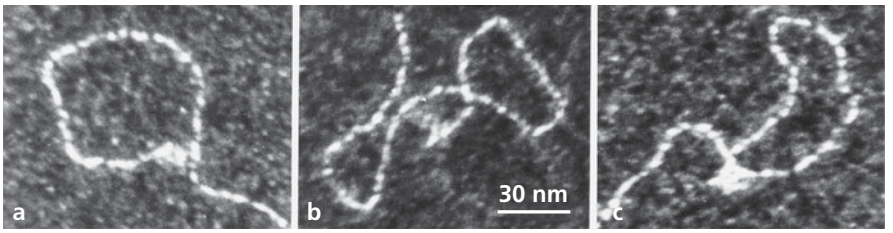


Figure 21.15 ▸
Organization of the genes that encode proteins required to metabolize lactose. The coding regions for three proteins—LacZ, LacY, and LacA—constitute the *lac* operon and are co-transcribed from a single promoter (P_{lac}). The gene that encodes *lac* repressor, *lacI*, is located upstream of the *lac* operon and has its own promoter, P; *lac* repressor binds to the operators O_1 and O_2 near P_{lac}; t denotes the transcription termination sequence.



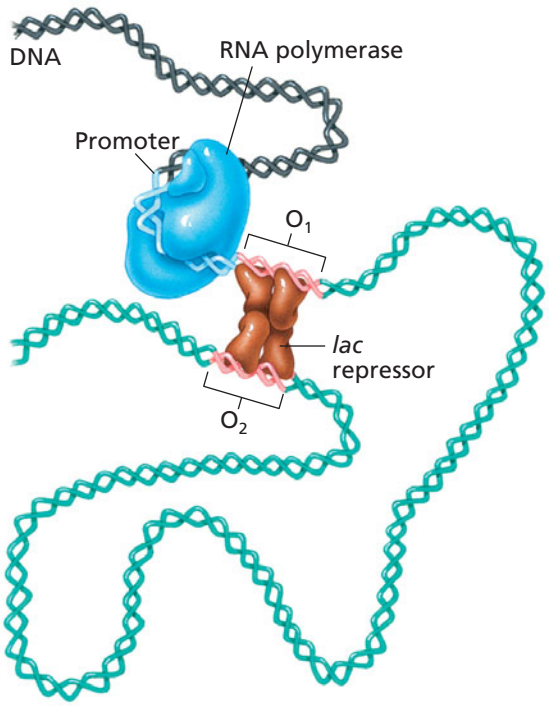
◀ **Figure 21.16**
Electron micrographs of DNA loops. These loops were formed by mixing *lac* repressor with a fragment of DNA bearing two synthetic *lac* repressor–binding sites. One binding site is located at one end of the DNA fragment, and the other is 535 bp away. DNA loops 535 bp in length form when the tetrameric repressor binds simultaneously to the two sites.

in electron micrographs of the complex formed between *lac* repressor and DNA (Figure 21.16). The interaction of *lac* repressor with the operator sequences may block transcription by preventing the binding of RNA polymerase to the *lac* promoter. However, it is now known that, in some cases, both *lac* repressor and RNA polymerase can bind to the promoter at the same time. Thus, the repressor may also block transcription initiation by preventing formation of the open complex and promoter clearance. A schematic diagram of *lac* repressor bound to DNA in the presence of RNA polymerase is shown in Figure 21.17. The diagram illustrates the relationship between the operators and the promoter and the DNA loop that forms when the repressor binds to DNA.

At any given time, one molecule of repressor is bound to the operator and nine molecules are bound non-specifically to DNA.

The repressor locates an operator by binding nonspecifically to DNA and searching by sliding or hopping in one dimension. The non-specific equilibrium constant is about $10^6\ M^{-1}$—comparable to that of RNA polymerase (Section 21.3C). (Recall from Section 21.3C that RNA polymerase also uses this kind of searching mechanism.) The equilibrium association constant for the specific binding of *lac* repressor to O_1 *in vitro* is very high ($K_a \approx 10^{13}M^{-1}$). As a result, the repressor blocks transcription very effectively. (*lac* repressor binds to the O_2 site with lower affinity.) A bacterial cell contains only about ten molecules of *lac* repressor but the repressor searches for and finds an operator so rapidly that when a repressor dissociates spontaneously from the operator, another occupies the site within a very short time. However, during this brief interval, one transcript of the operon can be made since RNA polymerase is poised at the promoter. This low level of transcription, called escape synthesis, ensures that small amounts of lactose permease and β-galactosidase are present in the cell.

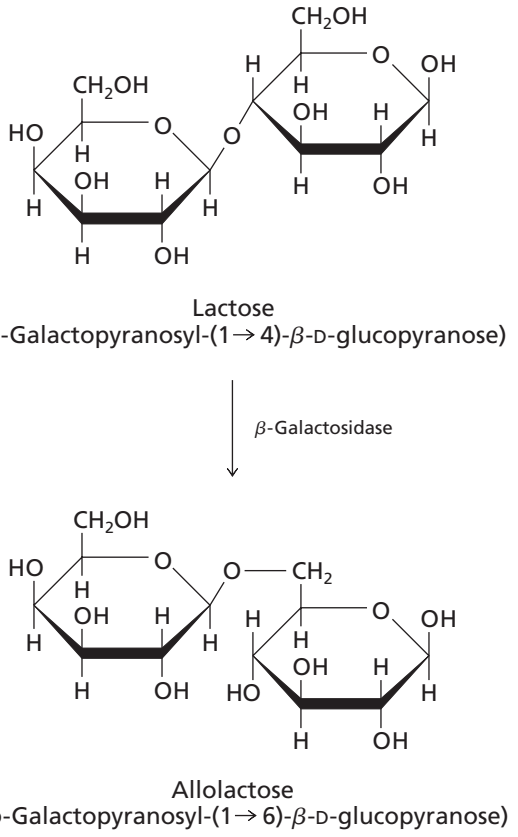
In the absence of lactose, *lac* repressor blocks expression of the *lac* operon, but when β-galactosides are available as potential carbon sources, the genes are transcribed. Several β-galactosides can act as inducers. If lactose is the available carbon source, the inducer is allolactose, which is produced from lactose by the action of β-galactosidase (Figure 21.18). Allolactose binds tightly to *lac* repressor and causes a conformational change that reduces the affinity of the repressor for the operators ($K_a \approx 10^{10}M^{-1}$). In the presence of the inducer, *lac* repressor dissociates from the DNA, allowing RNA polymerase to initiate transcription. (Note that because of escape synthesis, lactose can be taken up and converted to allolactose even when the operon is repressed.)



▲ **Figure 21.17**
Binding of *lac* repressor to the *lac* operon. The tetrameric *lac* repressor interacts simultaneously with two sites near the promoter of the *lac* operon. As a result, a loop of DNA forms. RNA polymerase can still bind to the promoter in the presence of the *lac* repressor–DNA complex.

B. The Structure of *lac* Repressor

The role of *lac* repressor in regulating expression of the *lac* operon has been known since the 1960s. However, the structure of this important protein was solved only in the 1990s after the development of new techniques for determining the structure of large molecules. The structure of part of the *lac* repressor bound to one operator sequence is shown in Figure 21.19. The complete protein contains four identical subunits arranged as two pairs, and each pair of subunits binds to a different operator sequence. Inside the cell these two fragments of DNA are part of a single DNA molecule—and repressor binding forms a loop of DNA at the 5′ end of the *lac* operon.



▲ Figure 21.18
Formation of allolactose from lactose, catalyzed by β-galactosidase. This is a minor or side reaction. The main enzymatic activity of β-galactosidase is to cleave disaccharides into monomers that can be converted into substrates for glycolysis.

The subunits are joined together at a hinge region. The X-ray crystallographic structure reveals that the two pairs of subunits are stacked on top of one another (Figure 21.17) and not extended away from the hinge region as was expected. This makes a more compact protein that is less symmetric than many other tetrameric proteins.

Each subunit contains a helix-turn-helix motif at the ends farthest from the hinge region. When bound to DNA, one of the α helices lies in the major groove where amino acid side chains interact directly with the specific base pairs of the operator sequence. The two helices from each pair of subunits are positioned about one turn of DNA apart (about 10 bp), and each one interacts with half of the operator sequence. This binding strategy is similar to that of restriction endonuclease *Eco*RI (Section 19.6C).

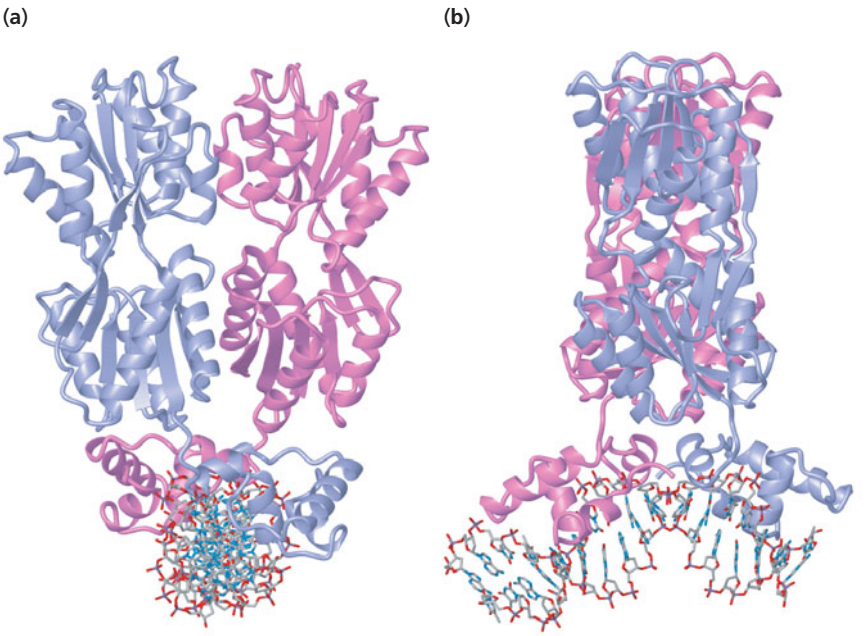
In the absence of DNA the distal regions of the *lac* repressor subunits are disordered (Section 4.7D). This is one reason why it took such a long time to work out the structure. The structure of the helix-turn-helix motif can only be seen when the protein is bound to DNA. There are now many examples of such interactions in which the stable structure of the protein is significantly altered by ligand binding. In the presence of inducers, such as allolactose or IPTG, the repressor adopts a slightly different conformation and can no longer bind to the DNA operators.

C. cAMP Regulatory Protein Activates Transcription

Transcription of the *lac* operon in *E. coli* depends not only on the presence of β-galactosides but also on the concentration of glucose in the external medium. The *lac* operon is transcribed maximally when β-galactosides, such as lactose, are the only carbon source; transcription is reduced 50-fold when glucose is also present. The decreased rate of transcription of operons when glucose is present is termed catabolite repression.

Catabolite repression is a feature of many operons encoding metabolic enzymes. These operons characteristically have weak promoters from which transcription is initiated inefficiently in the presence of glucose. In the absence of glucose, however, the rate of transcription initiation increases dramatically due to an activator that converts the relatively weak promoter to a stronger one. No repressor is involved, despite the

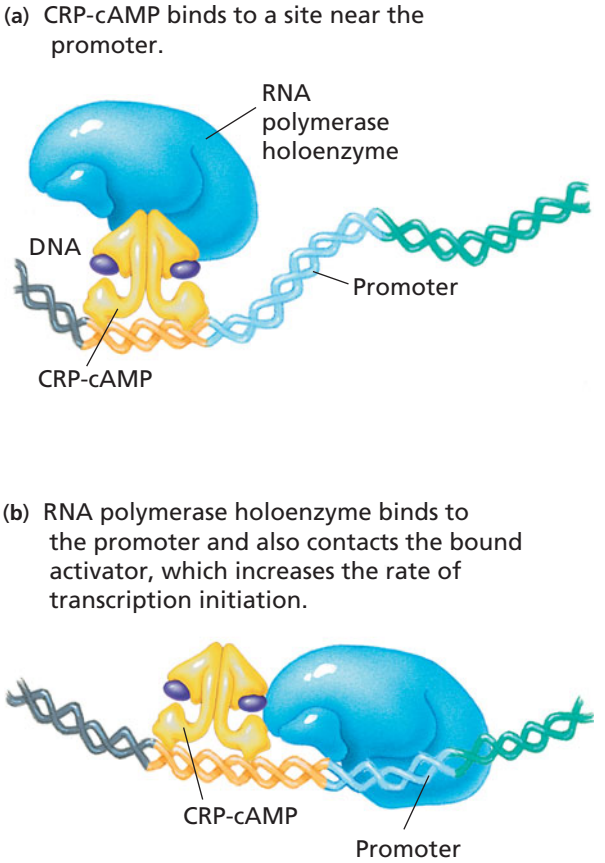
Figure 21.19 ►
Structure of *E. coli lac* repressor. This figure shows a dimer of *lac* repressor subunits bound to DNA. *Lac* repressor is a tetramer *in vivo*, containing two DNA-binding sites. **(a)** End-on view of the DNA molecule. **(b)** Side view showing the *lac* repressor α helix in the major groove. [PDB 1EFA].



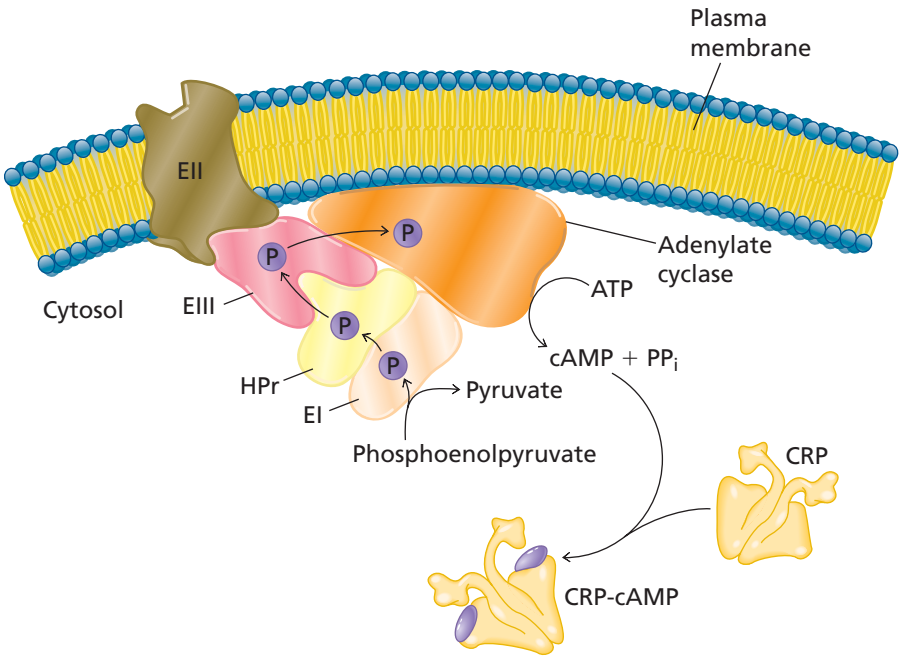
use of the term *catabolite repression*. In fact, this is a well-studied example of an activation mechanism.

The activator is cyclic AMP regulatory (or receptor) protein (CRP), also known as catabolite activator protein (CAP). CRP is a dimeric protein whose activity is modulated by cyclic AMP. In the absence of cAMP, CRP has low affinity for DNA but when cAMP is present it binds to CRP and converts it to a sequence-specific DNA-binding protein. The CRP-cAMP complex interacts with specific DNA sequences near the promoters of more than 30 genes including the *lac* operon. Because the genome contains many more binding sites for CRP-cAMP than for *lac* repressor, it is not surprising that there are at least 1000 molecules of CRP per cell compared to only about 10 molecules of *lac* repressor. The CRP-cAMP binding sites are often just upstream of the −35 regions of the promoters they activate. While bound to DNA, CRP-cAMP can contact RNA polymerase at the promoter site, leading to increased rates of transcription initiation (Figure 21.20). Most of the protein–protein interactions are between bound CRP-cAMP and the α subunits of RNA polymerase. This is typical of most interactions between activators and RNA polymerase. (There are many different transcriptional activators in bacterial cells.) The net effect of CRP-cAMP is to increase the production of enzymes that can use substrates other than glucose. In the case of the *lac* operon, activation by CRP-cAMP occurs only when β-galactosides are available. At other times, transcription of the operon is repressed.

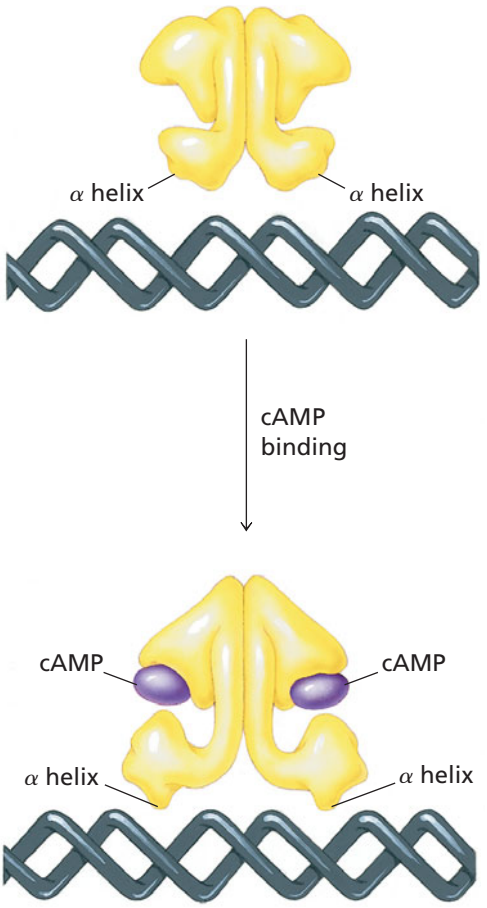
The concentration of cAMP inside an *E. coli* cell is controlled by the concentration of glucose outside the cell. When glucose is available, it is imported into the cell and phosphorylated by a complex of transport proteins collectively known as the phosphoenolpyruvate-dependent sugar phosphotransferase system. When glucose is not available, one of the glucose transport enzymes, enzyme III, catalyzes the transfer of a phosphoryl group, ultimately derived from phosphoenolpyruvate, to adenylate cyclase, leading to its activation (Figure 21.21).



▲ **Figure 21.20**
Activation of transcription initiation at the *lac* promoter by CRP-cAMP.



◀ **Figure 21.21**
cAMP production. In the absence of glucose, enzyme III (EIII) transfers a phosphoryl group, originating from phosphoenolpyruvate, to membrane-bound adenylate cyclase. Phosphorylated adenylate cyclase catalyzes the conversion of ATP to cAMP. cAMP binds to CRP, and CRP-cAMP activates the transcription of a number of genes encoding enzymes that compensate for the lack of glucose as a carbon source.



▲ **Figure 21.22**
Conformational changes in CRP caused by cAMP binding. Each monomer of the CRP dimer contains a helix-turn-helix motif. In the absence of cAMP, the α helices cannot fit into adjacent sections of the major groove of DNA and cannot recognize the CRP-cAMP binding site. When cAMP binds to CRP, the two α helices assume the proper conformation for binding to DNA.

Adenylate cyclase (also know as adenylyl cyclase) then catalyzes the conversion of ATP to cAMP thereby increasing the levels of cAMP in the cell. As molecules of cAMP are produced, they bind to CRP stimulating transcription initiation at promoters that respond to catabolite repression. Similar mechanisms for responding to external stimuli operate in eukaryotes where molecules such as cAMP act as second messengers (Section 9.12B).

Each subunit of the CRP dimer contains a helix-turn-helix DNA binding motif. In the presence of cAMP, two helices—one from each monomer—fit into adjacent sections of the major groove of DNA and contact the nucleotides of the CRP-cAMP binding site. This is the same general binding strategy used by *lac* repressor and *Eco*RI. In the absence of cAMP, the conformation of CRP changes so that the two α helices can no longer bind to the major groove (Figure 21.22). When CRP-cAMP is bound to the activator sequence, the DNA is bent slightly to conform to the surface of the protein (Figure 21.23).

21.8 Post-transcriptional Modification of RNA

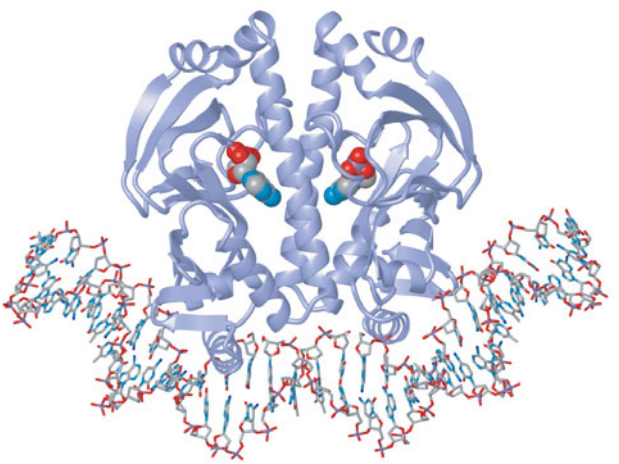
In many cases, RNA transcripts must be extensively altered before they can adopt their mature structures and functions. These alterations fall into three general categories: (1) removal of nucleotides from primary RNA transcripts; (2) addition of nucleotides not encoded by the corresponding genes; and (3) covalent modification of certain bases. The reactions that transform a primary RNA transcript into a mature RNA molecule are referred to collectively as **RNA processing**. RNA processing is crucial for the function of most RNA molecules and is an integral part of gene expression.

A. Transfer RNA Processing

Mature tRNA molecules are generated in both eukaryotes and prokaryotes by processing primary transcripts. In prokaryotes, the primary transcript often contains several tRNA precursors. These precursors are cleaved from the large primary transcripts and trimmed to their mature lengths by ribonucleases, or RNases. Figure 21.24 summarizes the processing of prokaryotic tRNA precursors.

The endonuclease RNase P catalyzes the initial cleavage of most tRNA primary transcripts. The enzyme cleaves the transcript on the 5′ side of each tRNA sequence, releasing monomeric tRNA precursors with mature 5′ ends. Digestion with RNase P *in vivo* is rapid and occurs while the transcript is still being synthesized.

Figure 21.23 ►
Structure of a complex between CRP-cAMP and DNA. Both subunits contain a molecule of cAMP bound at the allosteric site. Each subunit has an α helix positioned in the major groove of DNA at the CRP-cAMP binding site. Note that binding induces a slight bend in the DNA. [PDB 1CGP].



RNase P was one of the first specific ribonucleases studied in detail and much is known about its structure. The enzyme is actually a ribonucleoprotein. In *E. coli,* it is composed of a 377-nucleotide RNA molecule (M_r 130,000) and a small polypeptide (M_r 18,000). In the absence of protein the RNA component is catalytically active *in vitro* (under certain conditions). It was one of the first RNA molecules shown to have enzymatic activity and is an example of the fourth class of RNA molecules described in Section 21.1. The protein component of RNase P helps maintain the three-dimensional structure of the RNA. Sidney Altman was awarded the Nobel Prize in 1989 for showing that the RNA component of RNase P had catalytic activity.

Other endonucleases cleave tRNA precursors near their 3′ ends. Subsequent processing of the 3′ end of a tRNA precursor requires the activity of an exonuclease, such as RNase D. This enzyme catalyzes the sequential removal of nucleotides from the 3′ end of a monomeric tRNA precursor until the 3′ end of the tRNA is reached.

All mature prokaryotic and eukaryotic tRNA molecules must contain the sequence CCA as the final three nucleotides at their 3′ ends. In some cases, these nucleotides are added post-transcriptionally after all other types of processing at the 3′ end have been completed. The addition of these three nucleotides is catalyzed by tRNA nucleotidyltransferase and is one of the few examples of the addition of nucleotides that are not encoded by a gene.

Processing of tRNA precursors also involves covalently modifying some of the nucleotide bases. Mature tRNA molecules exhibit a greater diversity of covalent modifications than any other class of RNA molecule. Typically 26 to 30 of the approximately 80 nucleotides in a tRNA molecule are covalently modified. Each type of covalent modification usually occurs in only one location on each molecule. Some examples of the sites of modification of nucleotides are shown in Figure 21.25.

B. Ribosomal RNA Processing

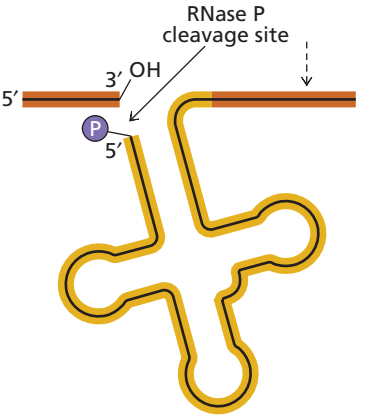
Ribosomal RNA molecules in all organisms are produced as large primary transcripts that require subsequent processing, including methylation and cleavage by endonucleases, before the mature molecules can adopt their active forms. This processing of ribosomal RNA is coupled to ribosome assembly.

The primary transcripts of prokaryotic rRNA molecules are about 30S in size and contain one copy each of the 16S, 23S, and 5S rRNAs. The transcripts also contain interspersed tRNA precursors. (Note that S is the symbol for the Svedberg unit, a measure of the rate at which particles move in the gravitational field established in an ultracentrifuge. Large S values are associated with large masses. The relationship between S and mass is not linear; therefore, S values are not additive.) Since the three rRNAs are derived from a single transcript, this processing ensures that there are equimolar amounts of each of the mature ribosomal RNAs.

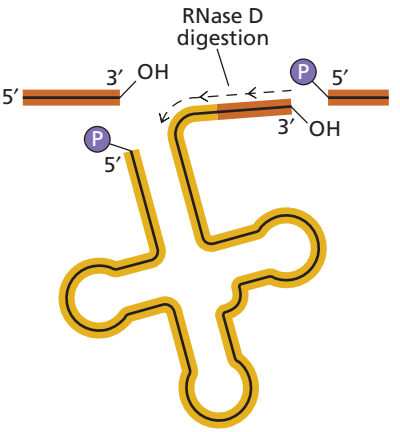
The 5′ and 3′ ends of each mature rRNA molecule are usually found in base-paired regions in the primary transcript. In prokaryotes, the endonuclease RNase III binds to these regions and cleaves the precursor near the ends of the 16S and 23S rRNAs. Following the initial cleavage, the ends of the rRNA molecules are trimmed by the actions of specific endonucleases (Figure 21.26).

Eukaryotic ribosomal RNAs are also produced by processing a larger precursor. The primary transcripts are between 35S and 47S in size and contain a copy of each of three eukaryotic rRNA species: 18S, 5.8S, and 28S. (The fourth eukaryotic rRNA, 5S rRNA, is transcribed as a monomer by RNA polymerase III and is processed separately.) The primary transcripts are synthesized in the region of the nucleus called the nucleolus, where initial processing occurs. Each rRNA precursor partially folds up and binds to some of its ribosomal protein partners before the processing cleavages take place.

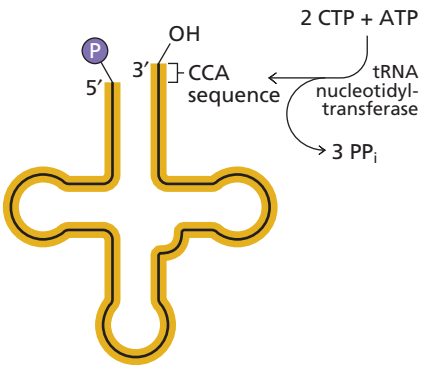
(a) RNase P and other endonucleases cleave the primary transcript.



(b) RNase D trims the 3′ end.



(c) tRNA nucleotidyl transferase adds CCA to the 3′ end.



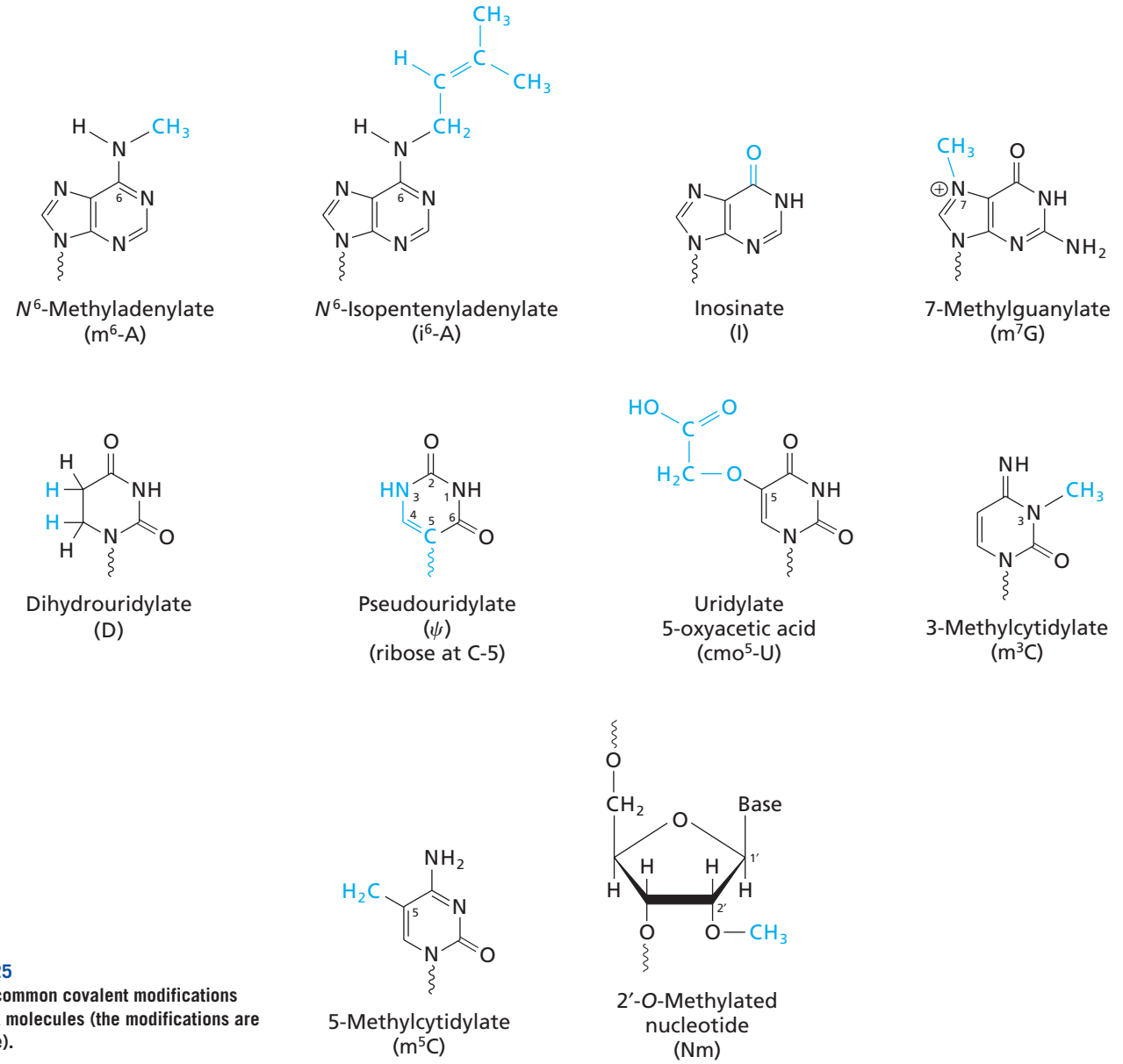
▲ **Figure 21.24**
Summary of prokaryotic tRNA processing.

21.9 Eukaryotic mRNA Processing

The processing of mRNA precursors is one of the biochemical features that distinguishes prokaryotes from eukaryotes. In prokaryotes, the primary mRNA transcripts are translated directly, often initiating translation before transcription is

KEY CONCEPT

Unmodified mRNAs are inherently unstable in a cell and would be rapidly degraded by ribonucleases.



▲ Figure 21.25
Examples of common covalent modifications found in tRNA molecules (the modifications are shown in blue).

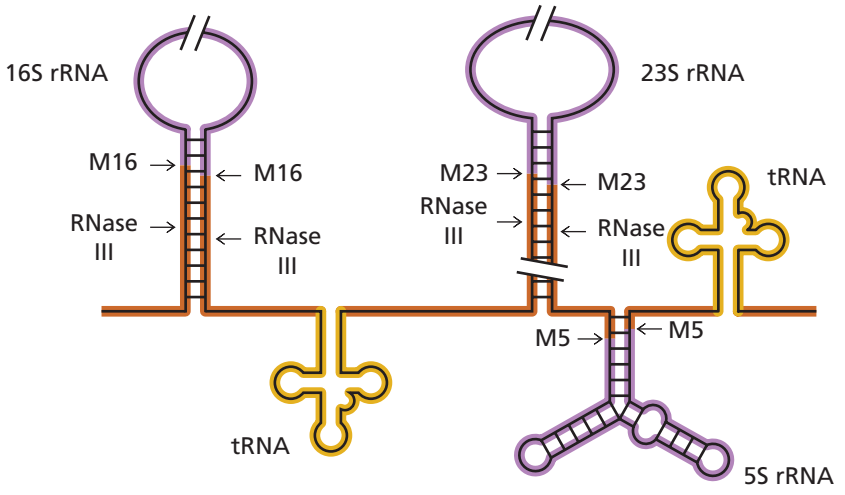


Figure 21.26 ►
Endonucleolytic cleavage of ribosomal RNA precursors in *E. coli*. The primary transcript contains a copy of each of the three ribosomal RNAs and may also contain several tRNA precursors. The large rRNA precursors are cleaved from the large primary transcript by the action of RNase III. The ends of the 16S, 23S, and 5S rRNAs are trimmed by the action of endonucleases M16, M23, and M5, respectively. (Slash marks indicate that portions of the rRNA primary transcript have been deleted for clarity.)

complete. In eukaryotes, on the other hand, transcription occurs in the nucleus, and translation takes place in the cytoplasm. This compartmentalization of functions in eukaryotic cells allows nuclear processing of mRNA precursors without disrupting translation.

Mature eukaryotic mRNA molecules are often derived from much larger primary transcripts. Subsequent processing of these primary transcripts includes some of the same steps that we saw in the previous section, namely: cleavage of a precursor, addition of terminal nucleotides, and covalent modification of nucleotides. Often, specific nucleotides (called intervening sequences, or introns) from the middle of an mRNA primary transcript are actually excised, or removed, and the resulting fragments are ligated together to produce the mature mRNA. This step, called **splicing**, is common in most eukaryotic species. Splicing also occurs during the processing of some eukaryotic tRNA and rRNA precursors (although these post-transcriptional modifications use a different splicing mechanism).

A. Eukaryotic mRNA Molecules Have Modified Ends

All eukaryotic mRNA precursors undergo modifications that increase the stability of the mature mRNAs and make them better substrates for translation. One way to increase the stability of mRNAs is to modify their ends so that they are no longer susceptible to cellular exonucleases that degrade RNA.

The 5′ ends are modified while the mRNA precursors are still being synthesized. The 5′ end of the primary transcript is a nucleoside triphosphate residue (usually a purine) that was the first nucleotide incorporated by RNA polymerase II. Modification of this end begins when the gamma-phosphate group is removed by the action of a phosphohydrolase (Figure 21.27). The resulting 5′-diphosphate group then reacts with the *a*-phosphorus atom of a GTP molecule to create a 5′–5′ triphosphate linkage. This reaction is catalyzed by guanylyltransferase and produces a structure called a **cap**. The cap is often further modified by methylating the newly added guanylate. The 2′-hydroxyl groups of the first two nucleotides in the original transcript may also be methylated. Methyl groups for these reactions are donated by *S*-adenosylmethionine (Section 7.3).

KEY CONCEPT

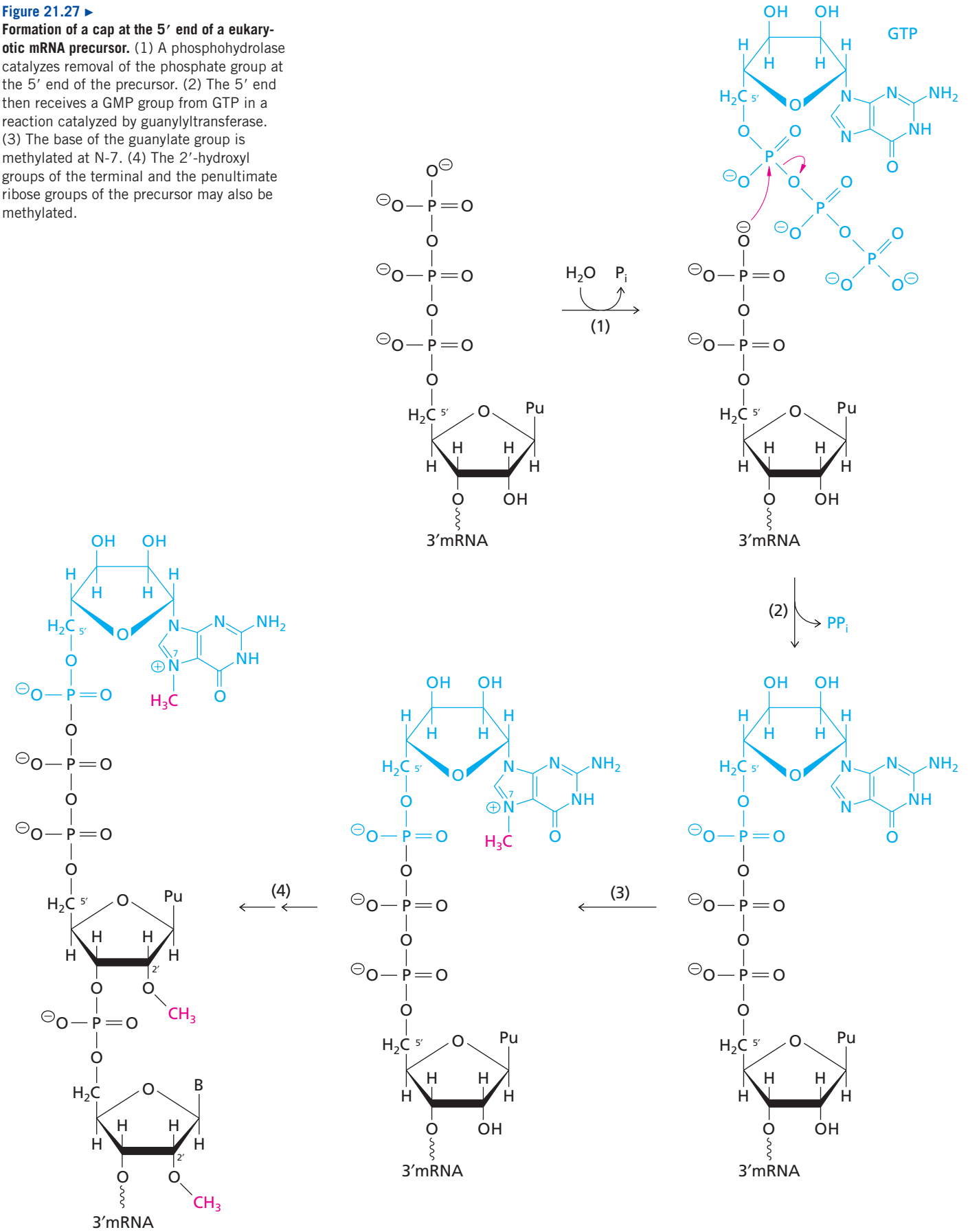
Many eukaryotic coding sequences are interrupted by introns.

The 5′–5′ triphosphate linkage protects the mRNA molecule from 5′ exonucleases by blocking its 5′ end. The cap also converts mRNA precursors into substrates for other processing enzymes in the nucleus, such as those that catalyze splicing. In mature mRNA, the cap is the site where ribosomes attach during protein synthesis. Capping is a cotranscriptional process that is confined to the nucleus. The capping enzymes shown in Figure 21.27 interact directly with RNA polymerase II transcription complexes but not with RNA polymerase I or RNA polymerase III complexes, ensuring that mRNA precursors are the only capped RNAs (i.e., tRNA and rRNA are not substrates for capping).

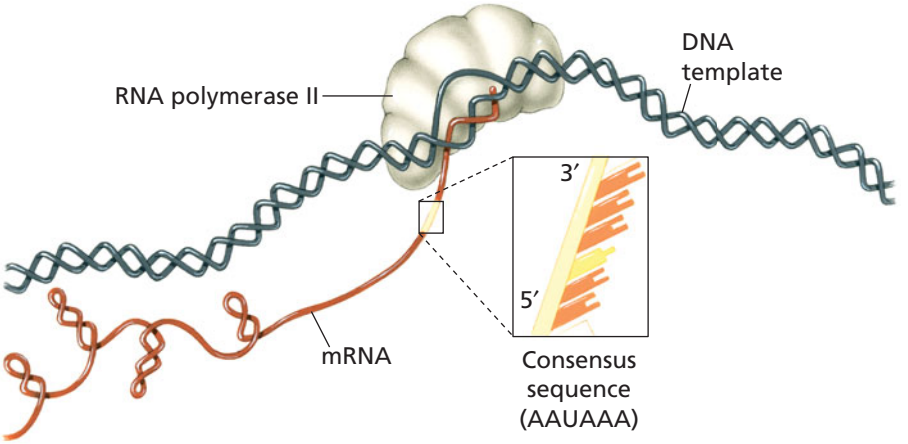
Eukaryotic mRNA precursors are also modified at their 3′ ends. Once RNA polymerase II has transcribed past the 3′ end of the coding region of DNA, the newly synthesized RNA is cleaved by an endonuclease downstream of a specific site whose consensus recognition sequence is AAUAAA. This sequence is bound by a cleavage and polyadenylation specificity factor (CPSF), a protein that also interacts with the endonuclease and a polymerase (Figure 21.28). After cleaving the RNA, the endonuclease dissociates and multiple adenylate residues are added to the newly generated 3′ end of the molecule. The addition reactions are catalyzed by poly A polymerase, which adds adenylate residues using ATP as a substrate. Up to 250 nucleotides can be added to form a stretch of polyadenylate known as a **poly A tail**.

With a few rare exceptions, all mature mRNA molecules in eukaryotes contain poly A tails. The length of the tail varies, depending on the species and possibly on the type of mRNA and the developmental stage of the cell. The length also depends on the age of the mRNA since the poly A tail is progressively shortened by the action of 3′ exonucleases. In fact, the tail has already been shortened by 50 to 100 nucleotides by the time the mature mRNA reaches the nuclear pores. The presence of the poly A tail increases the time required for the exonucleases to reach the coding region.

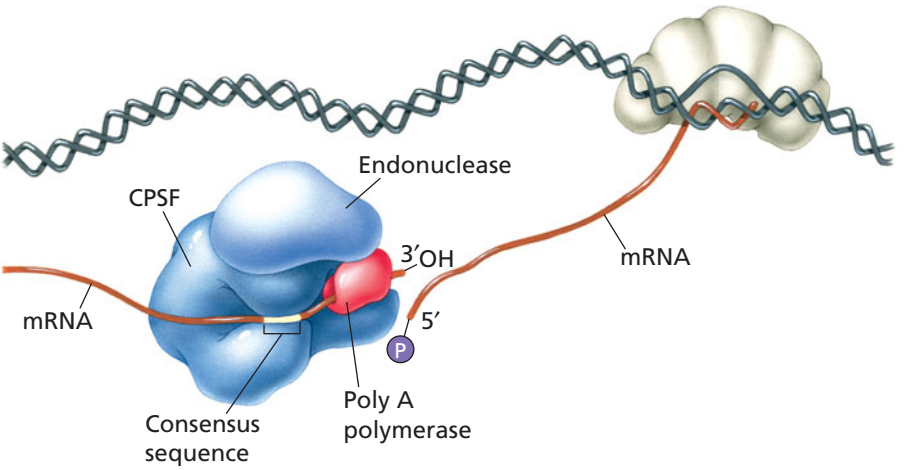
Figure 21.27 ▶
Formation of a cap at the 5′ end of a eukaryotic mRNA precursor. (1) A phosphohydrolase catalyzes removal of the phosphate group at the 5′ end of the precursor. (2) The 5′ end then receives a GMP group from GTP in a reaction catalyzed by guanylyltransferase. (3) The base of the guanylate group is methylated at N-7. (4) The 2′-hydroxyl groups of the terminal and the penultimate ribose groups of the precursor may also be methylated.



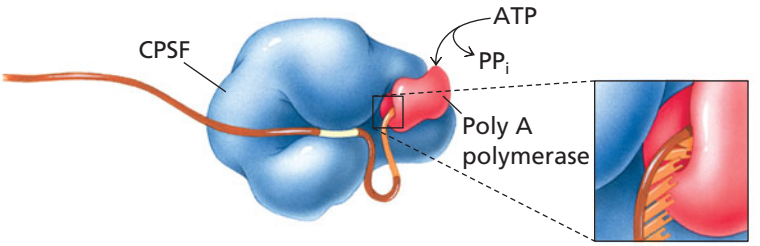
(a) Polyadenylation begins when RNA polymerase II transcription complex transcribes through a polyadenylation signal at the 3′ end of an mRNA precursor.



(b) CPSF binds to the consensus sequence and forms a complex containing an RNA endonuclease. The endonuclease catalyzes cleavage of the transcript downstream of the polyadenylation sequence, forming a new 3′ end. Poly A polymerase can then bind to the end of the mRNA precursor.



(c) The endonuclease dissociates and the new 3′ end of the RNA is polyadenylated by the activity of poly A polymerase.

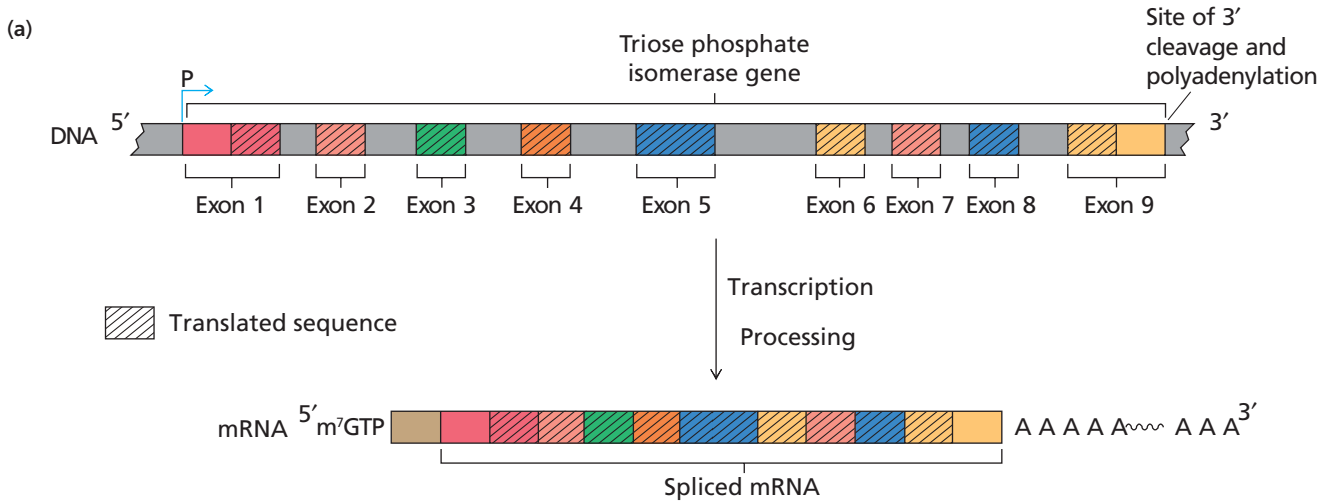


▲ **Figure 21.28**
Polyadenylation of a eukaryotic mRNA precursor.

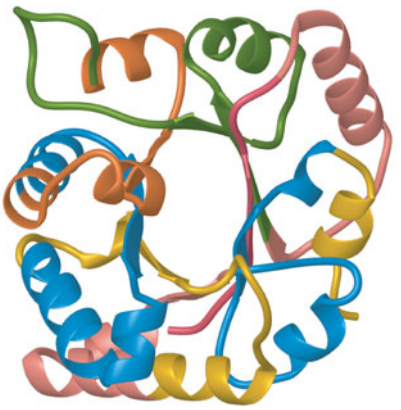
B. Some Eukaryotic mRNA Precursors Are Spliced

Splicing is rare in prokaryotes but it is the rule in animals and flowering plants. Internal sequences that are removed from the primary RNA transcript are called **introns**. Sequences that are present in the primary RNA transcript and in the mature RNA molecule are called **exons**. The words *intron* and *exon* also refer to the regions of the gene (DNA) that encode corresponding RNA introns and exons. Since DNA introns are transcribed, they are considered part of the gene. The junctions of introns and exons are known as **splice sites** since these are the sites where the mRNA precursor is cut and joined.

Because of the loss of introns, mature mRNA is often a fraction of the size of the primary transcript. For example, the gene for triose phosphate isomerase from maize contains nine exons and eight introns and spans over 3400 bp of DNA. The mature



(b)

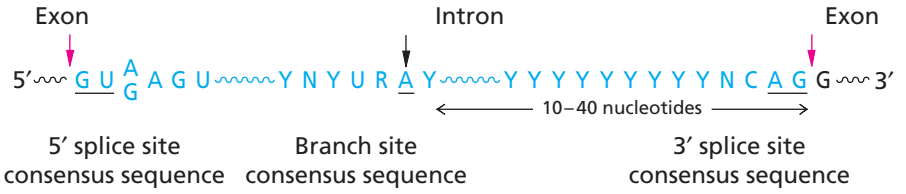


▲ **Figure 21.29**
Triose phosphate isomerase gene from maize and the encoded enzyme. (a) Diagram of the gene showing nine exons and eight introns. Some exons contain both translated and untranslated sequences. **(b)** Three-dimensional structure of the protein showing the parts of the protein encoded by each exon.

mRNA, which includes a poly A tail, is only 1050 nucleotides long (Figure 21.29). The enzyme itself contains 253 amino acid residues.

It used to be thought that there was a correlation between the intron/exon organization of a gene and the structure of the protein that the gene encodes. According to this hypothesis, exons encode protein domains and the presence of introns reflects the primitive organization of the gene. In other words, introns arose early in evolution. However, as shown in Figure 21.29b, there is no obvious correlation between exons and protein structure. Most biochemists and molecular biologists now believe that introns have been inserted at random locations during the evolution of a gene. The "introns late" hypothesis states that most primitive genes did not have introns and postulates that introns arose much later during the evolution of eukaryotes.

Introns can vary in length from as few as 42 bp to as many as 10,000 bp (the lower limit varies with each species; for example, most *C. elegans* introns are too small to be accurately spliced in either a vertebrate cell or cell-free extract). The nucleotide sequences at splice sites are similar in all mRNA precursors, but the sequence of the rest of the intron is not conserved. The vertebrate consensus sequences at the two splice sites are shown in Figure 21.30. Another short consensus sequence is found within the intron near the 3′ end. This sequence, known as the **branch site** or the branch-point sequence, also plays an important role in splicing.



◀ **Figure 21.30**
Consensus sequences at splice sites in vertebrates. Highly conserved nucleotides are underlined. Y represents a pyrimidine (U or C), R represents a purine (A or G), and N represents any nucleotide. The splice sites, where the RNA precursor is cut and joined, are indicated by red arrows, and the branch site is indicated by a black arrow. The intron is highlighted in blue.

The splicing of an mRNA precursor to remove a single intron requires two transesterification reactions: one between the 5′ splice site and the branch-site adenylate residue, and one between the 5′ exon and the 3′ splice site. The products of these two reactions are (1) the joined exons and (2) the excised intron in the form of a lariat-shaped molecule. These splicing reactions are catalyzed by a large RNA-protein complex called the **spliceosome**. The spliceosome helps to not only retain the intermediate splicing products but also position the splice sites so that the exons can be precisely joined (Figure 21.31).

The spliceosome is a large, multisubunit complex. It contains over 100 proteins and five molecules of RNA whose total length is about 5000 nucleotides. These RNA molecules are called small nuclear RNA (snRNA) molecules and are associated with proteins to form small nuclear ribonucleoproteins, or snRNPs (pronounced "snurps"). snRNPs are important not only in the splicing of mRNA precursors but also in other cellular processes.

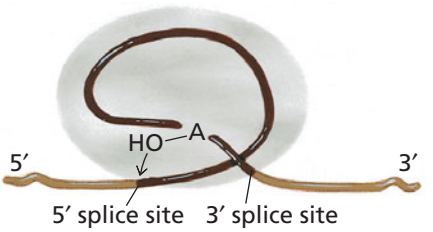
There are five different types of snRNAs—U1, U2, U4, U5, and U6. (U stands for uracil, a common base in these small RNA molecules.)—and a diploid vertebrate nucleus contains more than 100,000 total copies of snRNA. All five snRNA molecules are extensively base-paired and contain modified nucleotides. Each snRNP contains one or two snRNAs plus a number of proteins. Some of these proteins are common to all snRNPs; others are found in only one class of snRNP.

Biochemical experiments *in vitro* using purified components have led to a sequential model for spliceosome assembly (Figure 21.32). Spliceosome formation begins when a U1 snRNP binds to the newly synthesized 5′ splice site of the mRNA precursor. This interaction depends on base pairing between the 5′ splice site and a complementary sequence near the 5′ end of the U1 snRNA. A U2 snRNP then binds to the branch site of the intron, forming a stable complex that covers about 40 nucleotides. Next, a U5 snRNP associates with the 3′ splice site. Finally, a U4/U6 snRNP joins the complex, and all snRNPs are drawn together to form the spliceosome. Because several groups have now discovered that these same snRNPs are found preassembled in a much larger complex, prior to splicing, this pathway may not accurately reflect the splicing cycle *in vivo*.

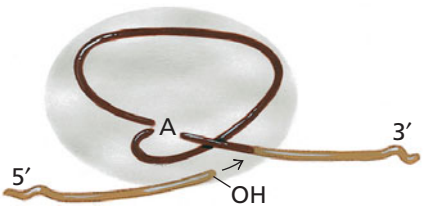
Binding of the U1, U2, and U5 snRNPs to consensus sequences at the 5′ splice site, branch site, and 3′ splice site of the intron positions these three interactive sites properly for the splicing reaction. The spliceosome then prevents the 5′ exon from diffusing away after cleavage and positions it to be joined to the 3′ exon. Once a spliceosome has formed at an intron, it is quite stable and can be purified from cell extracts.

Since spliceosomes can be observed on nascent transcripts, it is thought that intron removal is the rate limiting step in RNA processing. Since the spliceosome, which is almost as large as a ribosome, is too large to fit through the nuclear pores, the mRNA precursors are prevented from leaving the nucleus before processing is complete. Once an intron is excised, the spliceosome gets recycled and will repeat the catalytic cycle on the next intron it encounters.

(a) The spliceosome positions the adenylate residue at the branch site near the 5′ splice site. The 2′-hydroxyl group of the adenylate attacks the 5′ splice site.



(b) The 2′-hydroxyl group is attached to the 5′ end of the intron, and the newly created 3′-hydroxyl group of the exon attacks the 3′ splice site.



(c) As a result, the ends of the exons are joined, and the intron, a lariat-shaped molecule, is released.

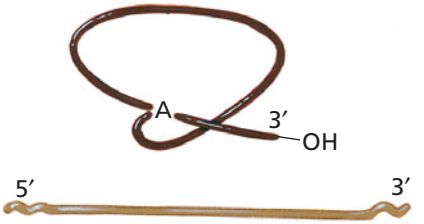
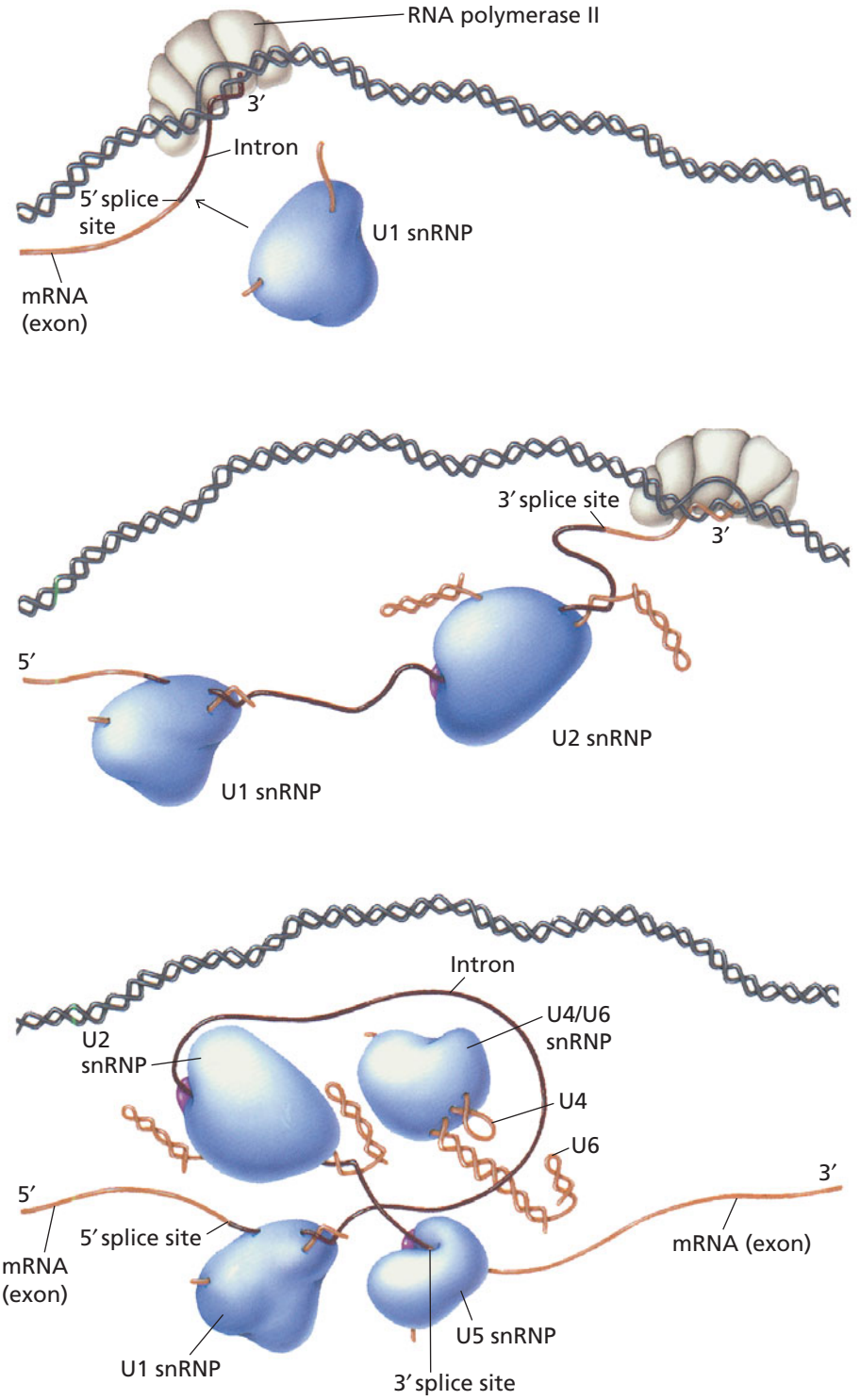


Figure 21.31 ▶
Intron removal in mRNA precursors. The spliceosome, a multicomponent RNA-protein complex, catalyzes splicing.
[Adapted from Sharp, P. A. (1987). Splicing of messenger RNA precursors. Science 235:766–771.]



◀ **Figure 21.32**
Formation of a spliceosome.

(a) As soon as the 5′ splice site exits the transcription complex, a U1 snRNP binds to it.

(b) Next, a U2 snRNP binds to the branch site within the intron.

(c) When the 3′ splice site emerges from the transcription complex, a U5 snRNP binds, and the complete spliceosome assembles around a U4/U6 snRNP.

Summary

1. A gene is a sequence of DNA that is transcribed. Housekeeping genes encode proteins and RNA molecules that are essential for normal cellular activities.
2. Cells contain several types of RNA, including transfer RNA, ribosomal RNA, messenger RNA, and small RNA molecules.
3. DNA-directed RNA synthesis, or transcription, is catalyzed by RNA polymerase. Ribonucleoside triphosphates are added in nucleotidyl-group–transfer reactions using a DNA strand as a template.
4. Transcription begins at a promoter sequence and proceeds in the 5′ → 3′ direction. A promoter consensus sequence indicates the nucleotides most commonly found at each position. The σ subunit of *E. coli* RNA polymerase increases the affinity of the core polymerase for a promoter and decreases the affinity for nonpromoter sequences. During initiation, a transcription bubble forms and a short stretch of RNA is synthesized. The σ subunit dissociates in the transition from initiation to chain elongation.

5. Transcription termination in *E. coli* occurs near pause sites, often when the RNA forms a hairpin structure. Some terminations require *rho*, which binds to single-stranded RNA.
6. In eukaryotes, several different RNA polymerases carry out transcription. Transcription factors interact with the promoter and RNA polymerase to initiate transcription.
7. Some genes are expressed constitutively, but the transcription of other genes is regulated. Transcription may be regulated by a repressor or an activator. These are often allosteric proteins.
8. Transcription of the three genes of the *lac* operon is blocked when *lac* repressor binds to two operators near the promoter. The repressor dissociates from the DNA when it binds the inducer allolactose. Transcription is activated by a complex of cAMP and CRP (cAMP regulatory protein).
9. RNA transcripts are frequently modified by processing, which includes the removal, addition, or modification of nucleotide residues. Primary transcripts of prokaryotic tRNA and rRNA are processed by nucleolytic cleavage and covalent modification.
10. Processing of mRNA in eukaryotes includes the addition of a 5′ cap and a 3′ poly A tail to protect the molecule from nuclease digestion. In some cases, introns are removed by splicing. The two transesterification reactions of splicing are catalyzed by the spliceosome, a complex containing small nuclear ribonucleoproteins (snRNPs).

Problems

1. A bacterial RNA polymerase elongates RNA at a rate of 70 nucleotides per second, and each transcription complex covers 70 bp of DNA.
   (a) What is the maximum number of RNA molecules that can be produced per minute from a gene of 6000 bp? (Assume that initiation is not rate limiting.)
   (b) What is the maximum number of transcription complexes that can be bound to this gene at one time?
2. The *E. coli* genome is approximately 4600 kb in size and contains about 4000 genes. The mammalian genome is approximately 33×10^6 kb in size and contains at most 30,000 genes. An average gene in *E. coli* is 1000 bp long.
   (a) Calculate the percentage of *E. coli* DNA that is not transcribed.
   (b) Although many mammalian genes are larger than bacterial genes, most mammalian gene products are the same size as bacterial gene products. Calculate the percentage of DNA in exons in the mammalian genome.
3. There are a variety of methods that will allow you to introduce an intact eukaryotic gene (e.g., the triose phosphate isomerase gene) into a prokaryotic cell. Would you expect this gene to be properly transcribed by prokaryotic RNA polymerase? What about the converse situation, where an intact prokaryotic gene is introduced into a eukaryotic cell; will it be properly transcribed by a eukaryotic transcription complex?
4. Assume that, in a rare instance, a typical eukaryotic triose phosphate isomerase gene contains the correct sequences to permit accurate transcription in a prokaryotic cell. Would the resulting RNA be properly translated to yield the intact enzyme?
5. Describe how the rate of transcription of the *lac* operon is affected when *E. coli* cells are grown in the presence of (a) lactose plus glucose, (b) glucose alone, and (c) lactose alone.
6. In the promoter of the *E. coli lac* operon the −10 region has the sequence 5′-TATGTT-3′. A mutation named UV5 changes this sequence to 5′-TATAAT-3′ (see Figure 21.6). Transcription from the *lac* UV5 promoter is no longer dependent on the CRP-cAMP complex. Why?
7. When β-3^{32}P4-ATP is incubated with a eukaryotic cell extract that is capable of transcription and RNA processing, where does the label appear in mRNA?
8. Unlike DNA polymerase, RNA polymerase does not have proofreading activity. Explain why the lack of proofreading activity is not detrimental to the cell.
9. Mature mRNA from eukaryotic cells is often purified from other components in the cell with the use of columns containing oligo (dT) cellulose. These columns contain short segments of single-stranded deoxyribose thymidylate residues, oligo(dT), attached to a cellulose matrix. Explain the rationale for use of these columns to purify mature mRNA from a mixture of components.
10. Rifampicin is a semisynthetic compound made from rifamycin B, an antibiotic isolated from *Streptomyces mediterranei*. Rifampicin is an approved anti-mycobacterial drug that is a standard component of combination regimens for treating tuberculosis and staphylococci infections that resist penicillin. Recent studies have suggested that rifampicin-resistant tuberculosis is becoming more common. For example, 2% of samples from a survey in Botswana were found to be resistant to the drug. The table below gives some results from wild type *E. coli* and *E. coli* with a single amino acid change in the β subunit of RNA polymerase (Asp to Tyr at amino acid position 516) and their growth response to media that contained rifampicin. [Severinov, K., Soushko, M., Goldfarb, A., and Nikiforov, V. (1993). Rifampicin region revisited. *J. Biol. Chem.* 268:14820–14825].

E. coli	**Rifampicin[a] (μg/ml)**
Wild type	<5
Asp516Tyr in β subunit	>50

[a]Rifampicin concentration at the point of growth arrest of the *E. coli.*

   (a) What is your interpretation of the data?
   (b) What role does the β subunit have in RNA polymerase?
   (c) Describe one mechanism for rifampicin-resistant bacteria.
11. A segment of DNA from the middle of an *E. coli* gene has the sequence below. Write the mRNA sequences that can be produced by transcribing this segment in either direction.

CCGGCTAAGATCTGACTAGC

12. Does the definition of a gene given on page 638 5e [first page of Chapter 21] apply to the rRNA and tRNA genes whose primary transcript is shown in Figure 21.26?

13. In general, if we know the genomic DNA sequence of a gene we can reliably predict the nucleotide sequence of the RNA encoded by that gene. Is this statement also true for tRNAs in prokaryotes? What about tRNAs in eukaryotes?
14. Assume that a spliceosome assembles at the first intron of the gene for triose phosphate isomerase in maize (Figure 21.29) almost as soon as the intron is transcribed (i.e., after about 500 nucleotides of RNA have been synthesized). How long must the spliceosome be stable if the splicing reaction cannot occur until transcription terminates? Assume that the rate of transcription by RNA polymerase II in maize is 30 nucleotides per second.
15. CRP-cAMP represses transcription of the *crp* gene. Predict the location of the CRP-cAMP binding site relative to the promoter of the *crp* gene.
16. Why are mutations within an intron of a protein-coding gene sometimes detrimental?
17. A deletion in one of the introns in the gene for the triose phosphate isomerase moves the branch site to a new location seven nucleotides away from the 3′-splice acceptor sequence. Will this deletion have any affect on splicing of the gene?

Selected Readings

General

Alberts, B., Johnson, A., Lewis, J., and Raff, M. (2007). *Molecular Biology of the Cell*, 5th ed. (New York: Garland).

Krebs, J., Goldstein, L., and Kilpatrick, S. (2009). *Lewin's Genes X* (New York: Jones & Bartlett).

RNA Polymerases and Transcription

Ardehali, M. B., and Lis, J. T. (2009). Tracking rates of transcription and splicing *in vivo*. *Nature Structural & Molecular Biology* 16:1123–1124.

Bushnell, D. A., and Kornberg, R. D. (2003). Complete, 12-subunit RNA polymerase II and 4.1-A resolution: implications for the initiation of transcription. *Proc. Natl. Acad. Sci. (U.S.A.)* 100:6969–6973.

Kornberg, R. D. (1999). Eukaryotic transcriptional control. *Trends Cell Biol.* 9:M46–M49.

Lisser, S., and Margalit, H. (1993). Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.* 21:1507–1516.

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296:1285–1290.

Richardson, J. P. (1993). Transcription termination. *Crit. Rev. Biochem.* 28:1–30.

Regulation of Transcription

Becker, P. B., and Horz, W. (2002). ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* 71:247–273.

Bushman, F. D. (1992). Activators, deactivators and deactivated activators. *Curr. Biol.* 2:673–675.

Fuda, N. J., Behfar, M., and Lis, J. T. (2009). Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* 461:186–192.

Harrison, S. C., and Aggarwal, A. K. (1990). DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem.* 59:933–969.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3: 318–356.

Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. (1993). Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* 62:749–795.

Myers, L. C., and Kornberg, R. D. (2000). Mediator of transcriptional regulation. *Annu. Rev. Biochem.* 69:729–749.

Pan, Y., Tsai, C.-J., Ma, B., and Nussinov, R. (2009). How do transcription factors select specific binding sites in the genome? *Nature Structural & Molecular Biology* 16:1118–1120.

Wolfe, A. P., and Guschin, D. (2000). Review: chromatin structural features and targets that regulate transcription. *J. Struct. Biol.* 129:102–122.

Workman, J. L., and Kingston, R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* 67: 545–579.

RNA Processing

Apirion, D., and Miczak, A. (1993). RNA processing in prokaryotic cells. *BioEssays* 15:113–120.

Collins, C. A., and Guthrie, C. (2000). The question remains: is the spliceosome a ribozyme? *Nature Struct. Biol.* 7: 850–854.

James, B. D., Olsen, G. J., Liu, J., and Pace, N. R. (1988). The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell* 52:19–26.

Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. *Molecular Cell* 12:5–14.

McKeown, M. (1993). The role of small nuclear RNAs in RNA splicing. *Curr. Biol.* 5:448–454.

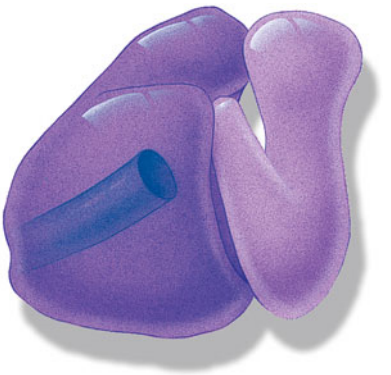
Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *BioEssays* 25:1147–1149.

Proudfoot, N. (2000). Connecting transcription to messenger RNA processing. *Trends Biochem. Sci.* 25:290–293.

Shatkin, A. J., and Manley, J. L. (2000). The ends of the affair: capping and polyadenylation. *Nature Struct. Biol.* 7: 838–842.

Wahle, E. (1992). The end of the message: 3′-end processing leading to polyadenylated messenger RNA. *BioEssays* 14:113–118.





Protein Synthesis

We are now ready to examine the final stage of biological information flow: the translation of mRNA and the polymerization of amino acids into proteins. The essential features of the biochemistry of protein synthesis were worked out in the decade between 1955 and 1965. It was clear that there was a genetic code that had to be used to translate a nucleotide sequence into a sequence of amino acids. In 1955, Francis Crick proposed that the first step in this process was the attachment of an amino acid to a small adapter RNA. Shortly after that, the adapters, now known as transfer RNAs, were identified. Ribosomes and the other essential components of the translation machinery were discovered by fractionating cells and reconstituting protein synthesis *in vitro*. Workers in several laboratories demonstrated that messenger RNA is one of the key intermediates in the flow of information from DNA to protein. By 1961, the most important missing ingredient was the nature of the genetic code.

We begin this chapter with a discussion of the genetic code and tRNA structure. Next, we examine how mRNA, tRNA, ribosomes, and accessory proteins participate in protein synthesis. We will also present some examples of the regulation of translation and post-translational processing.

The results indicate that polyuridylic acid contains the information for the synthesis of a protein having many of the characteristics of poly-L-phenylalanine. . . . One or more uridylic acid residues therefore appear to be the code for phenylalanine. Whether the code is of the singlet, triplet, etc., type has not yet been determined. Polyuridylic acid seemingly functions as a synthetic template or messenger RNA, and this stable, cell-free E. coli *system may well synthesize any protein corresponding to meaningful information contained in added RNA.*

—M. Nirenberg and H. Matthaei, 1961

22.1 The Genetic Code

George Gamow first proposed the basic structural units of the genetic code. He reasoned that since the DNA "alphabet" consists of only four "letters" (A, T, C, and G) and since these four letters encode 20 amino acids, the genetic code might contain "words," or **codons**, with a uniform length of three letters. Two-letter words constructed from any combination of the four letters produce a vocabulary of only 16 words (4^2), not enough for all 20 amino acids. In contrast, four-letter words produce a vocabulary of 256 words (4^4), far more than are needed. Three-letter words allow a possible vocabulary of 64 words (4^3), more than sufficient to specify each of the 20 amino acids but not excessive.

Top: *Escherichia coli* ribosome. The ribosome, a complex of RNA and protein, is the site where genetic information is translated into protein.



▲ **The enigma cryptography machine used by German armed forces during the Second World War.** This mechanical typewriter permitted the user to adjust its three large dials to encrypt outgoing messages before being sent by telegraph. The recipients could decode the message by setting the dials on their enigma machine to match. This type of encryption is extremely difficult to decipher, but when Allied forces were able to capture an intact enigma machine they could listen in on all their enemy's transmissions.

The "cracking" of the genetic code began with a chance observation by Marshall Nirenberg and J. Heinrich Matthaei. They discovered that polyuridylate (poly U) could direct the synthesis of polyphenylalanine *in vitro*. By showing that UUU encodes phenylalanine, they identified the first codon.

Between 1962 and 1965, the rest of the code was deciphered by a number of workers, chiefly Nirenberg and H. Gobind Khorana. Overall, it took ten years of hard work to learn how mRNA encodes proteins. The development of methods for sequencing genes and proteins has allowed direct comparison of the primary sequences of proteins with the nucleotide sequences of their corresponding genes. Each time a new protein and its gene are characterized, the genetic code is confirmed.

Transfer RNA (tRNA) plays an important role in interpreting the genetic code and translating a nucleotide sequence into an amino acid sequence. tRNAs are the adapters between mRNA and proteins. One region of a tRNA molecule is covalently linked to a specific amino acid, while another region on the same tRNA molecule interacts directly with an mRNA codon by complementary base pairing. It is this processive joining of the amino acids specified by an mRNA template that allows the precise synthesis of proteins.

In principle, a genetic code made up of three-letter words can be either overlapping or nonoverlapping (Figure 22.1). If the codons overlap, then each letter is part of more than one word and mutating a single letter changes several words simultaneously. For example, in the sequence shown in Figure 22.1a, each letter is part of three different words in an overlapping code. One of the advantages of a nonoverlapping code (Figure 22.1b) is that each letter is part of only one word; therefore, mutating a single nucleotide affects only one codon. All living organisms use a nonoverlapping genetic code.

Even with a nonoverlapping code, a sequence can be translated in many different ways, depending on where translation begins. (We will see later that translation does not typically begin with the very first nucleotide in an mRNA.) Each potential translation initiation point defines a unique sequence of three-letter words, or **reading frame**, in the mRNA. The correct translation of the "message" transcribed, or written, in the genetic code depends on establishing the correct reading frame for translation (Figure 22.2).

The standard genetic code is shown in Figure 22.3. With a few minor exceptions, all living organisms use this genetic code, suggesting that all modern species are descended from a common ancestor that also used the standard genetic code. This ancestral species probably lived billions of years ago, making the genetic code one of the most ancient remnants of early life.

By convention, all nucleotide sequences are written in the 5′ → 3′ direction. Thus, UAC specifies tyrosine, and CAU specifies histidine. The term *codon* usually refers to triplets of nucleotides in mRNA but it can also apply to triplets of nucleotides in the DNA sequence of a gene. For example, one DNA codon for tyrosine is TAC.

Codons are always translated 5′ → 3′, beginning near the 5′ end of the message (i.e., the end synthesized first) and proceeding to the end of the coding region that is

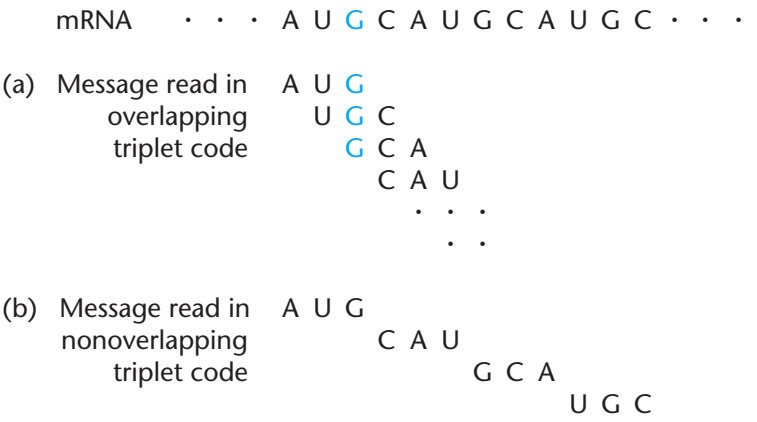


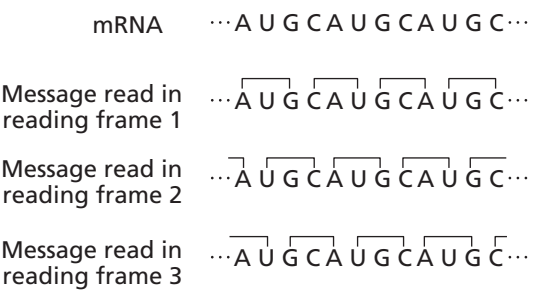
Figure 22.1 ►
Message read in (a) overlapping and (b) nonoverlapping three-letter codes. In an overlapping code, each letter is part of three different three-letter words (as indicated for the letter G in blue); in a nonoverlapping code, each letter is part of only one three-letter word.

usually near the 3′ end of the mRNA. The correct reading frame is specified by special punctuation signals that mark the beginning and the end.

The standard genetic code has several prominent features:

1. The genetic code is unambiguous. In a particular organism or organelle each codon corresponds to one, and only one, amino acid.
2. There are multiple codons for most amino acids. For example, leucine is the most abundant amino acid found in proteins (Table 3.3) and has six codons. Because of the existence of several codons for most amino acids, the genetic code is said to be **degenerate**. Different codons that specify the same amino acid (e.g., UCU and CGU both specify Ser; ACA, ACC, ACG, and ACU all specify Thr) are known as **synonymous codons**.
3. The first two nucleotides of a codon are often enough to specify a given amino acid. For example, the four codons for glycine (GGU, GGC, GGA, and GGG) all begin with GG.
4. Codons with similar sequences specify chemically similar amino acids. For example, the codons for threonine differ from four of the codons for serine by only a single nucleotide at the 5′ position and the codons for aspartate and glutamate begin with GA and differ only at the 3′ position. In addition, codons with pyrimidines at the second position usually encode hydrophobic amino acids. Therefore, mutations that alter either the 5′ or the 3′ position of these codons usually result in the incorporation of a chemically similar amino acid into the protein.
5. Only 61 of the 64 codons specify amino acids. The three remaining codons (UAA, UGA, and UAG) are **termination codons**, or **stop codons**. Termination codons are not normally recognized by any tRNA molecules in the cell. Instead, they are recognized by specific proteins that cause newly synthesized peptides to be released from the translation machinery. The methionine codon, AUG, also specifies the initiation site for protein synthesis and is often called the **initiation codon**.

Since the completion of the first draft of the human genome in 2000, it has been common to read in the popular press of "deciphering the code of life" or "unlocking the human genetic code." Strictly speaking, the information in the human genome is encoded using the same "universal" genetic code discovered 50 years ago. Sequencing projects actually reveal the *messages* encoded by the genes and not the code itself.



▲ **Figure 22.2**
One mRNA contains three different reading frames. The same string of letters read in three different reading frames will be translated into three different "messages" or protein sequences. Thus, translation of the correct message requires selecting the correct reading frame.

INTERNATIONAL MORSE CODE
Time of dash equals three dots

A	• —	N	— •	1	• — — — —
B	— • • •	O	— — —	2	• • — — —
C	— • — •	P	• — — •	3	• • • — —
D	— • •	Q	— — • —	4	• • • • —
E	•	R	• — •	5	• • • • •
F	• • — •	S	• • •	6	— • • • •
G	— — •	T	—	7	— — • • •
H	• • • •	U	• • —	8	— — — • •
I	• •	V	• • • —	9	— — — — •
J	• — — —	W	• — —	0	— — — — —
K	— • —	X	— • • —		
L	• — • •	Y	— • — —		
M	— —	Z	— — • •		

▲ **Morse code permitted text to be sent by telegraph.** Messages written in the Latin alphabet and/or Arabic numerals could be transmitted via electrical wires using a code invented by Samuel Morse. In the Morse code the most common letters in English language text are coded by the shortest sequence of dashes and dots (allowing messages to be sent with the fewest number of symbols).

First position (5′ end)	U	C	A	G	Third position (3′ end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

◀ **Figure 22.3**
Standard genetic code. The standard genetic code is composed of 64 triplet codons. The left-hand column indicates the nucleotide found at the first (5′) position of the codon; the top row indicates the nucleotide found at the second (middle) position of the codon; and the right column indicates the nucleotide found at the third (3′) position of the codon. The codon AUG specifies methionine (Met) and is also used to initiate protein synthesis. STOP indicates a termination codon.

22.2 Transfer RNA

Transfer RNA molecules are the interpreters of the genetic code. They are the crucial link between the sequence of nucleotides in mRNA and the sequence of amino acids in the corresponding polypeptide. In order for tRNA to fulfill this role, every cell must contain at least 20 different tRNA species (one for every amino acid) and each tRNA must recognize at least one codon.

A. The Three-Dimensional Structure of tRNA

The nucleotide sequences of different tRNA molecules from many organisms have been determined. The sequences of almost all these molecules are compatible with the secondary structure shown in Figure 22.4. This "cloverleaf" structure contains several arms that are composed of a loop or a loop with a hydrogen-bonded stem. The double-stranded region of each arm forms a short, stacked, right-handed helix similar to that of double-stranded DNA.

The 5′ end and the region near the 3′ end of the tRNA molecule are base-paired to each other forming the *acceptor stem* (or amino acid stem). The activated amino acid will be covalently attached to tRNA on the 3′ end of this stem. The amino acid's carboxyl group gets linked to the terminal adenylate's ribose on either its 2′- or 3′-hydroxyl group (Recall from Section 21.8A that mature tRNA molecules are produced by processing a larger primary transcript and that the nucleotides at the 3′ end of a mature tRNA molecule are invariably CCA.) All tRNA molecules have a phosphorylated nucleotide on the 5′ end.

The single-stranded loop opposite the acceptor stem in the cloverleaf structure is called the anticodon loop. It contains the **anticodon**, the three-base sequence that binds to a complementary codon in mRNA. The arm of the tRNA molecule that contains the anticodon is called the *anticodon arm.* The remaining two arms of the tRNA molecule are named for the covalently modified nucleotides found within them. (See Figure 21.25

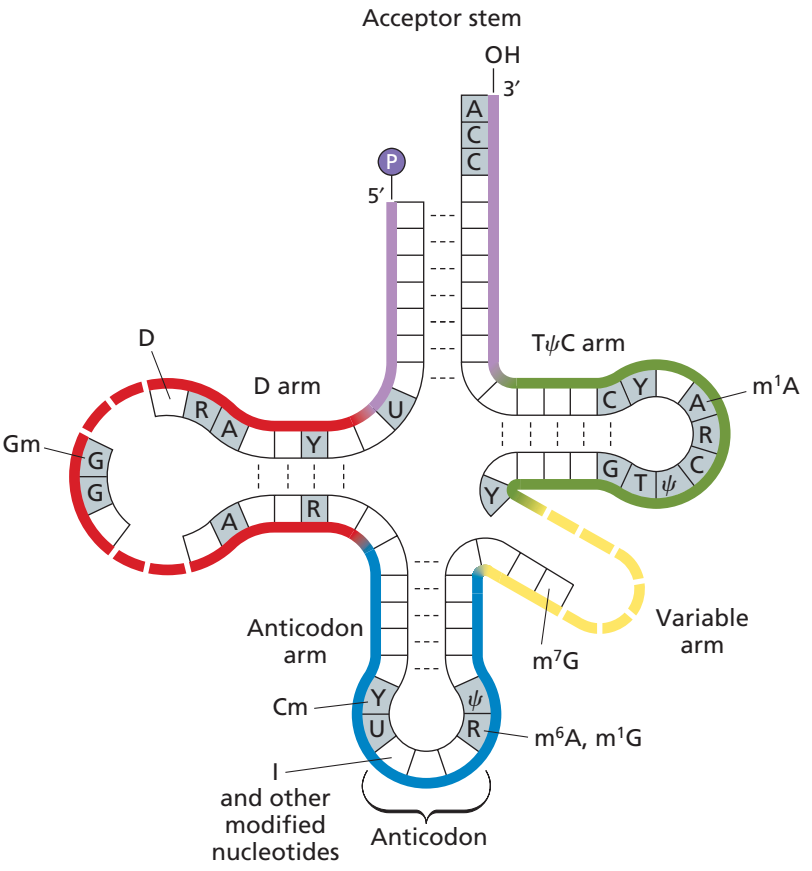
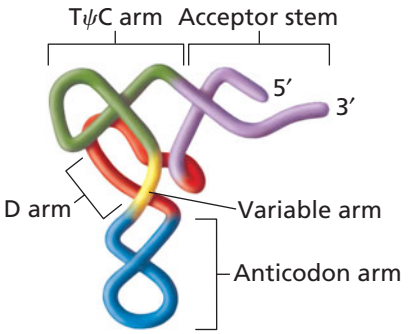


Figure 22.4 ►
Cloverleaf secondary structure of tRNA. Watson-Crick base pairing is indicated by dashed lines between nucleotide residues. The molecule is divided into an acceptor stem and four arms. The acceptor stem is the site of amino acid attachment, and the anticodon arm is the region of the tRNA molecule that interacts with mRNA codons. The D and TψC arms are named for modified nucleotides that are conserved within these arms. The number of nucleotide residues in each arm is more or less constant (except in the variable arm). Conserved bases (gray) and positions of common modified nucleotides are noted. Abbreviations other than standard nucleotides: R, a purine nucleotide; Y, a pyrimidine nucleotide; m^1A, 1-methyladenylate; m^6A, N^6-methyladenylate; Cm, 2′-O-methylcytidylate; D, dihydrouridylate; Gm, 2′-O-methylguanylate; m^1G, 1-methylguanylate; m^7G, 7-methylguanylate; I, inosinate; ψ, pseudouridylate; T, thymine ribonucleotide.

for the structures of these nucleotides.) One of the arms, called the TψC arm, always contains the triplet sequence ribothymidylate (T), pseudouridylate (ψ), and cytidylate (C). Dihydrouridylate (D) residues lend their name to the *D arm*. tRNA molecules also have a *variable arm* between the anticodon arm and the *TψC arm*. The variable arm ranges in length from about 3 to 21 nucleotides. With a few rare exceptions, tRNA molecules are between 73 and 95 nucleotides long.

The cloverleaf diagram of tRNA is a two-dimensional representation of a three-dimensional molecule. In three dimensions, the tRNA molecule is folded into a sideways "L" shape (Figures 22.5 and 22.6). The acceptor stem is at one end of the L-shaped molecule, and the anticodon is located in a loop at the opposite end. The resulting structure is compact and very stable, in part because of hydrogen bonds between the nucleotides in the D, TψC, and variable arms. This base pairing differs from normal Watson-Crick base pairing. Most of the nucleotides in tRNA are part of two perpendicular stacked helices. The interactions between the adjacent stacked base pairs are additive and make a major contribution to tRNA stability (analogous to the role of base stacking interactions in the 3D structure of double-stranded DNA we described in Section 19.2C).

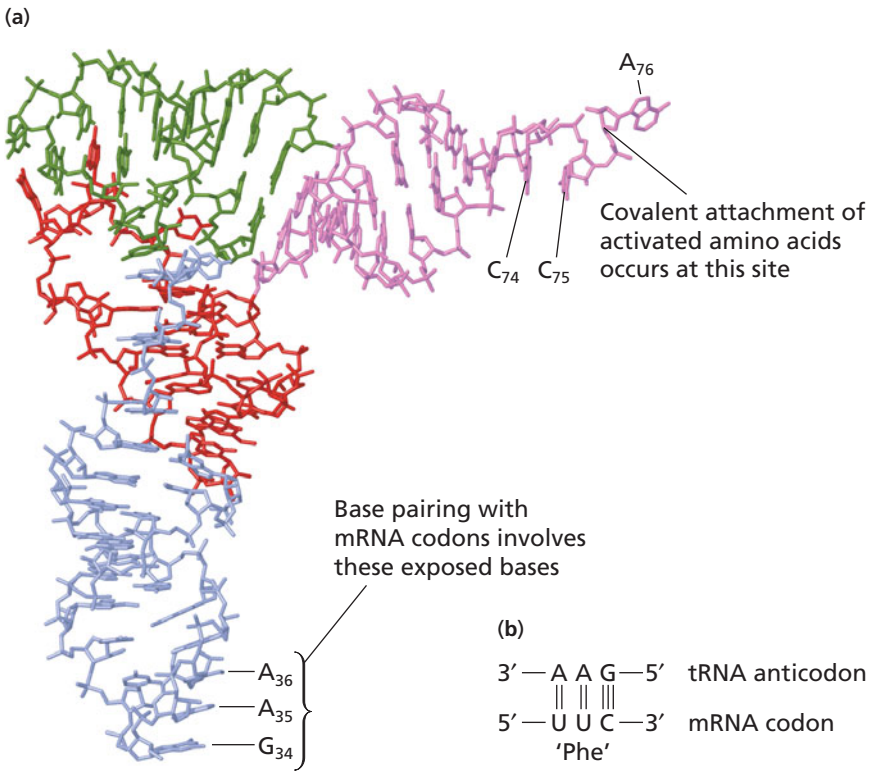


▲ **Figure 22.5**
Tertiary structure of tRNA. The cloverleaf-shaped molecule shown in Figure 22.4 actually folds up into this three-dimensional shape. The tertiary structure of tRNA results from base pairing between the TψC loop and the D loop, and two stacking interactions that **(a)** align the TψC arm with the acceptor arm, and **(b)** align the D arm with the anticodon arm. For clarity, only the ribose-phosphate backbone is shown here.

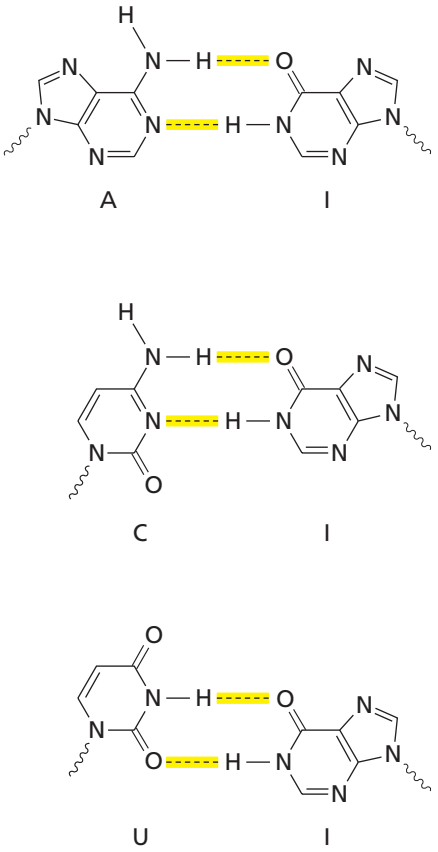
B. tRNA Anticodons Base-Pair with mRNA Codons

tRNA mediated decoding of the information stored in mRNA molecules requires base-pairing interactions between tRNA anticodons and complementary mRNA codons. The anticodon of a tRNA molecule therefore determines where the amino acid attached to its acceptor stem is added to a growing polypeptide chain. Transfer RNA molecules are named for the amino acid they carry. For example, the tRNA molecule shown in Figure 22.6 has the anticodon GAA that binds to the phenylalanine codon UUC. Prior to protein synthesis, phenylalanine is covalently attached to the acceptor stem of this tRNA. The molecule is therefore designated $tRNA^{Phe}$.

Much of the base pairing between the codon and the anticodon is governed by the rules of Watson-Crick base pairing: A pairs with U, G pairs with C, and the strands in the base-paired region are antiparallel. However, some exceptions to these rules led Francis



◀ **Figure 22.6**
Structure of $tRNA^{Phe}$ from the yeast *Saccharomyces cerevisiae*. **(a)** Stick model showing base pairs and the position of the D arm (red) relative to the TψC arm (green). Note that there are two double-stranded RNA helices arrayed at right angles to each other to form an L-shaped structure. **(b)** Diagram showing the complement base-pairing between $tRNA^{Phe}$ and a phe codon to generate a double-stranded, antiparallel RNA helix during decoding. [NDB TRNA10].



▲ **Figure 22.7**
Inosinate base pairs. Inosinate (I) is often found at the 5′ (wobble) position of a tRNA anticodon. Inosinate can form hydrogen bonds with A, C, or U. This versatility in hydrogen bonding allows a tRNA carrying a single anticodon to recognize more than one synonymous codon.

Table 22.1 Predicted base pairing between the 5′ (wobble) position of the anticodon and the 3′ position of the codon

Nucleotide at 5′ (wobble) position of anticodon	Nucleotide at 3′ position of codon
C	G
A	U
U	A or G
G	U or C
I[a]	U, A, or C

[a] I = Inosinate.

Crick to propose that complementary Watson-Crick base pairing is required for only two of the three base pairs formed. The codon must form Watson-Crick base pairs with the 3′ and middle bases of the anticodon but other types of base pairing are permitted at the 5′ position of the anticodon. This alternate pairing suggests that the 5′ position is conformationally flexible. Crick dubbed this flexibility "wobble" and the 5′ position of the anticodon is sometimes called the wobble position.

Table 22.1 summarizes the allowable base pairs between the wobble position of an anticodon and the third nucleotide of an mRNA codon. When G is at the wobble position, for example, it can pair with either C or U (!). The base at the wobble position of many anticodons is covalently modified permitting additional flexibility in codon recognition. For example, in several tRNA molecules, G at the 5′ anticodon position is deaminated at C-2 to form inosinate (I), which can hydrogen-bond with A, C, or U (Figure 22.7). The presence of I at the 5′ position of the anticodon explains why $tRNA^{Ala}$ with the anticodon IGC can bind to three different codons specifying alanine: GCU, GCC, and GCA (Figure 22.8).

Wobble allows some tRNA molecules to recognize more than one codon but several different tRNA molecules are often required to recognize all synonymous codons. Different tRNA molecules that can attach to the same amino acid are called **isoacceptor tRNA molecules**. The term *isoacceptor* describes not only tRNA molecules with different anticodons that are covalently attached to the same activated amino acid but also tRNA molecules with the same anticodon but different primary sequences. Isoacceptor tRNAs are identified by Roman numerals or by the codons they recognize (i.e., $tRNA_I^{Ala}$, $tRNA_{II}^{Ala}$, or $tRNA_{GCG}^{Ala}$).

Genome sequencing data reveal that bacterial genomes encode 30 to 60 different tRNAs and that eukaryotic genomes have genes for as many as 80 different tRNA molecules. Many of the eukaryotic tRNA genes are present in multiple copies, especially those genes that encode abundant tRNAs used most frequently in protein synthesis.

22.3 Aminoacyl-tRNA Synthetases

Like DNA and RNA synthesis, protein synthesis can be divided into three distinct stages: initiation, chain elongation, and termination. However, our description of translation includes a step prior to the initiation of polymerization, namely, aminoacylation of tRNA. The activation of amino acids is considered part of the overall translation process because it is such an important part of the flow of biological information from nucleic acid to protein.

Each of the 20 amino acids is covalently attached to the 3′ end of its respective tRNA molecules. The product of this reaction is called an **aminoacyl-tRNA**. The amino acid is said to be "activated" for subsequent transfer to a growing polypeptide chain because the aminoacyl-tRNA is a "high-energy" molecule. A specific aminoacyl-tRNA molecule is identified by naming both the tRNA and the attached amino acid;

for example, aminoacylated $tRNA^{Ala}$ is called alanyl-$tRNA^{Ala}$. The various enzymes that catalyze the aminoacylation reaction are called aminoacyl-tRNA synthetases (e.g., alanyl-tRNA synthetase).

Most species have at least 20 different aminoacyl-tRNA synthetases in each cell since there are 20 different amino acids. A few species have two different aminoacyl-tRNA synthetases for the same amino acid. Some bacteria don't have glutaminyl- or asparaginyl-tRNA synthetases. In these species, the glutaminyl- and asparaginyl-tRNAs are synthesized by modifying glutamate and aspartate residues after they have been covalently attached to $tRNA^{Gln}$ and $tRNA^{Asn}$ by glutamyl- and aspartyl-tRNA synthetases (Glutamate and aspartate residues that are bound to their proper tRNAs are not modified.)

Although each synthetase is specific for a particular amino acid, it can recognize many isoacceptor tRNA molecules. For example, there are six codons for serine and several different isoacceptor $tRNA^{Ser}$ molecules. All these different $tRNA^{Ser}$ molecules are recognized by the organism's single seryl-tRNA synthetase enzyme. The accuracy of protein synthesis depends on the ability of aminoacyl-tRNA synthetases to catalyze attachment of the correct amino acid to its corresponding tRNA.

A. The Aminoacyl-tRNA Synthetase Reaction

The activation of an amino acid by its specific aminoacyl-tRNA synthetase requires ATP. The overall reaction is:

$$\text{Amino Acid} + \text{tRNA} + \text{ATP} \longrightarrow \text{Aminoacyl-tRNA} + \text{AMP} + PP_i \qquad \textbf{(22.1)}$$

The amino acid is covalently attached to the tRNA molecule by the formation of an ester linkage between the carboxylate group of the amino acid and a hydroxyl group of the ribose at the 3′ end of the tRNA molecule. Since all tRNAs end in —CCA, the attachment site is always an adenylate residue.

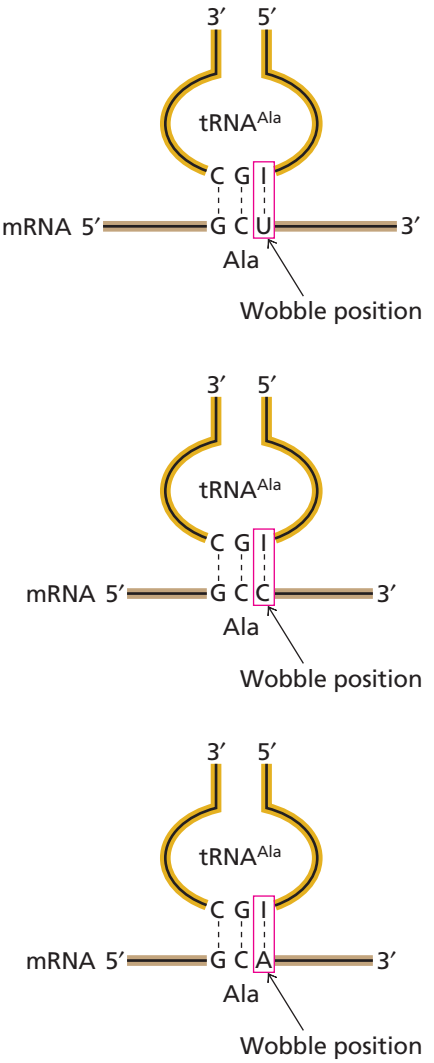
Aminoacylation proceeds in two discrete steps (Figure 22.9). In the first step, the amino acid is activated by formation of a reactive aminoacyl-adenylate intermediate. The intermediate remains tightly but noncovalently bound to the aminoacyl-tRNA synthetase. Rapid hydrolysis of the liberated pyrophosphate strongly favors the forward reaction. The second step of aminoacyl-tRNA formation is aminoacyl-group transfer from the aminoacyl-adenylate intermediate to tRNA. The amino acid is attached to either the 2′- or the 3′-hydroxyl group of the terminal adenylate residue of tRNA, depending on the specific aminoacyl-tRNA synthetase catalyzing the reaction. If the amino acid is initially attached to the 2′-hydroxyl group, it is shifted to the 3′-hydroxyl group in an additional step. The amino acid must be attached to the 3′ position to function as a protein synthesis substrate.

Formation of the aminoacyl-tRNA is favored under cellular conditions and the intracellular concentration of free tRNA is very low. The Gibbs free energy of hydrolysis of an aminoacyl-tRNA is approximately equivalent to that of a phosphoanhydride bond in ATP. The energy stored in the aminoacyl-tRNA is ultimately used in the formation of a peptide bond during protein synthesis. Note that the two ATP equivalents consumed during each aminoacylation reaction contribute to the energetic cost of protein synthesis.

B. Specificity of Aminoacyl-tRNA Synthetases

Attaching a specific amino acid to its corresponding tRNA is a crucial step in translating a genetic message. If there are errors at this step, the wrong amino acid could be incorporated into a protein.

Each aminoacyl-tRNA synthetase binds ATP and selects the proper amino acid based on its charge, size, and hydrophobicity. This initial selection eliminates most of the other amino acids. For example, tyrosyl-tRNA synthetase almost always binds tyrosine but rarely phenylalanine or any other amino acid. The synthetase then selectively binds a specific tRNA molecule. The proper tRNA is distinguished by features unique to its structure. In particular, the part of the acceptor stem that lies on the inner surface of



▲ **Figure 22.8**
Base pairing at the wobble position. The $tRNA^{Ala}$ molecule with the anticodon IGC can bind to any one of three codons specifying alanine (GCU, GCC, or GCA) because I can pair with U, C, or A. Note that the RNA strand containing the codon and the strand containing the anticodon are antiparallel. The wobble position is boxed in each example.

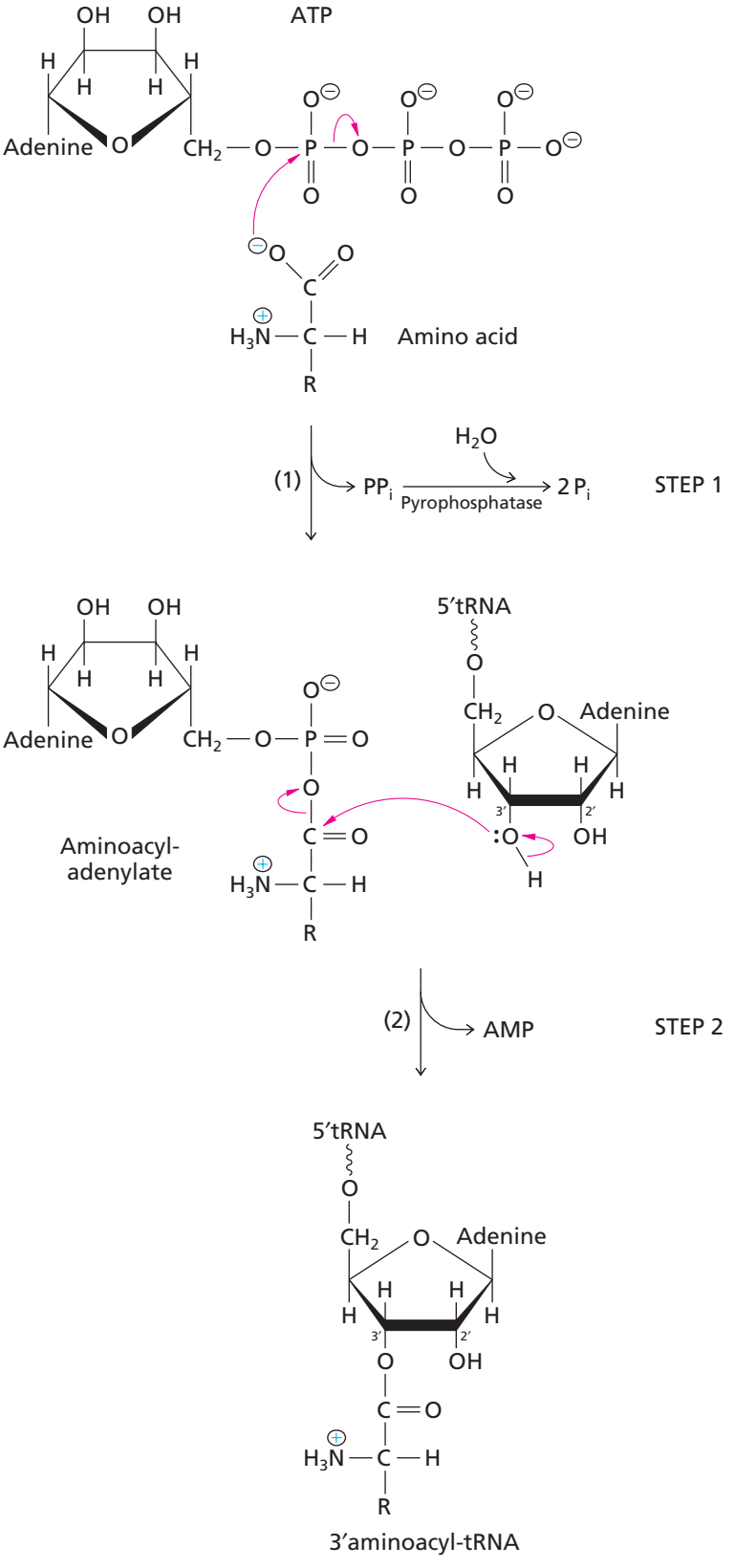


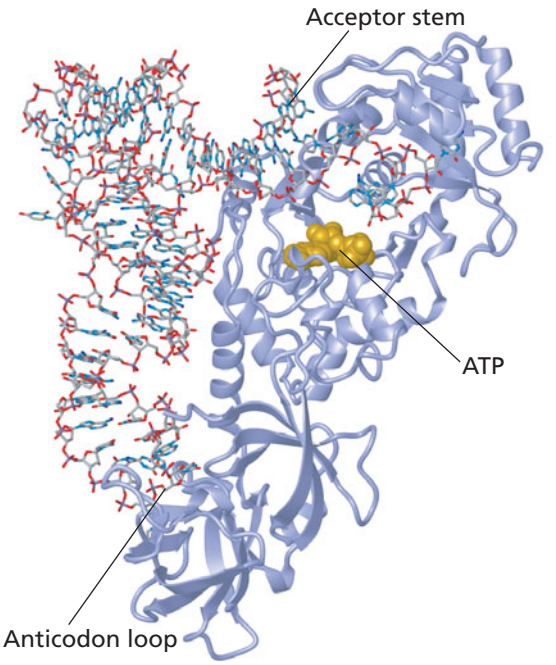
Figure 22.9 ►
Synthesis of an aminoacyl-tRNA molecule catalyzed by its specific aminoacyl-tRNA synthetase. In the first step, the nucleophilic carboxylate group of the amino acid attacks the α-phosphorus atom of ATP, displacing pyrophosphate and producing an aminoacyl-adenylate intermediate. In the second step, nucleophilic attack by the 3′-hydroxyl group of the terminal residue of tRNA leads to displacement of AMP and formation of an aminoacyl-tRNA molecule.

the L-shaped tRNA molecule is implicated in the binding of tRNA to the aminoacyl-tRNA synthetase (Figure 22.10).

In some cases, the synthetase enzyme recognizes not only the the acceptor stem of the tRNA but also the anticodon. For example, the glutaminyl-tRNA synthetase's ability to recognize Gln-tRNAs and to discriminate against the other 19 types of tRNAs ensures

that glutamine is specifically attached to the correct tRNA (shown in Figure 22.10). Note that glutaminyl-tRNA synthetase contacts both the acceptor stem and the anticodon region of $tRNA^{Gln}$. The crystal structure also shows a molecule of ATP bound in the active site near the 3′ end of the tRNA.

Half of the 20 different aminoacyl-tRNA synthetases resemble glutaminyl-tRNA synthetase. These enzymes bind the anticodon and aminoacylate tRNA at the 2′-hydroxyl group. A subsequent chemical rearrangement shifts the aminoacyl group to the 3′-hydroxyl group. Such enzymes are known as class I synthetases. Class II aminoacyl-tRNA synthetases are often more complex, multisubunit enzymes and they aminoacylate tRNA at the 3′-hydroxyl group. In all cases, the net effect of the interaction between tRNA and synthetase is to position the 3′ end of the tRNA molecule in the active site of the enzyme.



▲ **Figure 22.10**
Structure of *E. coli* $tRNA^{Gln}$ bound to glutaminyl-tRNA synthetase. The 3′ end of the tRNA is buried in a pocket on the surface of the enzyme. A molecule of ATP is also bound at this site. The enzyme interacts with both the tRNA acceptor stem and anticodon. [PDB 1QRS].

C. Proofreading Activity of Aminoacyl-tRNA Synthetases

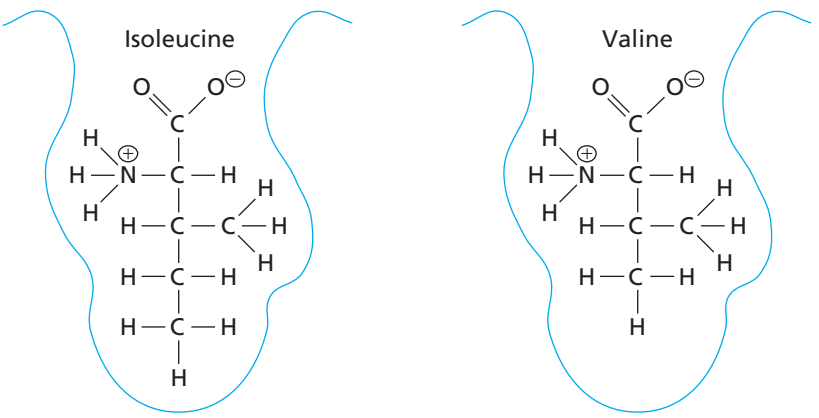
The error rate for most aminoacyl-tRNA synthetases is low because they make multiple contacts with a specific tRNA and a specific amino acid. However, isoleucine and valine are chemically similar amino acids, and both can be accommodated in the active site of isoleucyl-tRNA synthetase (Figure 22.11). Isoleucyl-tRNA synthetase mistakenly catalyzes the formation of the valyl-adenylate intermediate about 1% of the time. On the basis of this observation, we might expect valine to be attached to isoleucyl-tRNA and incorporated into protein in place of isoleucine about 1 time in 100 but the observed substitution of valine for isoleucine in polypeptide chains is only about 1 time in 10,000. This lower level of valine incorporation suggests that isoleucyl-tRNA synthetase also discriminates between the two amino acids after aminoacyl-adenylate formation. In fact, isoleucyl-tRNA synthetase carries out proofreading in the next step of the reaction. Although isoleucyl-tRNA synthetase may mistakenly catalyze the formation of valyl-adenylate, it usually catalyzes hydrolysis of the incorrect valyl-adenylate to valine and AMP or the hydrolysis of valyl-$tRNA^{Ile}$. The overall error rate of the reaction is 10^{-5} for most amino acyl-tRNA synthetases.

KEY CONCEPT

The accuracy of information flow from nucleic acids to protein depends, in part, on the accuracy of the amino acyl-tRNA synthetase reaction.

22.4 Ribosomes

Protein synthesis requires assembling four components that form an elaborate translation complex: the ribosome, which catalyzes peptide bond formation; its accessory protein factors, which help the ribosome in each step of the process; the mRNA, which carries the information specifying the protein's sequence; and the aminoacyl-tRNAs that carry the activated amino acids. Initiation involves assembly of the translation complex at the first codon in the mRNA. During polypeptide chain elongation the ribosome and associated components move, or translocate, along the template mRNA in the 5′ ⟶ 3′ direction.



◀ **Figure 22.11**
Model of the substrate-binding site in isoleucyl-tRNA synthetase. Despite the similar size and charge of isoleucine and valine, isoleucyl-tRNA synthetase binds to isoleucine about 100 times more readily than it binds to valine. A subsequent proofreading step also helps prevent the formation of valyl-$tRNA^{Ile}$.

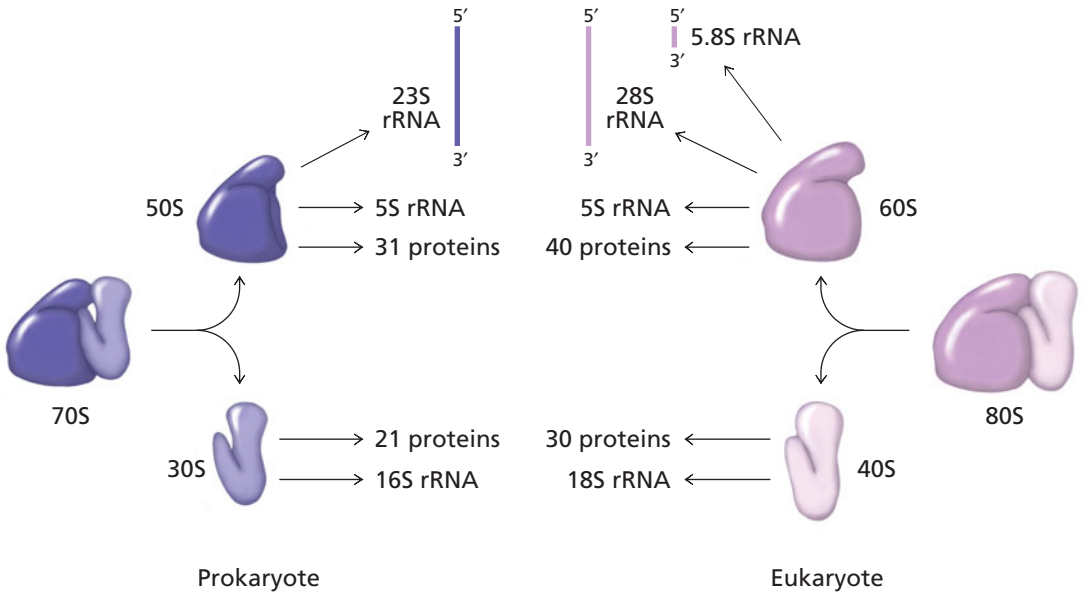
The polypeptide is synthesized from the N-terminus to its C-terminus. Finally, when synthesis of the protein is complete, the translation complex disassembles in a separate termination step. An important function of disassembly is to release the two ribosomal subunits from the mRNA so that they can participate in further rounds of translation.

A. Ribosomes Are Composed of Both Ribosomal RNA and Protein

All ribosomes contain two subunits of unequal size. In *E. coli,* the small subunit is called the 30S subunit and the large subunit is called the 50S subunit. (The terms 30S and 50S originally referred to the sedimentation rate of these subunits.) The 30S subunit is elongated and asymmetric, with overall dimensions of 5.5 × 22 × 22.5 nm. A narrow neck separates the head from the base and a protrusion extends from the base forming a cleft where the mRNA molecule appears to rest. The 50S ribosomal subunit is wider than the 30S subunit and has several protrusions; its dimensions are about 15 × 20 × 20 nm. The 50S subunit also contains a tunnel about 10 nm long and 2.5 nm in diameter. This tunnel extends from the site of peptide bond formation and accommodates the growing polypeptide chain during protein synthesis. The 30S and 50S subunits combine to form an active 70S ribosome.

In *E. coli,* the RNA component of the 30S subunit is a 16S rRNA of 1542 nucleotides. Although its exact length varies among species, the 16S rRNA contains extensive regions of secondary structure that are highly conserved in the ribosomes of all living organisms. There are 21 ribosomal proteins in the 30S subunit. The 50S subunit of the *E. coli* ribosome contains two molecules of ribosomal RNA: one 5S rRNA of 120 nucleotides and one 23S rRNA of 2904 nucleotides. There are 31 different proteins associated with the 5S and 23S rRNA molecules in the 50S subunit (Figure 22.12).

Eukaryotic ribosomes are similar in shape to bacterial ribosomes but they tend to be somewhat larger and more complex. Intact vertebrate ribosomes are designated 80S and are made up of 40S and 60S subunits (Figure 22.12). The small 40S subunit is analogous to the 30S subunit of the prokaryotic ribosome; it contains about 30 proteins and a single molecule of 18S rRNA. The large 60S subunit contains about 40 proteins and three ribosomal RNA molecules: 5S rRNA, 28S rRNA, and 5.8S rRNA. The 5.8S rRNA is about 160 nucleotides long and its sequence is similar to that of the 5′ end of prokaryotic 23S rRNA. This similarity implies that the 5.8S rRNA and the 5′ end of prokaryotic 23S



▲ **Figure 22.12**
Comparison of prokaryotic and eukaryotic ribosomes. Both types of ribosomes consist of two subunits, each of which contains ribosomal RNA and proteins. The large subunit of the prokaryotic ribosome contains two molecules of rRNA: 5S and 23S. The large subunit of almost all eukaryotic ribosomes contains three molecules of rRNA: 5S, 5.8S, and 28S. The sequence of the eukaryotic 5.8S rRNA is similar to the sequence of the 5′ end of the prokaryotic 23S rRNA.

rRNA are derived from a common ancestor and that there has been a fusion or splitting of rRNA genes during their evolution.

Both prokaryotic and eukaryotic genomes contain multiple copies of ribosomal RNA genes. The combination of a large number of copies and strong promoters for these genes allows cells to maintain a high level of ribosome synthesis. Eukaryotic ribosomal RNA genes, which are transcribed by RNA polymerase I (Section 21.5A), occur as tandem arrays of hundreds of copies. In most eukaryotes, these genes are clustered in the nucleolus, where processing of ribosomal RNA precursors and ribosome assembly occur (Section 21.8B). This processing is coupled to ribosome assembly, as shown in Figure 22.13 for the *E. coli* 30S subunit. Many of the ribosomal proteins contact RNA and bind specifically to regions of secondary structure in 16S rRNA. Others form protein–protein contacts and assemble into the complex only when other ribosomal proteins are present.

The structure of the 30S ribosomal subunit from the bacterium *Thermus thermophilus* is shown in Figure 22.14 on page 676. Note that most of the mass of the 30S subunit is due to the 16S ribosomal RNA, which forms a compact structure made up of multiple regions of double-stranded RNA. The ribosomal proteins bind to the surface of the RNA or to grooves and crevices between regions of RNA secondary structure.

Similarly, the assembly of the bacterial 50S subunit and of the 40S and 60S eukaryotic subunits are also coupled to the processing of their ribosomal RNA precursors. The structure of the 50S subunit from the archeon *Haloarcula marismortui* is also shown in Figure 22.14.

B. Ribosomes Contain Two Aminoacyl-tRNA Binding Sites

As discussed in Section 22.3, the substrates for peptide bond formation are not free amino acids but relatively large aminoacyl-tRNA molecules. A ribosome must align two adjacent aminoacyl-tRNA molecules so that their anticodons interact with the correct mRNA codons. The aminoacylated ends of these two tRNAs are positioned at the site of peptide bond formation. The ribosome must also hold the mRNA and the growing polypeptide chain, and it must accommodate the binding of several protein factors during protein synthesis. The ability to accomplish these tasks simultaneously explains, in part, why the ribosome is so large and complex.

The orientation of the two tRNA molecules during protein synthesis is shown in Figure 22.15 on page 677. The growing polypeptide chain is covalently attached to the tRNA positioned at the peptidyl site (P site), forming peptidyl-tRNA. The second aminoacyl-tRNA is bound at the aminoacyl site (A site). As the polypeptide chain is synthesized, it passes through the tunnel of the large ribosomal subunit and emerges on the outer surface of the ribosome.

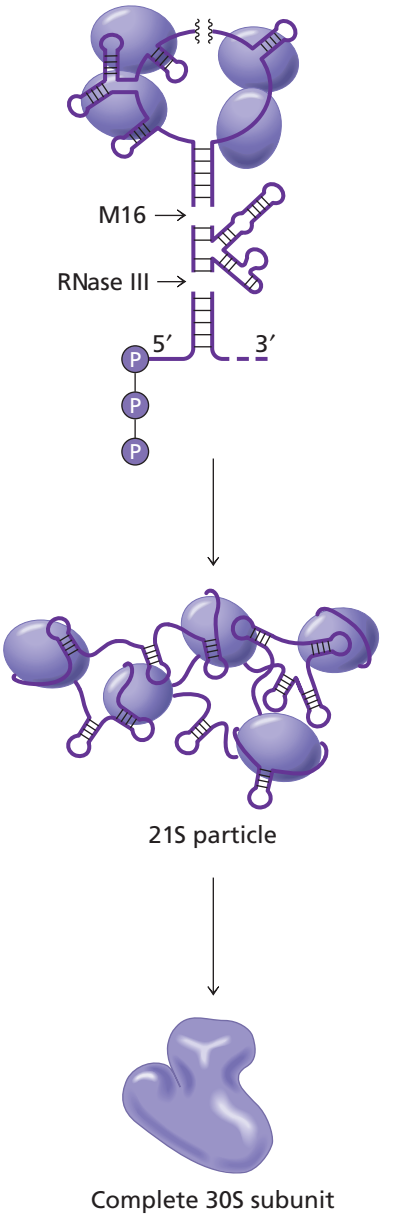
22.5 Initiation of Translation

The initiation of protein synthesis involves assembling a translation complex at the beginning of an mRNA's coding sequence. This complex consists of the two ribosomal subunits, an mRNA template to be translated, an initiator tRNA molecule, and several accessory proteins called *initiation factors*. This crucial initiation step ensures that the proper initiation codon (and therefore the correct reading frame) is selected before translation begins.

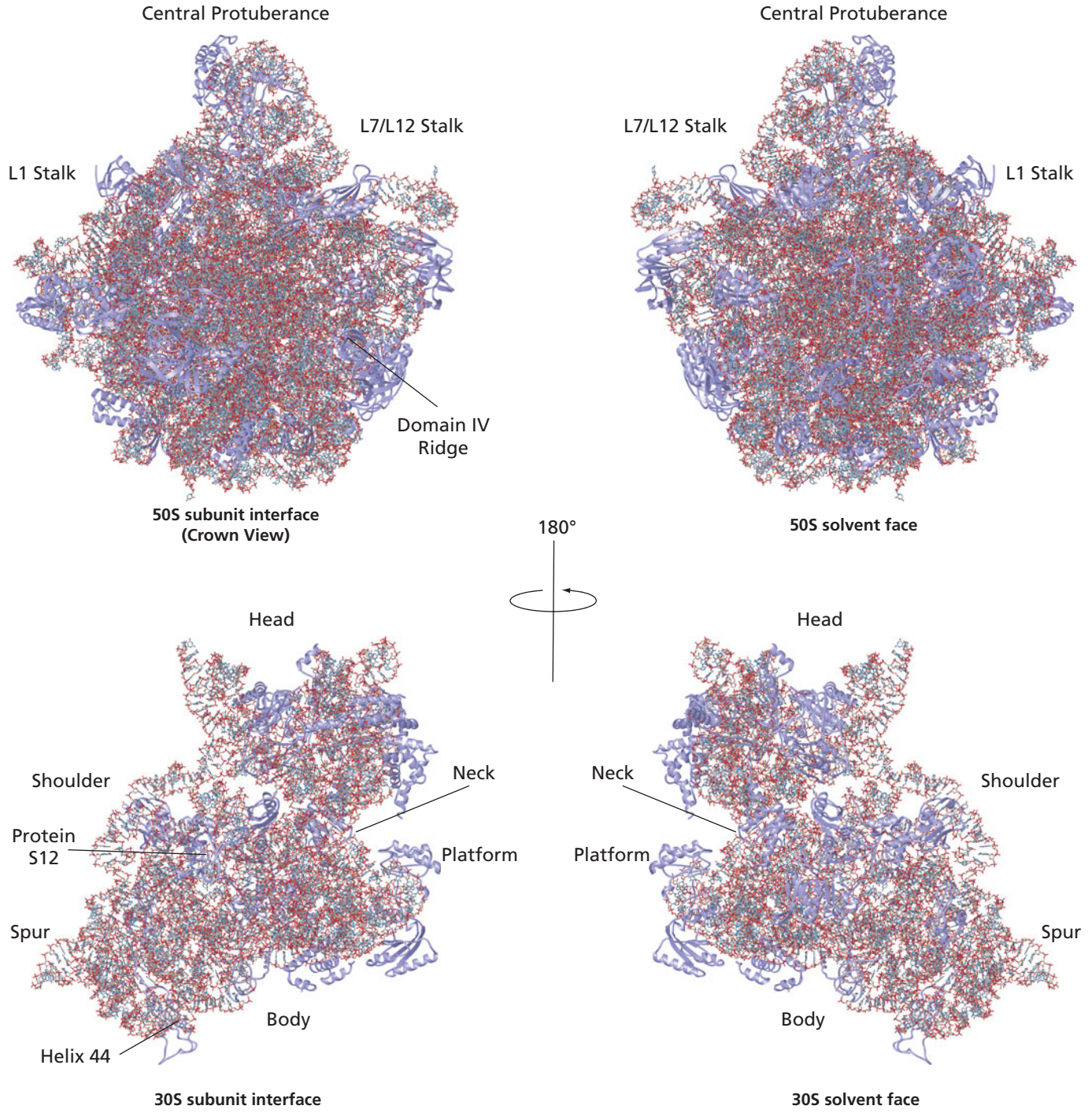
A. Initiator tRNA

As mentioned in Section 22.1, the first codon translated is usually AUG. Every cell contains at least two types of methionyl-tRNA^{Met} molecules that can recognize AUG codons. One type is used exclusively at initiation codons and is called the initiator tRNA. The other type only recognizes internal methionine codons. Although these two tRNA^{Met} molecules have different primary sequences, and distinct functions, both of them are aminoacylated by the same methionyl-tRNA synthetase.

In bacteria, the initiator tRNA is called tRNA$_f^{Met}$. The charged initiator tRNA (methionyl-tRNA$_f^{Met}$) is the substrate for a formyltransferase that catalyzes addition of a formyl group from 10-formyltetrahydrofolate to the methionine residue producing



▲ **Figure 22.13**
Assembly of the 30S ribosomal subunit and maturation of 16S rRNA in *E. coli*. Assembly of the 30S ribosomal subunit begins when six or seven ribosomal proteins bind to the 16S rRNA precursor as it is being transcribed, thereby forming a 21S particle. The 21S particle undergoes a conformational change, and the 16S rRNA molecule is processed to its final length. During this processing, the remaining ribosomal proteins of the 30S subunit bind (recall that M16 is a site-specific endonuclease involved in RNA processing that we discussed in Chapter 21).



▲ **Figure 22.14**
Three-dimensional structures of the *H. marismortui* 50S subunit (top) and the *T. thermophilus* 30S subunit (bottom).

N-formylmethionyl-$tRNA_f^{Met}$ ($fMet\text{-}tRNA_f^{Met}$) as shown in Figure 22.16 on page 681. In eukaryotes and archaebacteria, the initiator tRNA is called $tRNA_i^{Met}$. The methionine that begins protein synthesis in eukaryotes is not formylated.

N-Formylmethionine in bacteria—or methionine in other organisms—is the first amino acid incorporated into proteins. After protein synthesis is under way, the N-terminal methionine can be either deformylated or removed from the polypeptide chain altogether.

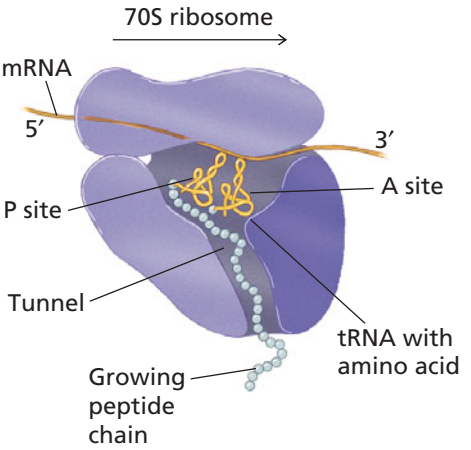
B. Initiation Complexes Assemble Only at Initiation Codons

There are three possible reading frames in an mRNA molecule but only one of them is correct. Establishing the correct reading frame during the initiation of translation is

critical for the accurate decoding of information from mRNA into protein. Shifting the reading frame by even a single nucleotide would alter the sequence of the entire polypeptide and result in a nonfunctional protein. The translation machinery must therefore accurately locate the initiation codon that serves as the start site for protein synthesis.

The ribosome needs to distinguish between the single correct initiation codon and all the other incorrect AUGs. These other AUGs specify either internal methionine residues in the correct reading frame or irrelevant methionine codons in the two other incorrect reading frames. It is important to appreciate that the initiation codon is not simply the first three nucleotides of the mRNA. Initiation codons can be located many nucleotides downstream of the 5′-end of the mRNA molecule.

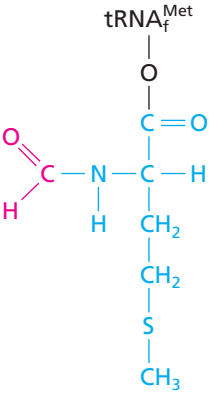
In prokaryotes, the selection of an initiation site depends on an interaction between the small subunit of the ribosome and the mRNA template. The 30S subunit binds to the mRNA template at a purine-rich region just upstream of the correct initiation codon. This region, called the Shine-Dalgarno sequence, is complementary to a pyrimidine-rich stretch at the 3′ end of the 16S rRNA molecule. During formation of the initiation complex, these complementary nucleotides pair to form a double-stranded RNA structure that binds the mRNA to the ribosome. The result of this interaction is to position the initiation codon at the P site on the ribosome (Figure 22.17). The initiation complex assembles exclusively at initiation codons because Shine-Dalgarno sequences are not found immediately upstream of internal methionine codons.



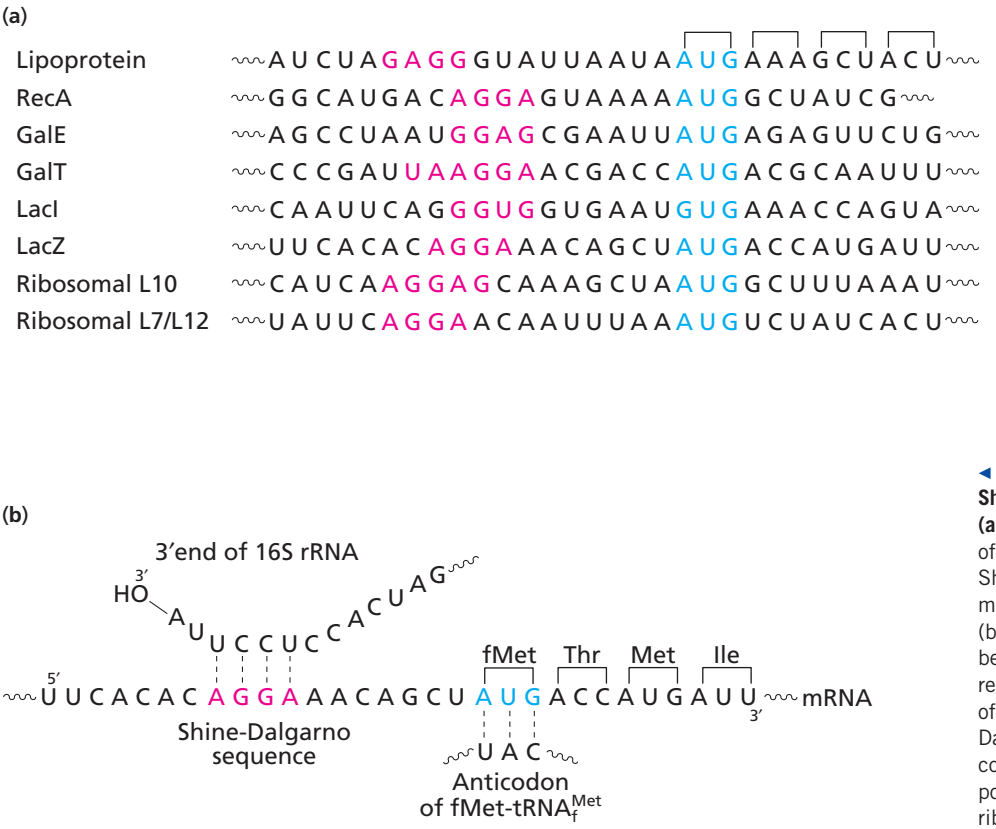
▲ **Figure 22.15**
Sites for tRNA binding in prokaryotic ribosomes. During protein synthesis, the P site is occupied by the tRNA molecule attached to the growing polypeptide chain, and the A site holds an aminoacyl-tRNA. The growing polypeptide chain passes through the tunnel of the large subunit.

C. Initiation Factors Help Form the Initiation Complex

Formation of the initiation complex requires several **initiation factors** in addition to ribosomes, initiator tRNA, and mRNA. Prokaryotes contain three initiation factors, designated IF-1, IF-2, and IF-3. There are at least eight eukaryotic initiation factors (eIF's). In both prokaryotes and eukaryotes, the initiation factors catalyze assembly of the protein synthesis complex at the initiation codon.



▲ **Figure 22.16**
Chemical structure of fMet-tRNA$_f^{Met}$. A formyl group (red) is added to the methionyl moiety (blue) of methionyl-tRNA$_f^{Met}$ in a reaction catalyzed by a formyltransferase.

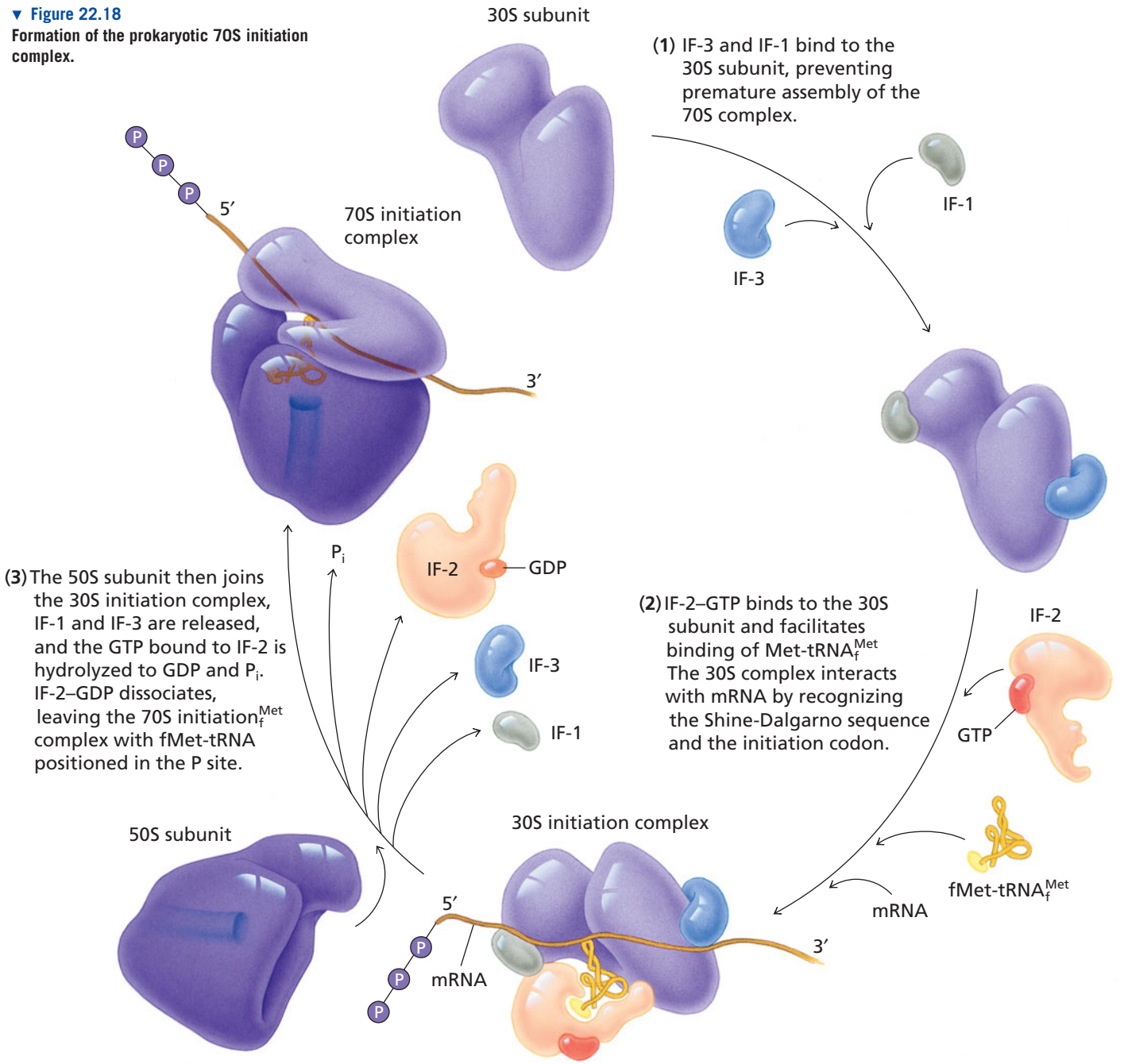


◀ **Figure 22.17**
Shine-Dalgarno sequences in *E. coli* mRNA. **(a)** Ribosome-binding sites at the 5′ end of mRNA for several *E. coli* proteins. The Shine-Dalgarno sequences (red) occur immediately upstream of initiation codons (blue). **(b)** Complementary base pairing between the 3′ end of 16S rRNA and the region near the 5′ end of an mRNA. Binding of the 3′ end of the 16S rRNA to the Shine-Dalgarno sequence helps establish the correct reading frame for translation by positioning the initiation codon at the ribosome's P site.

One of the roles of IF-3 is to maintain the ribosomal subunits in their dissociated state by binding to the small subunit. The ribosomal subunits bind separately to the initiation complex and the association of IF-3 with the 30S subunit prevents the 30S and 50S subunits from forming the 70S complex prematurely. IF-3 also helps position fMet-tRNA$_f^{Met}$ and the initiation codon at the P site of the ribosome. IF-2 selects the initiator tRNA from the pool of aminoacylated tRNA molecules in the cell. It binds GTP forming an IF-2–GTP complex that specifically recognizes the initiator tRNA and rejects all other aminoacyl-tRNA molecules. The third initiation factor, IF-1, binds to the 30S subunit and facilitates the actions of IF-2 and IF-3.

Once the 30S complex has been formed at the initiation codon, the 50S ribosomal subunit binds to the 30S subunit. Next, the GTP bound to IF-2 is hydrolyzed and P_i is released. The initiation factors dissociate from the complex when GTP is hydrolyzed. IF-2–GTP is regenerated when the bound GDP is exchanged for GTP. The steps in the formation of the 70S initiation complex are summarized in Figure 22.18.

▼ **Figure 22.18**
Formation of the prokaryotic 70S initiation complex.



The role of the prokaryotic initiation factors is to ensure that the aminoacylated initiator tRNA (fMet-$tRNA_i^{Met}$) is correctly positioned at the initiation codon. The initiation factors also mediate the formation of a complete initiation complex by reconstituting a 70S ribosome such that the initiation codon is positioned in the P site.

D. Translation Initiation in Eukaryotes

Eukaryotic mRNAs do not have distinct Shine-Dalgarno sequences that serve as ribosome binding sites. Instead, the first AUG codon in the message usually serves as the initiation codon. eIF-4 (eukaryotic initiation factor 4), also known as cap binding protein (CBP), binds specifically to the 7-methylguanylate cap (Figure 21.26) at the 5′ end of eukaryotic mRNA. Binding of eIF-4 to the cap structure leads to the formation of a preinitiation complex consisting of the 40S ribosomal subunit, an aminoacylated initiator tRNA, and several other initiation factors. The preinitiation complex then scans along the mRNA in the 5′ → 3′ direction until it encounters an initiation codon. When the search is successful, the small ribosomal subunit is positioned so that Met-$tRNA_i^{Met}$ interacts with the initiation codon in the P site. In the final step, the 60S ribosomal subunit binds to complete the 80S initiation complex and all the initiation factors dissociate. The dissociation of eIF-2—the eukaryotic counterpart of bacterial IF-2—is accompanied by GTP hydrolysis.

Most eukaryotic mRNA molecules encode only a single polypeptide since the normal mechanism of selecting the initiation codon by scanning along the mRNA from the 5′ end permits only one initiation codon per mRNA. In contrast, prokaryotic mRNAs often contain several coding regions. Each coding region begins with an initiation codon that is associated with its own upstream Shine-Dalgarno sequence. mRNA molecules that encode several polypeptides are said to be **polycistronic**.

22.6 Chain Elongation During Protein Synthesis Is a Three-Step Microcycle

At the end of the initiation step, the mRNA is positioned so that the next codon can be translated during the elongation stage of protein synthesis. The initiator tRNA occupies the P site in the ribosome and the A site is ready to receive an incoming aminoacyl-tRNA. During chain elongation each additional amino acid is added to the nascent polypeptide chain in a three-step microcycle. The steps in this microcycle are (1) positioning the correct aminoacyl-tRNA in the A site of the ribosome, (2) forming the peptide bond, and (3) shifting, or translocating, the mRNA by one codon relative to the ribosome (the two tRNAs in the ribosome's P and A sites also translocate).

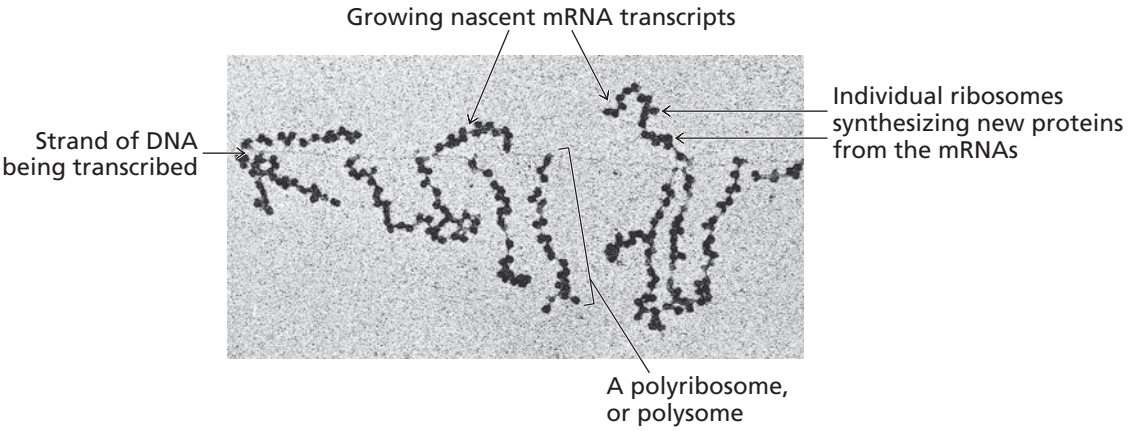
The translation machinery works relatively slowly compared to the enzyme systems that catalyze DNA replication. Proteins are synthesized at a rate of only 18 amino acid residues per second, whereas bacterial replisomes synthesize DNA at a rate of 1000 nucleotides per second. This difference in rates reflects, in part, the difference between polymerizing four types of nucleotides to make nucleic acids and polymerizing 20 types of amino acids to make proteins. Testing and rejecting all of the incorrect aminoacyl-tRNA molecules also takes time and slows protein synthesis.

The rate of transcription in prokaryotes is approximately 55 nucleotides per second. This corresponds to about 18 codons per second or the same rate at which the mRNA is translated. In bacteria, translation initiation occurs as soon as the 5′ end of an mRNA is synthesized and translation and transcription are coupled (Figure 22.19 on page 680). This tight coupling is not possible in eukaryotes because transcription and translation are carried out in separate compartments of the cell (the nucleus and the cytoplasm, respectively). Eukaryotic mRNA precursors must be processed in the nucleus (e.g., capped, polyadenylated, spliced) before they are exported to the cytoplasm for translation.

An *E. coli* cell contains about 20,000 ribosomes. Many large eukaryotic cells have several hundred thousand ribosomes. Large mRNA molecules can be translated simultaneously by many protein synthesis complexes forming a polyribosome or **polysome**, as

KEY CONCEPT

The A site of an actively translating ribosome spends the vast majority of its time bound to one of the 19 types of incorrect aminoacyl-tRNAs as it randomly samples the pool of charged tRNAs, seeking the correct tRNA.



▲ **Figure 22.19**
Coupled transcription and translation of an *E. coli* gene. The gene is being transcribed from left to right. Ribosomes bind to the 5′ end of the mRNA molecules as soon as they are synthesized. The large polysomes on the right are released from the gene when transcription terminates.

seen in Figure 22.19. The number of ribosomes bound to an mRNA molecule depends on the length of the mRNA and the efficiency of initiation of protein synthesis. At maximal efficiency the spacing between each translation complex in the polysome is about 100 nucleotides. On average, each mRNA molecule in an *E. coli* cell is translated 30 times, effectively amplifying the information it encodes by 30-fold.

A. Elongation Factors Dock an Aminoacyl-tRNA in the A Site

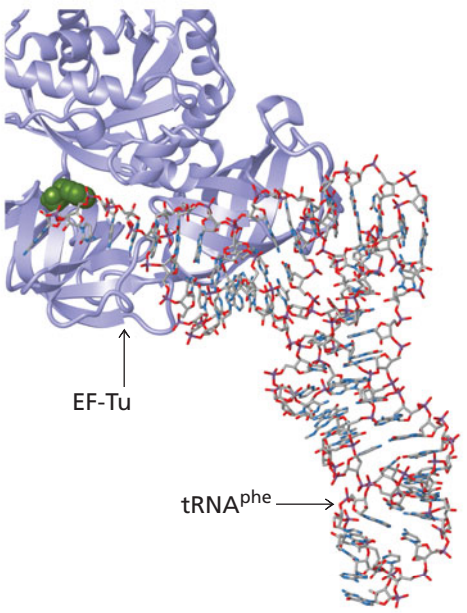
At the start of the first chain elongation microcycle, the A site is empty and the **P site** is occupied by the aminoacylated initiator tRNA. The first step in chain elongation is insertion of the correct aminoacyl-tRNA into the A site of the ribosome. In bacteria, this step is catalyzed by an elongation factor called EF-Tu. EF-Tu is a monomeric protein that contains a binding site for GTP. Each *E. coli* cell has about 135,000 molecules of EF-Tu, making it one of the most abundant proteins in the cell (emphasizing the importance of protein synthesis to a cell).

EF-Tu–GTP associates with an aminoacyl-tRNA molecule to form a ternary complex that fits into the A site of a ribosome. Almost all aminoacyl-tRNA molecules *in vivo* are found in such ternary complexes (Figure 22.20). The structure of EF-Tu is similar to that of IF-2 (which also binds GTP) and other G proteins (Section 9.12A), suggesting that they all evolved from a common ancestral protein.

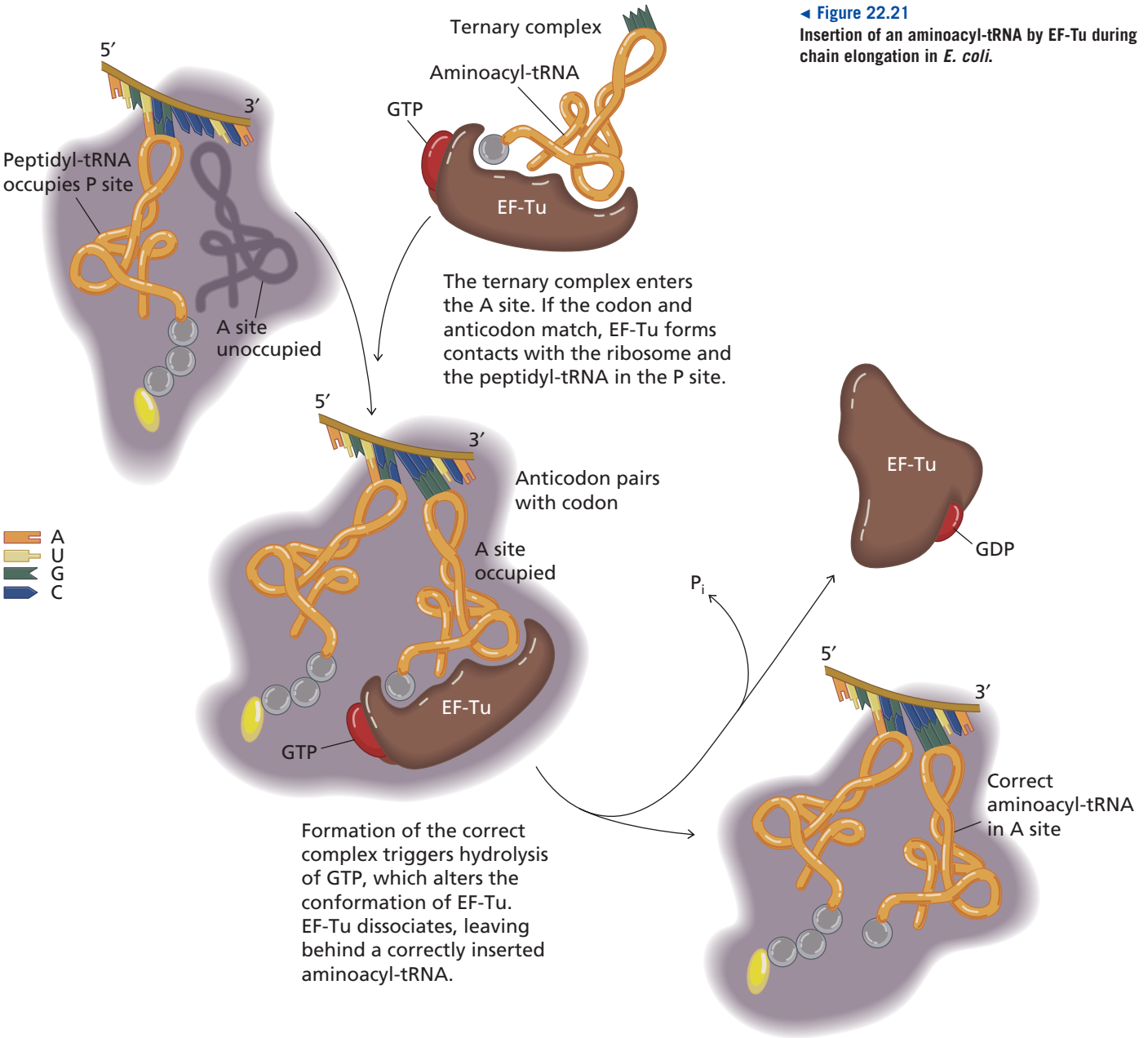
The EF-Tu–GTP complex recognizes common features of the tertiary structure of tRNA molecules and binds tightly to all aminoacyl-tRNA molecules except fMet-tRNA$_f^{Met}$. The fMet-tRNA$_f^{Met}$ molecule is distinguished from all other aminoacyl-tRNA molecules by the distinctive secondary structure of its acceptor stem.

A ternary complex of EF-Tu–GTP–aminoacyl-tRNA can diffuse freely into the A site in the ribosome. When correct base pairs form between the anticodon of the aminoacyl-tRNA and the mRNA codon in the A site, the complex is stabilized. EF-Tu–GTP can then contact sites in the ribosome as well as the tRNA in the P site (Figure 22.21, on page 681). These contacts trigger hydrolysis of GTP to GDP and P_i causing a conformational change in EF-Tu–GDP that releases the bound aminoacyl-tRNA. EF-Tu–GDP then dissociates from the chain elongation complex. The aminoacyl-tRNA remains in the A site where it is positioned for peptide bond formation.

EF-Tu–GDP cannot bind another aminoacyl-tRNA molecule until GDP dissociates. An additional elongation factor called EF-Ts catalyzes the exchange of bound GDP for GTP (Figure 22.22, on page 682). Note that one GTP molecule is hydrolyzed for every aminoacyl-tRNA that is successfully inserted into the A site.



▲ **Figure 22.20**
EF-Tu binds aminoacylated tRNAs. The EF-Tu–GTP complex binds to the acceptor end of aminoacylated tRNA (in this case phenylalanyl-tRNA^{Phe}). The phenylalanine residue is shown in green. This is how charged tRNAs commonly exist inside a cell.



◀ **Figure 22.21**
Insertion of an aminoacyl-tRNA by EF-Tu during chain elongation in *E. coli*.

B. Peptidyl Transferase Catalyzes Peptide Bond Formation

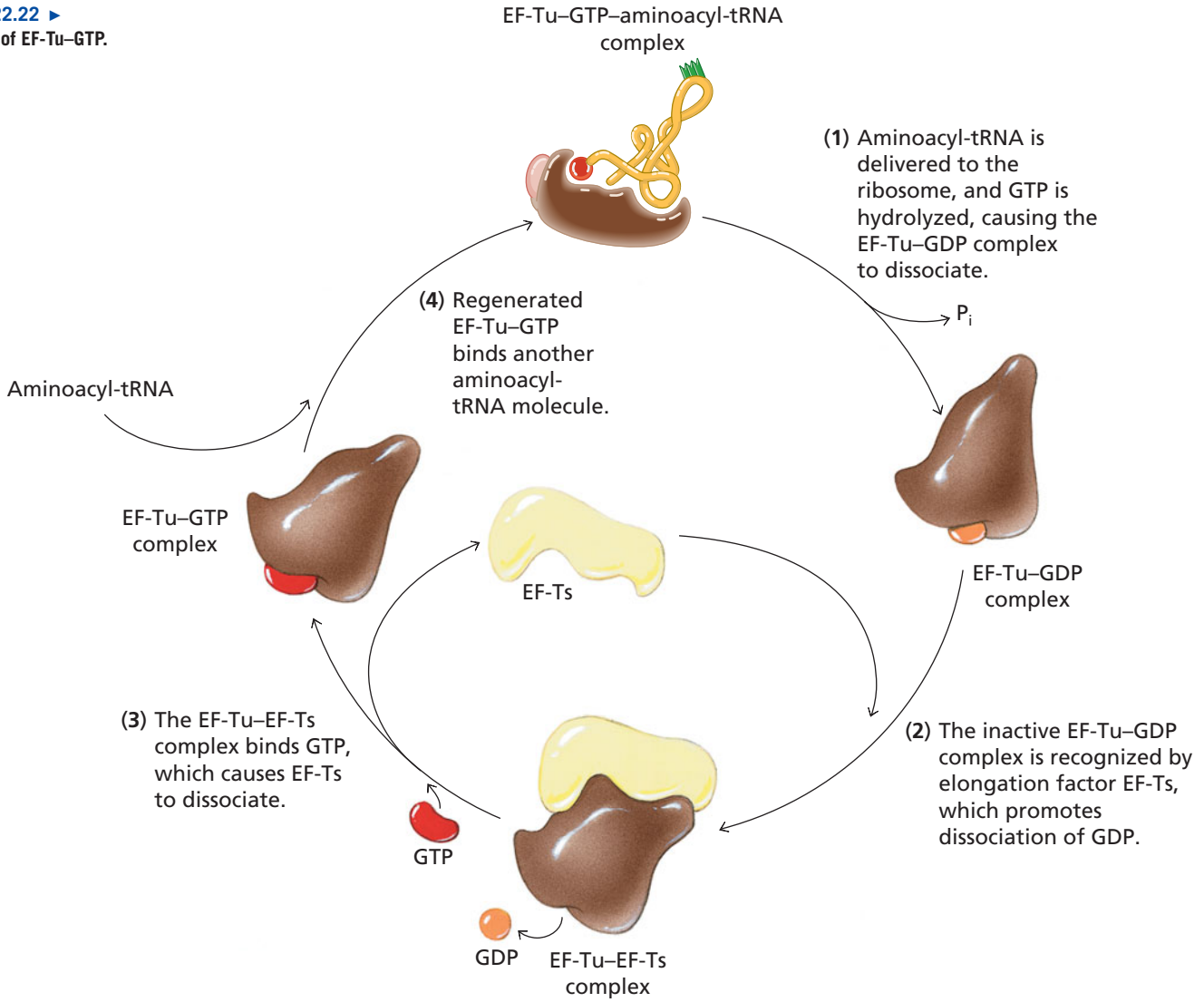
Binding of a correct aminoacyl-tRNA in the A site aligns the activated amino acid's α-amino group next to the ester bond's carbonyl on the peptidyl-tRNA in the neighboring P site. The nitrogen atom's lone pair of electrons execute a nucleophilic attack on the carbonyl carbon, resulting in the formation of a peptide bond via a displacement reaction. While it is straightforward to visualize how the ribosome's active site aligns these substrates, we do not understand precisely how the ribosome enhances the rate of this reaction. The peptide chain, now one amino acid longer, is transferred from the tRNA in the P site to the tRNA in the A site (Figure 22.23, on page 683). Formation of the peptide bond requires hydrolysis of the energy-rich peptidyl-tRNA linkage. Note that the growing polypeptide chain is covalently attached to the tRNA in the A site, forming a peptidyl-tRNA.

The enzymatic activity responsible for formation of the peptide bond is referred to as **peptidyl transferase**. This activity is contained within the large ribosomal subunit. Both the 23S rRNA molecule and the 50S ribosomal proteins contribute to the substrate binding sites, but the catalytic activity is localized to the RNA component. Thus, peptidyl transferase is yet another example of an RNA-catalyzed reaction.

KEY CONCEPT

Formation of the new peptide bond involves physically transferring the polypeptide attached to the P site tRNA onto the amino-terminus of the aminoacyl-tRNA bound in the ribosome's A site.

Figure 22.22 ►
Cycling of EF-Tu–GTP.

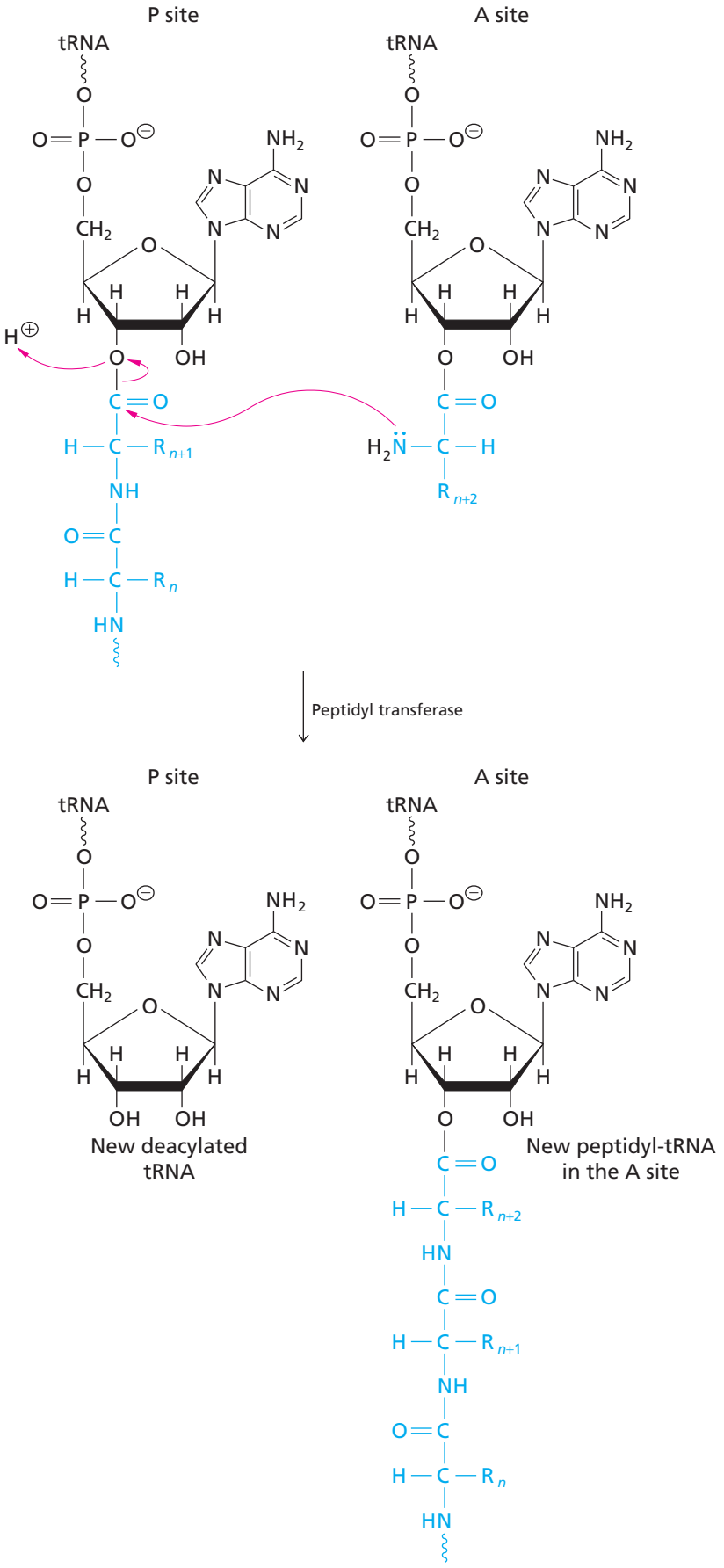


C. Translocation Moves the Ribosome by One Codon

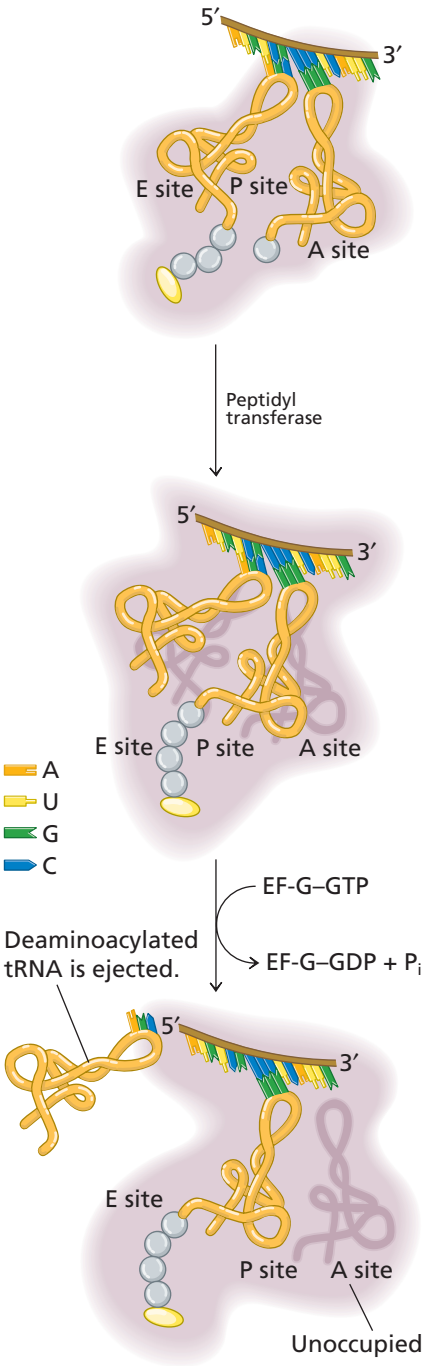
After the peptide bond has formed, the newly created peptidyl-tRNA is partially in the A site and partially in the P site (Figure 22.24, on page 684). The deaminoacylated tRNA has been displaced somewhat from the P site. It now occupies a position on the ribosome that is referred to as the exit site, or **E site**. Before the next codon can be translated, the deaminoacylated tRNA must be released and the peptidyl-tRNA must be completely transferred from the A site to the P site. At the same time, the mRNA must shift by one codon relative to the ribosome. This **translocation** is the third step in the chain elongation microcycle.

In prokaryotes, the translocation step requires a third elongation factor, EF-G. Like the other elongation factors, EF-G is an abundant protein; an *E. coli* cell contains approximately 20,000 molecules of EF-G, or roughly one for every ribosome. Like EF-Tu, EF-G has a binding site for GTP. Binding of EF-G–GTP to the ribosome completes the translocation of the peptidyl-tRNA from the A site to the P site and releases the deaminoacylated tRNA from the E site. EF-G itself is released from the ribosome only when its bound GTP is hydrolyzed to GDP and P_i is released. The dissociation of EF-G–GDP leaves the ribosome free to begin another microcycle of chain elongation.

The growing polypeptide chain extends from the peptidyl-tRNA in the P site through a tunnel in the 50S subunit, to exit on the exterior surface of the ribosome



◀ **Figure 22.23**
Formation of a peptide bond. The carbonyl carbon of the peptidyl-tRNA undergoes nucleophilic attack by the nitrogen atom of the amino group. This aminoacyl-group–transfer reaction results in growth of the peptide chain by one residue and transfer of the nascent peptide to the tRNA in the A site.



▲ **Figure 22.24**
Translocation during protein synthesis in prokaryotes.

top: Aminoacyl-tRNA is positioned in the A site.

middle: Following synthesis of the peptide bond, the newly formed peptidyl-tRNA is partly in the A site and partly in the P site.

bottom: Translocation shifts the peptidyl-tRNA completely into the P site, leaving the A site empty and ejecting the deaminoacylated tRNA from the E site.

(Figure 22.15). Each translocation step helps push the chain through the tunnel. The newly synthesized polypeptide doesn't begin to fold into its final shape until it emerges from the tunnel. This folding is assisted by chaperones, such as HSP70, that are associated with the translation machinery (Section 4.10D).

The elongation microcycle is repeated for each new codon in the mRNA being translated, resulting in the synthesis of a polypeptide chain that may be several hundred residues long. Eventually, the translation complex reaches the final codon at the end of the coding region, where translation is terminated.

The elongation reactions in eukaryotes are very similar to those in *E. coli*. Three accessory protein factors participate in chain elongation in eukaryotes: EF-1α, EF-1β, and EF-2. EF-1α docks the aminoacyl-tRNA in the A site; its activity thus parallels that of *E. coli* EF-Tu. EF-1β acts like bacterial EF-Ts, recycling EF-1α. EF-2 carries out translocation in eukaryotes. EF-Tu and EF-1α are highly conserved, homologous proteins, as are EF-G and EF-2. Eukaryotic and prokaryotic ribosomal RNAs are also very similar in sequence and in secondary structure. These similarities indicate that the common ancestor of prokaryotes and eukaryotes carried out protein synthesis in a manner similar to that seen in modern organisms. Thus, protein synthesis is one of the most ancient and fundamental biochemical reactions.

22.7 Termination of Translation

E. coli has three release factors (RF-1, RF-2, and RF-3) that participate in the termination of protein synthesis. After formation of the final peptide bond, the peptidyl-tRNA is translocated from the A site to the P site, as usual. The translocation positions one of the three termination codons (UGA, UAG, or UAA) in the A site. These termination codons are not recognized by any tRNA molecules so protein synthesis stalls at the termination codon. Eventually, one of the release factors diffuses into the A site. RF-1 recognizes UAA and UAG and RF-2 recognizes UAA and UGA. RF-3 binds GTP and enhances the effects of RF-1 and RF-2.

When the release factors recognize a termination codon, they cause hydrolysis of the peptidyl-tRNA. Release of the completed polypeptide is probably accompanied by GTP hydrolysis and dissociation of the release factors from the ribosome. At this point, the ribosomal subunits dissociate from the mRNA and initiation factors bind to the 30S subunit in preparation for the next round of protein synthesis.

22.8 Protein Synthesis Is Energetically Expensive

Protein synthesis is very expensive—it uses a large fraction of all ATP equivalents that are available in a cell. Where does all this energy go?

For each amino acid added to a polypeptide chain, four phosphoanhydride bonds are cleaved: ATP is hydrolyzed to AMP + 2 P_i during activation of the amino acid and two GTP molecules are hydrolyzed to 2 GDP + 2 P_i during chain elongation. The hydrolysis of GTP is coupled to conformational changes in the translation machinery. In this sense, GTP and GDP act as allosteric modulators. However, unlike most conformational changes induced by allosteric modulators, the conformational changes that occur during protein synthesis are associated with a considerable consumption of energy.

The hydrolysis of four phosphoanhydride bonds represents a large Gibbs free energy change—much more than is required for the formation of a single peptide bond. Most of the "extra" energy compensates for the loss of entropy during protein synthesis. The decrease in entropy is due primarily to the specific ordering of 20 different kinds of amino acids into a polypeptide chain. In addition, entropy is lost when an amino acid is linked to a particular tRNA and when an aminoacyl-tRNA associates with a specific codon.

22.9 Regulation of Protein Synthesis

Ribosomes moving
on messenger RNA
synthesize proteins
haiku by Sydney Brenner (2002)

One way gene expression can be regulated is by controlling the translation of mRNA into protein. Translation can be controlled at initiation, elongation, or termination. In general, translational control of gene expression is used to regulate the production of proteins that assemble into multisubunit complexes and proteins whose expression in the cell must be strictly and quickly controlled.

The rate of translation depends to some extent on the sequence of the template. An mRNA containing an abundance of rare codons, for example, is translated less rapidly (and therefore less frequently) than one containing the most frequently used codons. In addition, the rate of translation initiation varies with the nucleotide sequence at the initiation site. A strong ribosome binding site in bacterial mRNA leads to more efficient initiation. There is also evidence that the nucleotide sequence surrounding the initiation codon in eukaryotic mRNA influences the rate of initiation.

One difference between the initiation of translation and the initiation of transcription is that the formation of a translation complex can be influenced by secondary structure in the message. For example, the formation of intramolecular double-stranded regions in mRNA can mask ribosome binding sites and the initiation codon. Although structural properties can determine whether a given mRNA molecule is translated frequently or infrequently, this is not regulation in the strict sense. We use the term *translational regulation* to refer to cases where extrinsic factors modulate the frequency of mRNA translation.

Polypeptide synthesis is an example of head growth (Box 12.5).

KEY CONCEPT

mRNA codons in the ribosome's A site are also being continually tested by randomly diffusing release factors, which are seeking translation termination codons.

A. Ribosomal Protein Synthesis Is Coupled to Ribosome Assembly in *E. coli*

Every *E. coli* ribosome contains at least 52 ribosomal proteins. The genes encoding these ribosomal proteins are scattered throughout the genome in 13 operons and seven isolated genes. When multiple copies of genes encoding some of these ribosomal proteins are inserted into *E. coli*, the concentrations of the respective mRNAs increase sharply, yet the overall rate of ribosomal protein synthesis scarcely changes. Furthermore, the relative concentrations of ribosomal proteins remain unchanged even though the various mRNA molecules for ribosomal proteins are present in unequal amounts. These findings suggest that the synthesis of ribosomal proteins is tightly regulated at the level of translation.

Translational regulation of ribosomal protein synthesis is crucial since ribosomes cannot assemble unless all the proteins are present in the proper stoichiometry. The production of ribosomal proteins is controlled by regulating the efficiency with which their mRNAs are translated. Each of the large operons containing ribosomal protein genes encodes one ribosomal protein that inhibits translation of its own polycistronic mRNA by binding near the initiation codon of one of the first genes of the operon.

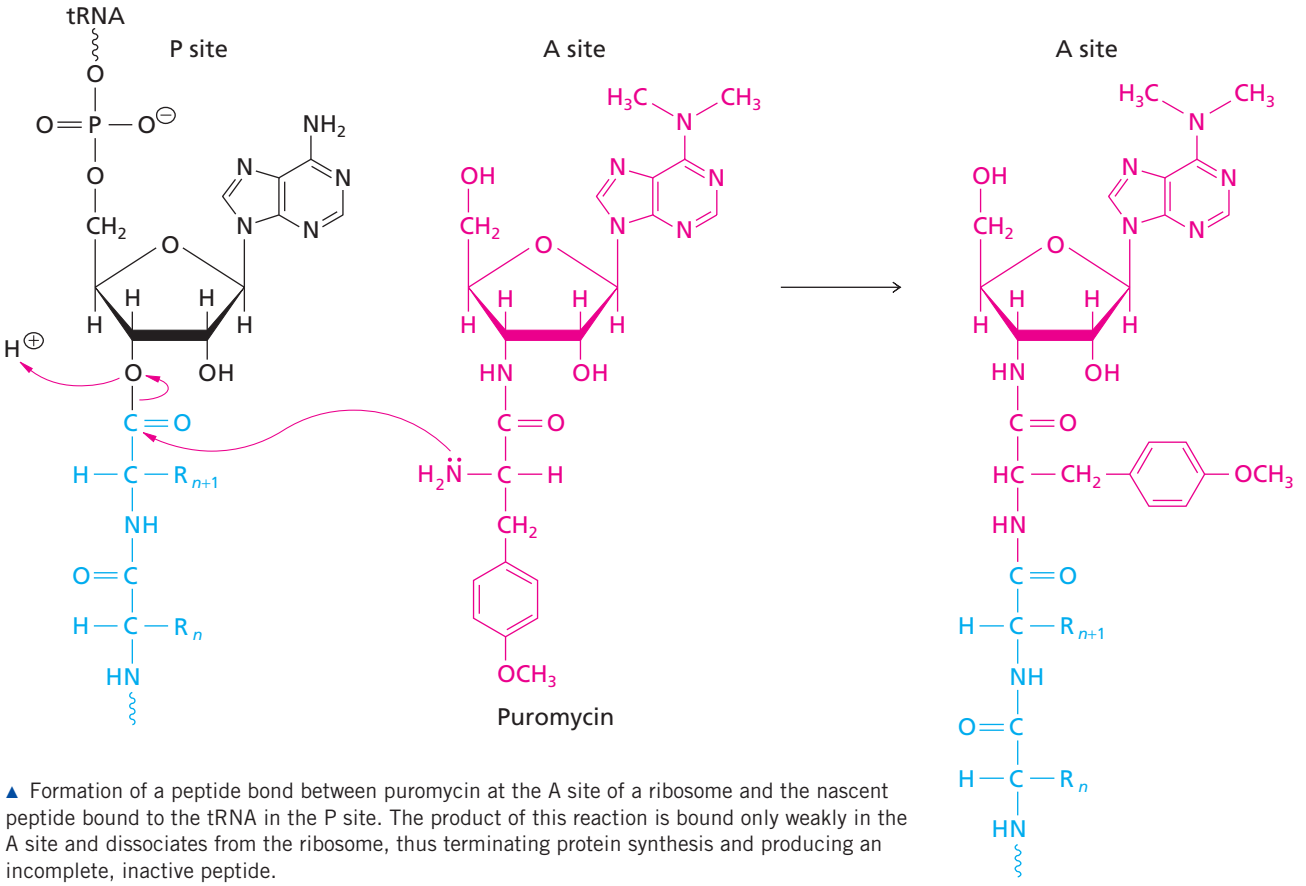
The interactions between the inhibiting ribosomal proteins and their mRNAs may resemble the interactions between these proteins and the ribosomal RNA to which they bind when assembled into mature ribosomes. For example, the mRNA transcript of the *str* operon, which includes the coding region for the ribosomal protein S7, contains some regions of RNA sequence that are identical to the S7 binding site of 16S rRNA. Moreover, the proposed secondary structure of the *str* mRNA resembles the proposed secondary structure of the 16S rRNA S7 binding site (Figure 22.25). S7 binds to this region of the *str* mRNA molecule and inhibits translation. It is likely that S7 recognizes analogous structural features in both RNA molecules. Similar mechanisms regulate the translation of mRNAs that encode the other ribosomal proteins.

The ribosomal proteins that inhibit translation bind more tightly to ribosomal RNA than to the similar sites on messenger RNA. Thus, the mRNA continues to be translated as long as newly synthesized ribosomal proteins are incorporated into ribosomes. However, as soon as ribosome assembly slows and the concentration of free ribosomal proteins increases within the cell, the inhibiting ribosomal proteins bind to their own mRNA molecules and block additional protein synthesis. In this way, synthesis of ribosomal proteins is coordinated with ribosome assembly.

BOX 22.1 SOME ANTIBIOTICS INHIBIT PROTEIN SYNTHESIS

Many microorganisms produce antibiotics, which they use as a chemical defense against competitors. Some antibiotics prevent bacterial growth by inhibiting the formation of peptide bonds. For example, the structure of the antibiotic puromycin closely resembles the structure of the 3′ end of an aminoacyl-tRNA molecule. Because of this similarity, puromycin can enter the A site of a ribosome. Peptidyl transferase then catalyzes the transfer of the nascent polypeptide to the free amino group of puromycin (see figure below). The peptidyl-puromycin is bound weakly in the A site and soon dissociates from the ribosome, thereby terminating protein synthesis.

Although puromycin effectively blocks protein synthesis in prokaryotes, it is not clinically useful since it also blocks protein synthesis in eukaryotes and is therefore poisonous to humans. Clinically important antibiotics, which include streptomycin, chloramphenicol, erythromycin, and tetracycline, are specific for bacteria and have little or no effect on eukaryotic protein synthesis. Streptomycin binds to one of the ribosomal proteins in the 30S subunit and inhibits the initiation of translation. Chloramphenicol interacts with the 50S subunit and inhibits peptidyl transferase. Erythromycin binds to the 50S subunit, inhibiting the translocation step. Tetracycline binds to the 30S subunit, preventing the binding of aminoacyl-tRNA molecules to the A site.



▲ Formation of a peptide bond between puromycin at the A site of a ribosome and the nascent peptide bound to the tRNA in the P site. The product of this reaction is bound only weakly in the A site and dissociates from the ribosome, thus terminating protein synthesis and producing an incomplete, inactive peptide.

B. Globin Synthesis Depends on Heme Availability

The synthesis of hemoglobin, the major protein in red blood cells, requires globin chains and heme in stoichiometric amounts (Section 4.12). One way globin synthesis is controlled is by regulation of translation initiation. Hemoglobin is initially synthesized in immature erythrocytes called rubriblasts. Mammalian rubriblasts lose their nuclei during maturation and eventually become reticulocytes, which are the

immediate precursors of erythrocytes. Hemoglobin continues to be synthesized in reticulocytes that are packed with processed, stable mRNA molecules encoding globin polypeptides.

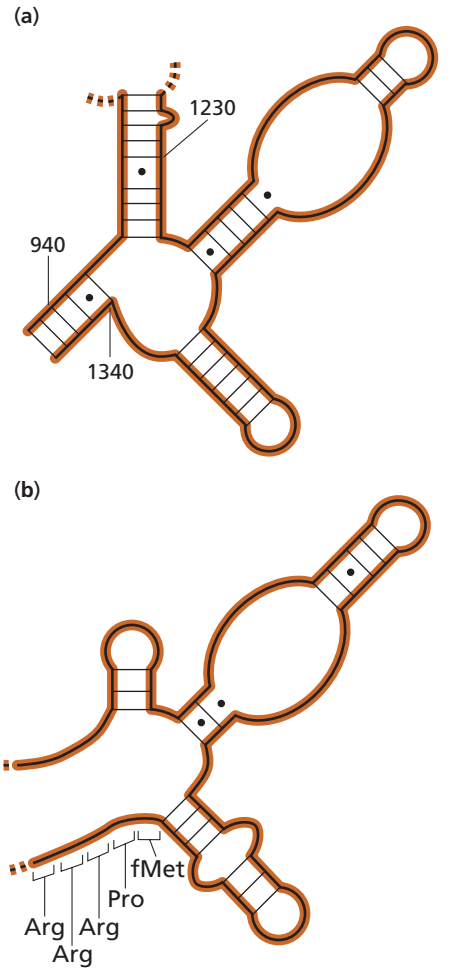
The rate of globin synthesis in reticulocytes is determined by the concentration of heme. When the concentration of heme decreases, the translation of globin mRNA is inhibited. The effect of heme on globin mRNA translation is mediated by a protein kinase called heme-controlled inhibitor (HCI) (Figure 22.26). Active HCI catalyzes transfer of a phosphoryl group from ATP to the translation initiation factor eIF-2. Phosphorylated eIF-2 is unable to participate in translation initiation and protein synthesis in the cell is inhibited.

During the initiation of translation, eIF-2 binds methionyl-$tRNA_i^{Met}$ and GTP. When the preinitiation complex encounters an initiation codon, methionyl-$tRNA_i^{Met}$ is transferred from eIF-2 to the initiation codon of the mRNA. This transfer reaction is accompanied by the hydrolysis of GTP and the release of eIF-2-GDP. An enzyme called guanine nucleotide exchange factor (GEF) catalyzes the replacement of GDP with GTP on eIF-2 and the attachment of another methionyl-$tRNA_i^{Met}$ to eIF-2. GEF binds very tightly to phosphorylated eIF-2–GDP, preventing the nucleotide exchange reaction. Protein synthesis is completely inhibited when all the GEF in the cell is bound because the active eIF-2–GTP complex cannot be regenerated.

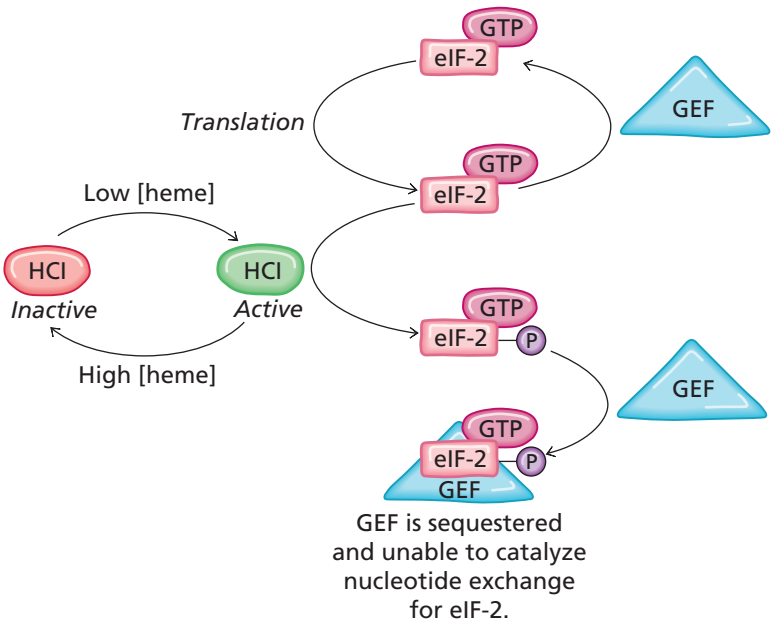
Heme regulates the synthesis of globin by interfering with the activation of HCI. When heme is abundant, HCI is inactive and globin mRNA can be translated. When heme is scarce, however, HCI is activated and translation of all mRNA within the cell is inhibited (Figure 22.26). Phosphorylation of eIF-2 also appears to regulate the translation of mRNA in other mammalian cell types. For example, during infection of human cells by RNA viruses, the presence of double-stranded RNA leads to the production of interferon, which in turn activates a protein kinase that phosphorylates eIF-2. This reaction inhibits protein synthesis in the virus-infected cell.

C. The *E. coli trp* Operon Is Regulated by Repression and Attenuation

The *trp* operon in *E. coli* encodes the proteins necessary for the biosynthesis of tryptophan. Most organisms synthesize their own amino acids but can also obtain them by degrading exogenous proteins. For this reason, most organisms have evolved mechanisms



▲ **Figure 22.25**
Comparison of proposed secondary structures of S7 binding sites. (a) S7 binding site on 16S rRNA. **(b)** S7 binding site on the *str* mRNA molecule.

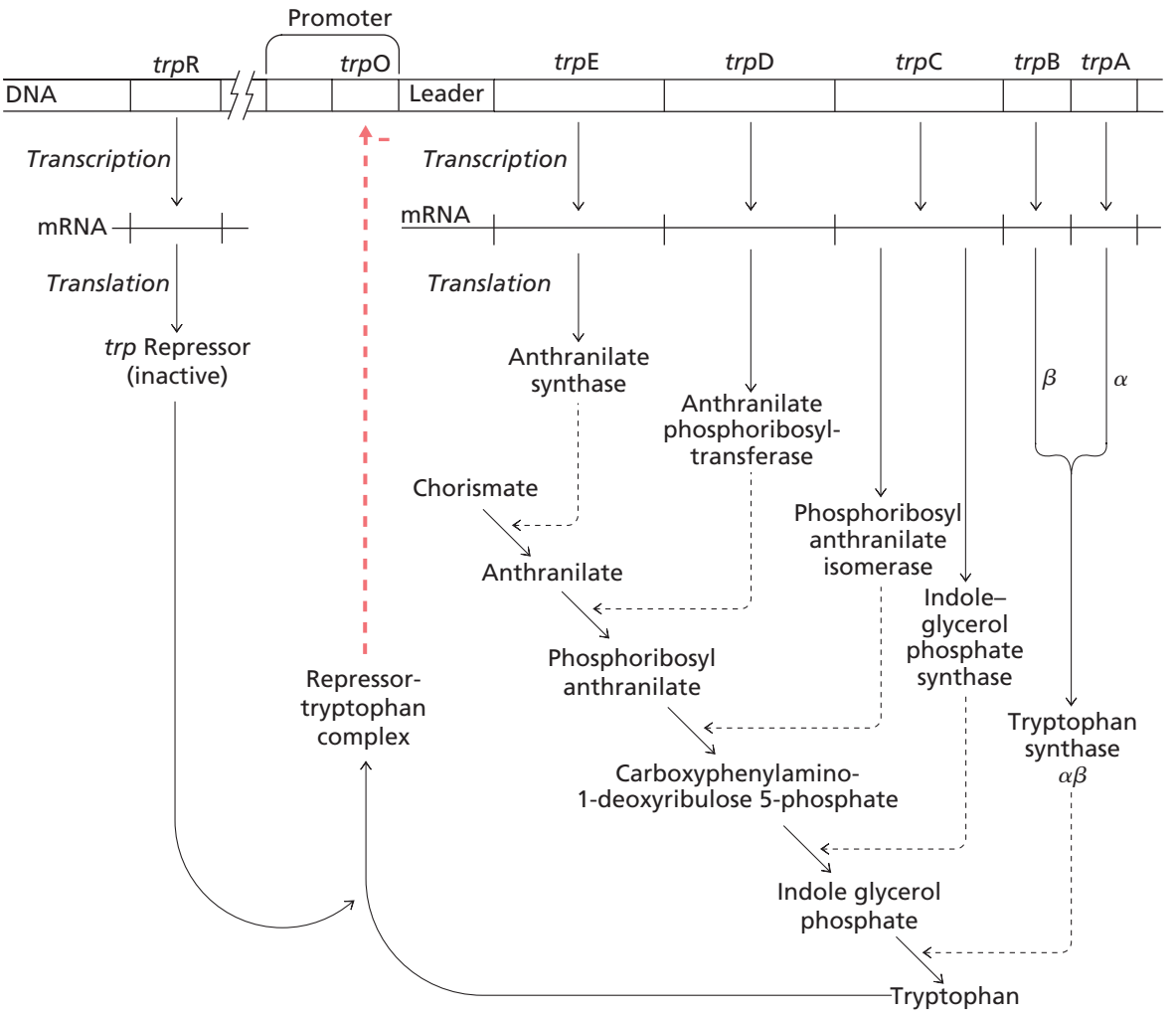


◄ **Figure 22.26**
Inhibition of protein synthesis by phosphorylation of eIF-2 in reticulocytes. When the concentration of heme is high, HCI is inactive and translation proceeds normally. When the concentration of heme is low, HCI catalyzes the phosphorylation of eIF-2. Phosphorylated eIF-2 binds the limiting amounts of GEF in the cell very tightly, sequestering the GEF and preventing translation of cellular mRNAs (including the globins).

to repress the synthesis of the enzymes required for *de novo* amino acid biosynthesis when the amino acid is available from exogenous sources. For example, in *E. coli,* tryptophan is a negative regulator of its own biosynthesis. In the presence of tryptophan, the *trp* operon is not expressed (Figure 22.27). Expression of the *trp* operon is inhibited in part by *trp* repressor, a dimer of two identical subunits. *trp* repressor is encoded by the *trpR* gene, which is located elsewhere on the bacterial chromosome and is transcribed separately. When tryptophan is abundant, a repressor-tryptophan complex binds to the operator *trpO*, which lies within the promoter. The bound repressor-tryptophan complex prevents RNA polymerase from binding to the promoter. Tryptophan is thus a corepressor of the *trp* operon.

Regulation of the *E. coli trp* operon is supplemented and refined by a second, independent mechanism called **attenuation**. This second mechanism depends on translation and helps determine whether transcription of the *trp* operon proceeds or terminates prematurely. The movement of RNA polymerase from the promoter into the *trpE* gene is governed by a 162 nucleotide sequence that lies between the promoter and *trpE*. This sequence, called the leader region (Figure 22.27), includes a stretch of 45 nucleotides that encodes a 14 amino acid peptide called the *leader peptide*. The mRNA transcript of the leader region contains two consecutive codons specifying tryptophan near the end of the coding region for the leader peptide. In addition, the

▼ Figure 22.27
Repression of the *E. coli trp* operon. The *trp* operon is composed of a leader region and five genes required for the biosynthesis of tryptophan from chorismate. The *trpR* gene, located upstream of the *trp* operon (*trpO*), encodes *trp* repressor, which is inactive in the absence of its corepressor, tryptophan. When tryptophan is present in excess, it binds to *trp* repressor, and the repressor-tryptophan complex binds to the *trp* operator (*trpO*). Once bound to the operator, the repressor-tryptophan complex prevents further transcription of the *trp* operon by excluding RNA polymerase from the promoter.



leader region contains four GC-rich sequences. The codons that specify tryptophan and the four GC-rich sequences regulate the synthesis of mRNA by affecting transcription termination.

When transcribed into mRNA, the four GC-rich sequences of the leader region can base-pair to form one of two alternative secondary structures (Figure 22.28, on the next page). The first possible secondary structure includes two RNA hairpins. These hairpins form between the sequences labeled 1 and 2 and between those labeled 3 and 4 in Figure 22.28a. The 1-2 hairpin is a typical transcription pause site. The 3-4 hairpin is followed by a string of uridylate residues, which is a typical rho-independent termination signal (Section 21.4). This particular termination signal is unusual, however, because it occurs upstream of the first gene in the *trp* operon. The other possible secondary structure includes a single RNA hairpin between sequences 2 and 3. This hairpin, which is more stable than the 3-4 hairpin, forms only when sequence 1 is not available for hairpin formation with sequence 2.

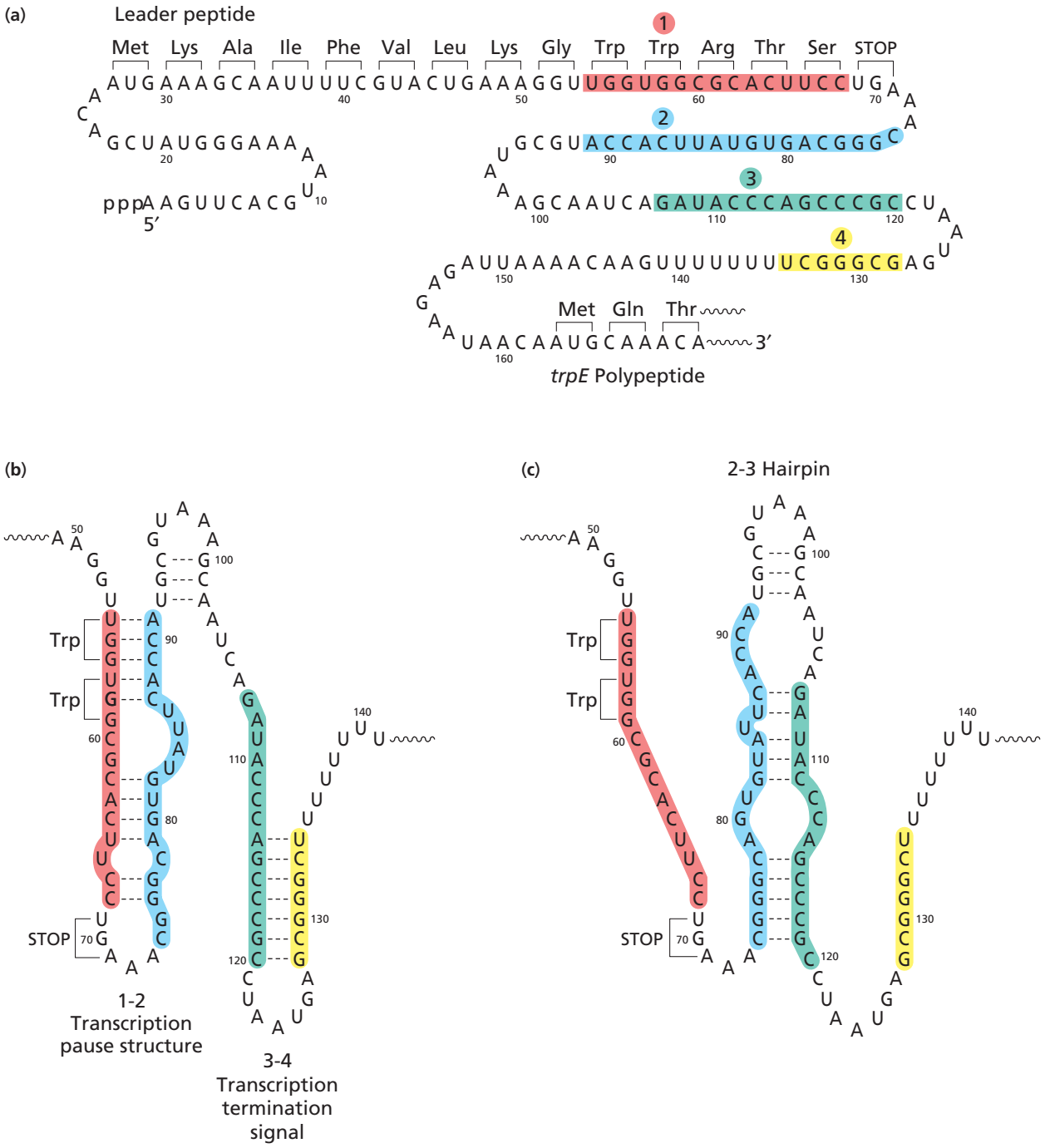
During transcription of the leader region, RNA polymerase pauses when the 1-2 hairpin forms. While RNA polymerase pauses, a ribosome initiates translation of the mRNA encoding the leader peptide. This coding region begins just upstream of the 1-2 RNA hairpin. Sequence 1 encodes the C-terminal amino acids of the leader peptide and also contains a termination codon. As the ribosome translates sequence 1, it disrupts the 1-2 hairpin, thereby releasing the paused RNA polymerase, which then transcribes sequence 3. In the presence of tryptophanyl-$tRNA^{Trp}$, the ribosome and RNA polymerase move at about the same rate. When the ribosome encounters the termination codon of the *trp* leader mRNA, it dissociates and the 1-2 hairpin re-forms. After the ribosome has disassembled, RNA polymerase transcribes sequence 4, which forms a transcription termination hairpin with sequence 3. This termination signal causes the transcription complex to dissociate from the DNA template before the genes of the *trp* operon have been transcribed.

When tryptophan is scarce, however, the ribosome and RNA polymerase do not move synchronously. When the concentration of cellular tryptophan falls, the cell becomes deficient in tryptophanyl-$tRNA^{Trp}$. Under these circumstances, the ribosome pauses when it reaches the two codons specifying tryptophan in sequence 1 of the mRNA molecule. RNA polymerase, which has already been released from the 1-2 pause site, transcribes sequences 3 and 4. While the ribosome is stalled and sequence 1 is covered, sequence 2 forms a hairpin loop with sequence 3. Since the 2-3 hairpin is more stable than the 3-4 hairpin, sequence 3 does not pair with sequence 4 to form the transcription termination hairpin. Under these conditions, RNA polymerase passes through the potential termination site (UGA in Figure 22.28a), and the rest of the *trp* operon is transcribed.

Attenuation appears to be a regulatory mechanism that has evolved relatively recently and is found only in enteric bacteria, such as *E. coli.* (Attenuation cannot occur in eukaryotes because transcription and translation take place in different parts of the cell.) Several *E. coli* operons, including the *phe, thr, his, leu,* and *ile* operons, are regulated by attenuation. Some operons, such as the *trp* operon, combine attenuation with repression, whereas others, such as the *his* operon, are regulated solely by attenuation. The leader peptides of operons whose genes are involved in amino acid biosynthesis may contain as many as seven codons specifying a particular amino acid.

22.10 Post-Translational Processing

As the translation complex moves along the mRNA template in the 5′ → 3′ direction, the nascent polypeptide chain lengthens. The 30 or so most recently polymerized amino acid residues remain buried in the ribosome, but amino acid residues closer to the N-terminus are extruded from the ribosome. The N-terminal residues start to fold into the native protein structure even before the C-terminus of the protein has



▲ **Figure 22.28**

***trp* leader region.** **(a)** mRNA transcript of the *trp* leader region. This 162 nucleotide mRNA sequence includes four GC-rich sequences and the coding region for a 14 amino acid leader peptide. The coding region includes two consecutive codons specifying tryptophan. The four GC-rich sequences can base-pair to form one of two alternative secondary structures. **(b)** Sequence 1 (red) and sequence 2 (blue) are complementary and, when base-paired, form a typical transcription pause site. Sequence 3 (green) and sequence 4 (yellow) are complementary and, when base-paired, form a rho-independent termination site. **(c)** Sequences 2 and 3 are also complementary and can form an RNA hairpin that is more stable than the 3-4 hairpin. This structure forms only when sequence 1 is not available for hairpin formation with sequence 2.

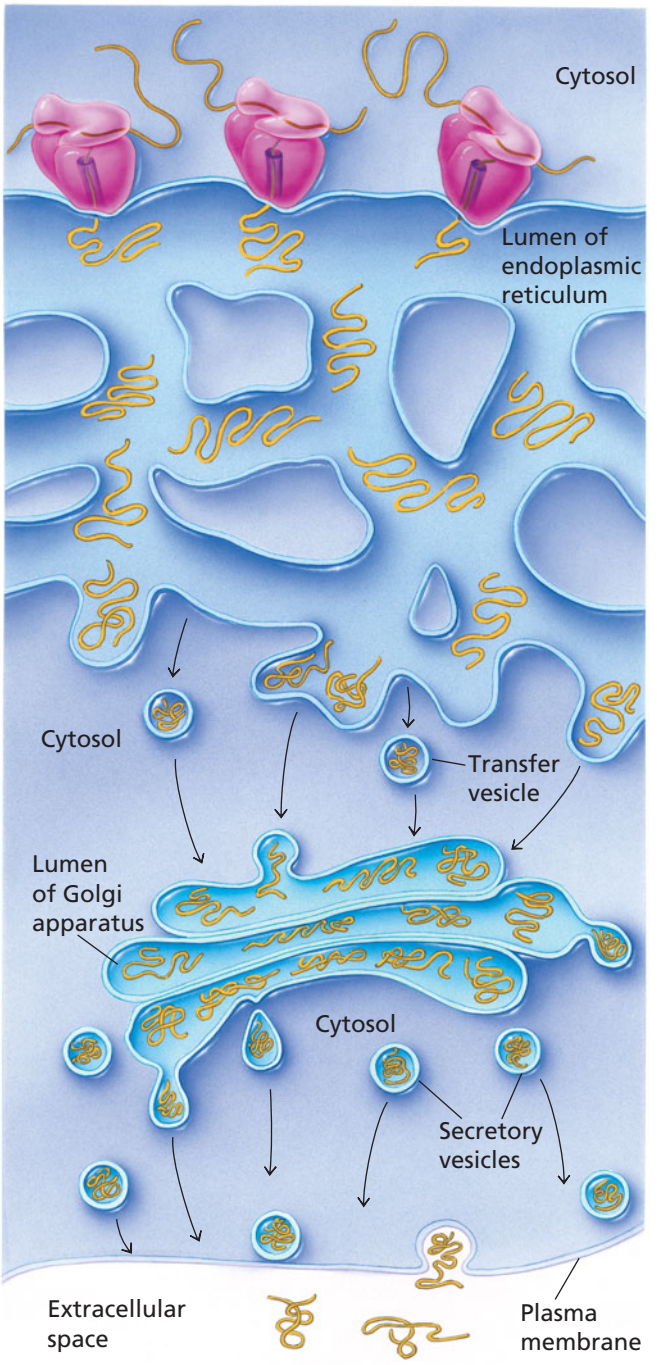
been synthesized. As these residues fold, they are acted on by enzymes that modify the nascent chain.

Modifications that occur before the polypeptide chain is complete are said to be **cotranslational**, whereas those that occur after the chain is complete are said to be **post-translational**. Some examples from the multitude of cotranslational and post-translational modifications include deformylation of the N-terminal residue in prokaryotic

proteins, removal of the N-terminal methionine from prokaryotic and eukaryotic proteins, formation of disulfide bonds, cleavage by proteinases, phosphorylation, addition of carbohydrate residues, and acetylation.

One of the most important events that occurs co- and post-translationally is the processing and transport of proteins through membranes. Protein synthesis occurs in the cytosol, but the mature forms of many proteins are embedded in membranes or are inside membrane bounded compartments. For example, many receptor proteins are embedded in the external membrane of the cell, with the bulk of the protein outside the cell. Other proteins are secreted from cells, and still others reside in lysosomes and other organelles inside eukaryotic cells. In each case, the protein synthesized in the cytosol must be transported across a membrane barrier. In fact, such proteins are synthesized by membrane bound ribosomes that are attached to the plasma membrane in bacteria and to the endoplasmic reticulum in eukaryotic cells.

The best-characterized transport system is the one that carries proteins from the cytosol to the plasma membrane for secretion (Figure 22.29). In eukaryotes, proteins destined for secretion are transported across the membrane of the endoplasmic reticulum into the lumen, which is topologically equivalent to the cell exterior. Once the protein has been transported into the endoplasmic reticulum, it can be transported by vesicles through the Golgi apparatus to the plasma membrane for release outside the cell.



▲ **Figure 22.29**
Secretory pathway in eukaryotic cells. Proteins whose synthesis begins in the cytosol are transported into the lumen of the endoplasmic reticulum. After further modification in the Golgi apparatus, the proteins are secreted.

A. The Signal Hypothesis

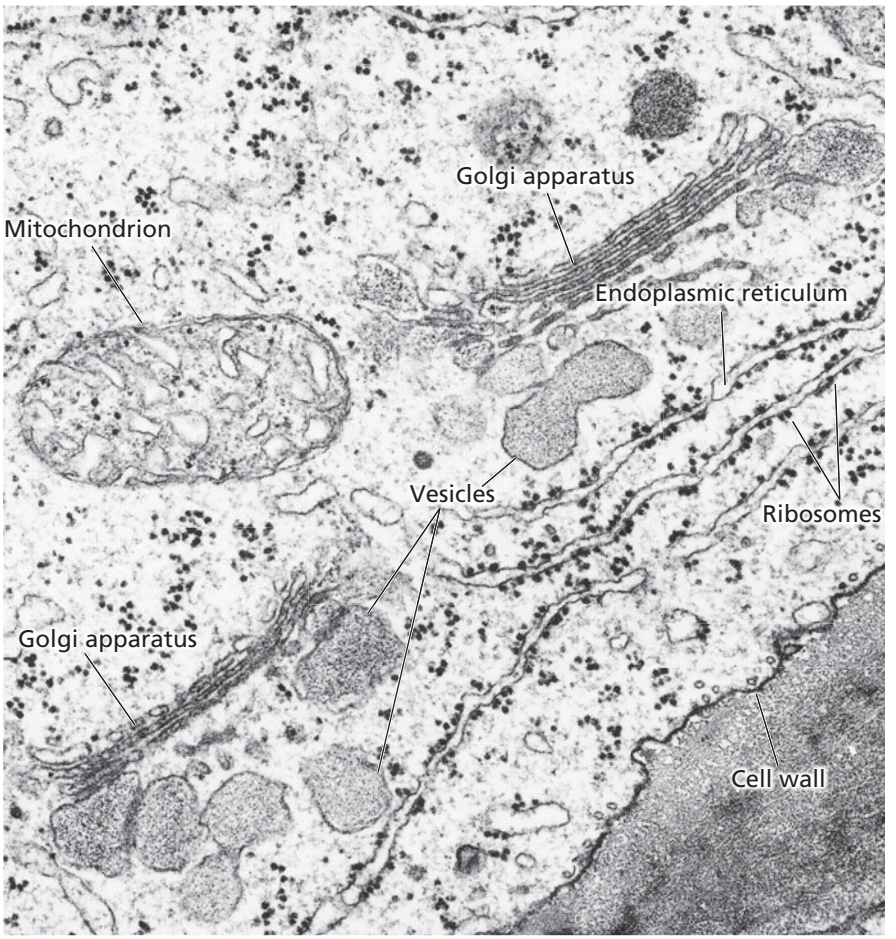
Secreted proteins are synthesized on the surface of the endoplasmic reticulum, and the newly synthesized protein is passed through the membrane into the lumen. In cells that make large amounts of secreted protein, the endoplasmic reticulum membranes are covered with ribosomes (Figure 22.30, on the next page).

The clue to the process by which many proteins cross the membrane of the endoplasmic reticulum appears in the first 20 or so residues of the nascent polypeptide chain. In most membrane bound and secreted proteins, these residues are present only in the nascent polypeptide, not in the mature protein. The N-terminal sequence of residues that is proteolytically removed from the protein precursor is called the **signal peptide** since it is the portion of the precursor that signals the protein to cross a membrane. Signal peptides vary in length and composition, but they are typically from 16 to 30 residues long and include 4 to 15 hydrophobic residues (Figure 22.31, on the next page).

In eukaryotes, many proteins destined for secretion appear to be translocated across the endoplasmic reticulum by the pathway shown in Figure 22.32 on page 693. In the first step, an 80S initiation complex—including a ribosome, a Met-tRNA$_i^{Met}$ molecule, and an mRNA molecule—forms in the cytosol. Next, the ribosome begins translating the mRNA and synthesizing the signal peptide at the N-terminus of the precursor. Once the signal peptide has been synthesized and extruded from the ribosome, it binds to a protein-RNA complex called a signal recognition particle (SRP).

SRP is a small ribonucleoprotein containing a 300 nucleotide RNA molecule called 7SL RNA and four proteins. SRP recognizes and binds to the signal peptide as it emerges from the ribosome. When SRP binds, further translation is blocked. The SRP-ribosome complex then binds to an SRP receptor protein (also known as docking protein) on the cytosolic face of the endoplasmic reticulum. The ribosome is anchored to the membrane of the endoplasmic reticulum by ribosome binding proteins called translocons, and the signal peptide is inserted into the membrane at a pore that is part

Figure 22.30 ►
Secretory vesicles in a maize rootcap cell. Large secretory vesicles containing proteins are budding off the Golgi apparatus (center). Note the abundance of ribosomes bound to the endoplasmic reticulum.



of the complex formed by the endoplasmic reticulum proteins at the docking site. Once the ribosome-SRP complex is bound to the membrane, the inhibition of translation is relieved and SRP dissociates in a reaction coupled to GTP hydrolysis. Thus, the role of SRP is to recognize nascent polypeptides containing a signal peptide and to target the translation complex to the surface of the endoplasmic reticulum.

Once the translation complex is bound to the membrane, translation resumes and the new polypeptide chain passes through the membrane. The signal peptide is then cleaved from the nascent polypeptide by a signal peptidase, an integral membrane protein associated with the pore complex. The transport of proteins across the membrane

Figure 22.31 ▼
Signal peptides from secreted proteins. Hydrophobic residues are shown in blue, and arrows mark the sites where the signal peptide is cleaved from the precursor. (OmpA is a bacterial membrane protein.)

Prelysozyme

$H_3\overset{\oplus}{N}$–Met–Arg–Ser–Leu–Leu–Ile–Leu–Val–Leu–Cys–Phe–Leu–Pro–Leu–Ala–Ala–Leu–Gly↓–Gly ~~

Preproalbumin

$H_3\overset{\oplus}{N}$–Met–Lys–Trp–Val–Thr–Phe–Leu–Leu–Leu–Leu–Phe–Ile–Ser–Gly–Ser–Ala–Phe–Ser↓–Arg ~~

Alkaline phosphatase

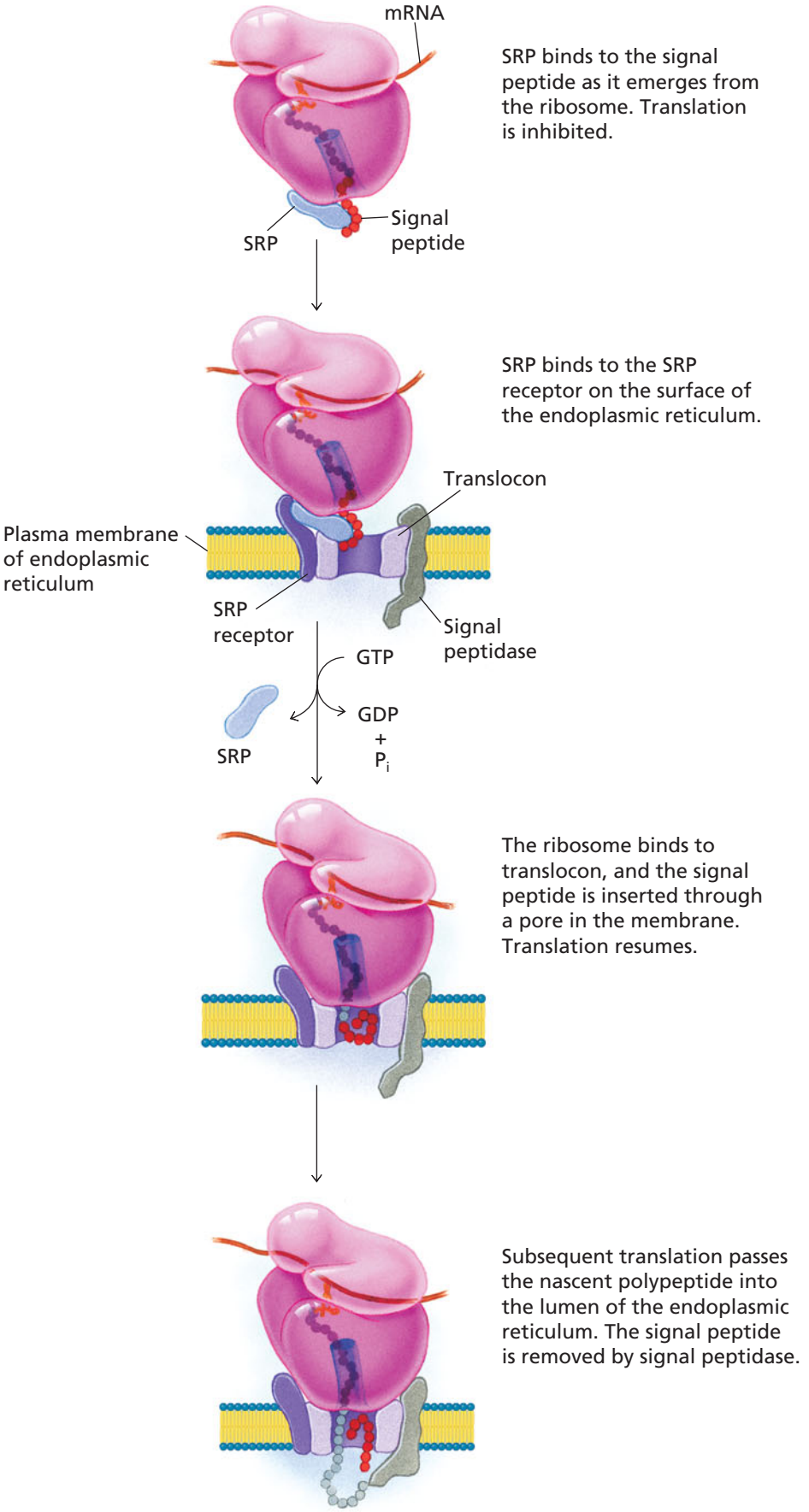
$H_3\overset{\oplus}{N}$–Met–Lys–Gln–Ser–Thr–Ile–Ala–Leu–Ala–Leu–Leu–Pro–Leu–Leu–Phe–Thr–Pro–Val–Thr–Lys–Ala↓–Arg ~~

Maltose-binding protein

$H_3\overset{\oplus}{N}$–Met–Lys–Ile–Lys–Thr–Gly–Ala–Arg–Ile–Leu–Ala–Leu–Ser–Ala–Leu–Thr–Thr–Met–Met–Phe–Ser–Ala–Ser–Ala–Leu–Ala↓–Lys ~~

OmpA

$H_3\overset{\oplus}{N}$–Met–Lys–Lys–Thr–Ala–Ile–Ala–Ile–Ala–Val–Ala–Leu–Ala–Gly–Phe–Ala–Thr–Val–Ala–Gln–Ala↓–Ala ~~



◀ **Figure 22.32**
Translocation of eukaryotic proteins into the lumen of the endoplasmic reticulum.

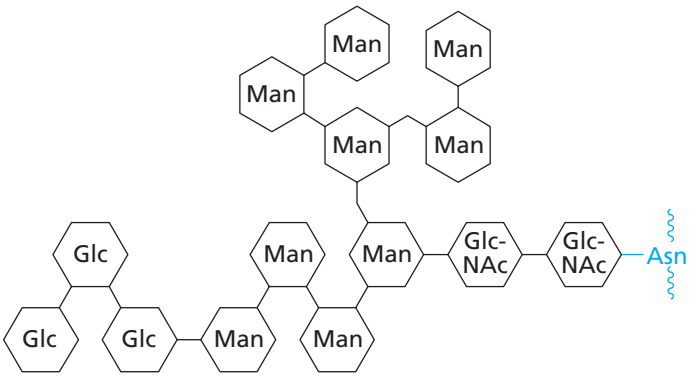


Figure 22.33 ►
Structure of a complex oligosaccharide linked to an asparagine residue. Abbreviations: Glc, glucose; GlcNAc, *N*-acetylglucosamine; Man, mannose.

is assisted by chaperones in the lumen of the endoplasmic reticulum. In addition to their role in protein folding, chaperones are required for translocation, and their activity requires ATP hydrolysis. When protein synthesis terminates, the ribosome dissociates from the endoplasmic reticulum, and the translation complex disassembles.

B. Glycosylation of Proteins

Many integral membrane proteins and secretory proteins contain covalently bound oligosaccharide chains. The addition of these chains to proteins is called protein glycosylation (Section 8.7C). Protein glycosylation is one of the major metabolic activities of the lumen of the endoplasmic reticulum and of the Golgi apparatus and is an extension of the general process of protein biosynthesis. A glycoprotein can contain dozens, indeed hundreds, of monosaccharide units. The mass of the carbohydrate portion may account for as little as 1% or as much as 80% of the mass of the glycoprotein.

A common glycosylation reaction involves the covalent attachment of a complex oligosaccharide to the side chain of an asparagine residue (Figure 22.33). During subsequent transit through the endoplasmic reticulum and the Golgi apparatus, proteins may be covalently modified in many ways, including the formation of disulfide bonds and proteolytic cleavage. The complex oligosaccharides attached to the proteins are likewise modified during transit. A variety of different oligosaccharides can be covalently bound to proteins. In some cases, the structure of the oligosaccharide acts as a signal to target proteins to a specific location. For example, lysosomal proteins contain sites for the attachment of an oligosaccharide that targets these proteins to the lysosome. By the time they have traversed the Golgi apparatus, the proteins and their oligosaccharides are usually fully modified.

Summary

1. The genetic code consists of nonoverlapping, three-nucleotide codons. The code is unambiguous and degenerate; the first two nucleotides of the three-letter code are often sufficient; codons with similar sequences specify chemically similar amino acids; and there are special codons for the initiation and termination of peptide synthesis.
2. tRNA molecules are the adapters between mRNA codons and amino acids in proteins. All tRNA molecules have a similar cloverleaf secondary structure with a stem and three arms. The tertiary structure is L-shaped. The anticodon loop is at one end of the structure, and the acceptor stem is at the other. The anticodon in tRNA base-pairs with a codon in mRNA. The 5′ (wobble) position of the anticodon is conformationally flexible.
3. An aminoacyl-tRNA synthetase catalyzes the addition of a specific amino acid to the acceptor stem of the appropriate tRNA, producing an aminoacyl-tRNA. Some aminoacyl-tRNA synthetases carry out proofreading.
4. Ribosomes are the RNA-protein complexes that catalyze the polymerization of amino acids bound to aminoacyl-tRNA molecules. All ribosomes are composed of two subunits: prokaryotic ribosomes contain three rRNA molecules, and eukaryotic ribosomes contain four. The growing polypeptide chain is attached to a tRNA in the peptidyl (P) site of the ribosome, and the aminoacyl-tRNA molecule bearing the next amino acid to be added to the nascent polypeptide chain docks in the aminoacyl (A) site.
5. Translation begins with the formation of an initiation complex consisting of an initiator tRNA, the mRNA template, the ribosomal subunits, and several initiation factors. In prokaryotes, initiation occurs just downstream of Shine-Dalgarno sequences; in eukaryotes, initiation usually occurs at the initiation codon closest to the 5′ end of the mRNA.
6. The elongation step of translation requires accessory proteins called elongation factors. The three steps of elongation are (1) positioning of the correct aminoacyl-RNA in the A site,

(2) formation of the peptide bond by peptidyl transferase, and (3) translocation of the ribosome by one codon.

7. Release factors recognize termination codons and catalyze the termination of protein synthesis and disassembly of the translation complex.
8. Protein synthesis requires the energy of four phosphoanhydride bonds per residue.
9. The regulation of translation includes the formation of secondary structure in mRNA that influences the rate of initiation. Ribosomal RNA proteins can inhibit translation of their own mRNA by binding to such sites. Phosphorylation of an initiation factor regulates globin synthesis. Regulation of expression of the *E. coli trp* operon involves attenuation, in which translation of a leader mRNA governs transcription of the operon.
10. Many proteins are post-translationally modified. Some eukaryotic proteins destined for secretion contain N-terminal signals for transport into the endoplasmic reticulum. Many membrane and secreted proteins are glycosylated.

Problems

1. The standard genetic code is read in codons that are three nucleotides long. How many potential reading frames are there on a single piece of double-stranded DNA? If instead the genetic code was read in codons that were four nucleotides long, how many reading frames would there be on the same piece of double-stranded DNA?
2. Examine the sequences of the mRNAs transcribed from the DNA sequence in Problem 11 in Chapter 21. Assuming that the DNA segment is from the middle of a protein-coding gene, which of the possible mRNAs is most likely to be the actual transcript? What is the sequence of the encoded peptide?
3. Calculate the number of phosphoanhydride bonds that are hydrolyzed during synthesis of a 600 amino acid residue protein in *E. coli*. Do not include the energy required to synthesize the amino acids, mRNA, tRNA, or the ribosomes.
4. Polypeptide chain elongation on the ribosome can be broken down into three discrete steps (the microcycle): (1) binding of the correct aminoacyl-tRNA in the ribosome's A site, (2) peptide bond formation, and (3) translocation. What, specifically, is it that gets translocated in the third step of this cycle?
5. A prokaryotic mRNA may contain many AUG codons. How does the ribosome distinguish AUG codons specifying initiation from AUG codons specifying internal methionine?
6. Given that the genetic code is universal, would a plant mRNA be correctly translated in a prokaryotic cell like *E. coli*?
7. Bacterial genomes usually contain multiple copies of the genes for rRNA. These are transcribed very efficiently in order to produce large amounts of rRNA for assembly into ribosomes. In contrast, the genes that encode ribosomal proteins are present only as single copies. Explain the difference in the number of copies of rRNA and ribosomal protein genes.
8. Suppressor mutations suppress the effects of other mutations. For example, mutations that produce the stop codon UAG in the middle of a gene are suppressed by an additional mutation in a tRNA gene that gives rise to a mutant anticodon with the sequence CUA. Consequently, an amino acid is inserted at the mutant stop codon, and a protein is synthesized (although it may be only partially active). List all the tRNA species that could be mutated to a suppressor of UAG mutations by a single base change in the anticodon. How can a cell with a suppressor tRNA survive?
9. Transfer RNAs are absolutely essential for polypeptide synthesis. After reviewing the material in this chapter, name five different cellular components that can bind to (interact with) tRNA molecules.
10. On rare occasions, the translation machinery encounters a codon that cannot be quickly interpreted because of the lack of a particular tRNA or release factor. In these cases, the ribosome may pause and then shift by a single nucleotide and begin translating a different reading frame. Such an occurrence is known as translational frameshifting. The *E. coli* release factor RF-2, which is translated from mRNA that contains an internal UGA stop codon, is produced by translational frameshifting. Explain how this phenomenon might regulate RF-2 production.
11. The mechanism of attenuation requires the presence of a leader region. Predict the effect of the following changes on regulation of the *trp* operon:
    (a) The entire leader region is deleted.
    (b) The sequence encoding the leader peptide is deleted.
    (c) The leader region, an AUG codon, is mutated.
12. In Chapter 21, you learned of many different regulatory mechanisms that control transcription of the *lac* operon in *E. coli*. In Chapter 22, one of the mechanisms of translational regulation discussed was called attenuation. Would you predict that in some other bacterial species the *lac* operon might have evolved such that an attenuation mechanism was used to regulate expression levels from this operon?
13. In the operons that contain genes for isoleucine biosynthesis, the leader regions that precede the genes contain multiple codons that specify not only isoleucine but valine and leucine as well. Suggest a reason why this is so.
14. Suggest the steps involved in the synthesis and processing of a glycosylated, eukaryotic integral membrane protein with a C-terminal cytosolic domain and an N-terminal extracellular domain.
15. In Chapter 23, you will learn about recombinant DNA techniques that allow genes to be cut and pasted at will. If you could remove the coding region for a secretion signal sequence from one protein and place it such that it will now occupy the N-terminus of a cytosolic protein (e.g., β-galactosidase), would you expect the new hybrid protein to enter the cell's secretory pathway?
16. In some species of bacteria, the codon GUG initiates protein synthesis (e.g., LacI, Figure 22.17a). The completed proteins always contain methionine at the N-terminus. How can the initiator tRNA base-pair with the codon GUG? How is this phenomenon related to wobble?

Selected Readings

Aminoacyl-tRNA Synthetases

Carter, C. W., Jr. (1993). Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* 62:715–748.

Ibba, M., and Söll, D. (2000). Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* 69:617–650.

Jakubowski, H., and Goldman, E. (1992). Editing of errors in selection of amino acids for protein synthesis. *Microbiol. Rev.* 56:412–429.

Kurland, C. G. (1992). Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* 26:29–50.

Schimmel, P., and Ribas de Pouplana, L. (2000). Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* 25:207–209.

Ribosomes and Translation

Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4Å resolution. *Science* 289:905–919.

Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000). Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407:340–348.

Garrett, R. A., Douthwate, S. R., Matheson A. T., Moore, P. B., and Noller, H. F., eds. (2000). *The Ribosome: Structure, Function, Antibiotics and Cellular Interactions* (Washington, DC: American Society for Microbiology).

Hanawa-Suetsugu, K., Sekine, S., Sakai, H., Hori-Takemoto, C., Tevader, T., Unzai, S., Tame, J.R.H., Kuramitsu, S., Shirouzu, M., and Yokoyama, S. (2004). Crystal structure of elongation factor P from *Thermus thermophilus* HB8. *Proc. Natl. Acad. Sci.* 101:9595–9600.

Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996). The structure of the *Escherichia coli* EF-Tu • EF-Ts complex at 2.5 Å resolution. *Nature* 379:511–518.

Moore, P. B., and Steitz, T. A. (2003). The structural basis of large ribosomal subunit function. *Annu. Rev. Biochem.* 72:813–850.

Nirenberg, M.W., and Matthaei, J.H., (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci.* 47:1588–1602.

Noller, H. F. (1993). Peptidyl transferase: protein, ribonucleoprotein, or RNA? *J. Bacteriol.* 175:5297–5300.

Pestova, T. V., and Hellen, C. U. T. (1999). Ribosome recruitment and scanning: what's new? *Trends Biochem. Sci.* 24:85–87.

Ramakrishnan, V. (2009). Unravelling the structure of the ribosome. Nobel Lecture 135–160.

Selmer, M., Al-Karadaghi, S., Hirokawa, G., Kaji, A., and Liljas, A. (1999). Crystal Structure of *Thermotoga maritima* ribosome recycling factor: A tRNA mimic. *Science* 286:2349–2352.

Steitz, T.A. (2009). From the structure and function of the ribosome to new antibiotics. Nobel Lecture 179–204.

Regulation of Translation

Kozak, M. (1992). Regulation of translation in eukaryotic systems. *Annu. Rev. Cell Biol.* 8:197–225.

McCarthy, J. E. G., and Gualerzi, C. (1990). Translational control of prokaryotic gene expression. *Trends Genet.* 6:78–85.

Merrick, W. C. (1992). Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* 56:291–315.

Rhoads, R. E. (1993). Regulation of eukaryotic protein synthesis by initiation factors. *J. Biol. Chem.* 268:3017–3020.

Samuel, C. E. (1993). The eIF-2a protein kinases, regulators of translation in eukaryotes from yeasts to humans. *J. Biol. Chem.* 268:7603–7606.

Post-translational Modification

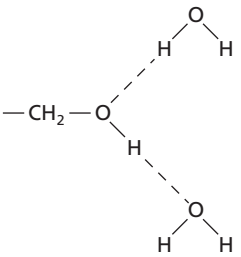
Hurtley, S. M. (1993). Hot line to the secretory pathway. *Trends Biochem. Sci.* 18:3–6.

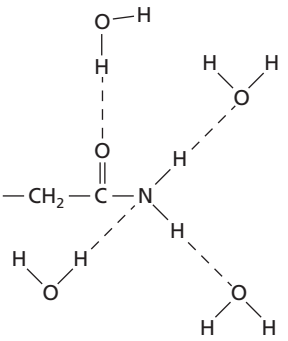
Parodi, A. J. (2000). Protein glycosylation and its role in protein folding. *Annu. Rev. Biochem.* 69:69–93.

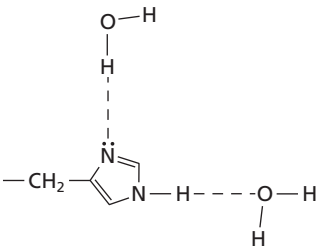
Solutions

Chapter 2 Water

1. Hydrogen bonds involve strongly electronegative atoms such as nitrogen, oxygen, or sulfur.

(a) 

(b) 

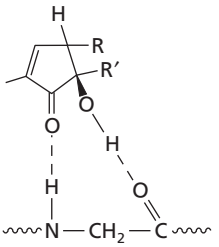
(c) 

2. (a) Glycerol is polar; it is not amphipathic; and it readily dissolves in water.
(b) Hexadecanoyl phosphate is polar; it is amphipathic; and it does not readily dissolve in water but forms micelles.
(c) Laurate is polar; it is amphipathic; and it does not readily dissolve in water but forms micelles.
(d) Glycine is polar; it is not amphipathic; and it readily dissolves in water.

3. There is a larger osmotic pressure inside the cells than outside because the molar concentration of solutes is much greater inside cells than outside. This results in a diffusion of water into cells, causing them to swell and burst.

4. If the pH of a solution is *below* the pK_a of any given ionizable group, the predominant species will be the one with the dissociable proton *on that group*. If the pH of a solution is *above* the pK_a of any given ionizable group, the predominant species will be the one with the dissociable proton *off of that group*.
(a) pH = 11 where the $-COO^{\ominus}$ form predominates.
(b) pH = 2 where the $H^{\oplus}$ form predominates.
(c) pH = 2 where the $H^{\oplus}$ form predominates.
(d) pH = 11 where the $R-O^{\ominus}$ form predominates.

5. (a) Tomato juice. For pH = 4.2, if pH = $-\log [H^{\oplus}]$, then
$[H^{\oplus}] = 10^{-pH}[H^{\oplus}] = 10^{-4.2} = 6.3 \times 10^{-5}$ M.
The ion-product constant of water (K_w) relates the concentrations of $OH^{\ominus}$ and $H^{\oplus}$ (Equation 2.6).
$[OH^{\ominus}] = K_w)/[H^{\oplus}] = 1.0 \times 10^{-14}\ M^2/6.3x^{-5}\ M = 1.6 \times 10^{-10}$ M.
(b) Human blood plasma. If the pH = 7.4, then
$[H^{\oplus}] = 10^{-7.4} = 4.0 \times 10^{-8}$ M. $[OH^{\ominus}] = K_w/[H^{\oplus}] =$
$1.0 \times 10^{-14}\ M^2/4.0x^{-8}\ M = 2.5 \times 10^{-7}$ M.
(c) 1M Ammonia. If the pH = 11.6, then
$[H^{\oplus}] = 10^{-11.6} = 2.5 \times 10^{-12}$ M. $[OH^{\ominus}] = K_w/[H^{\oplus}] =$
$1.0 \times 10^{-14}\ M^2/2.0x^{-12}\ M = 4 \times 10^{-3}$ M.

6. 

7. The total buffer species = [weak acid (HA)] + [conjugate base ($A^{\ominus}$)]

$$\text{Total buffer concentration} = 0.25\ \text{M} + 0.15\ \text{M} = 0.4\ \text{M}$$

The pH can be calculated from the pK_a and the concentrations given using the Henderson-Hasselbalch equation.

$$\text{pH} = \text{p}K_a + \log\frac{[A^{\ominus}]}{[HA]} = 3.90 + \log\frac{(0.15\ \text{M})}{(0.25\ \text{M})} = 3.90 - 0.22 = 3.68$$

8. The pK_a for the ionization of $H_2PO_4^{\ominus}$ is 7.2. The Henderson-Hasselbalch equation (Equation 2.18) indicates that when the concentrations of the acidic form ($H_2PO_4^{\ominus}$) and its conjugate base ($HPO_4^{2\ominus}$) are equivalent, the pH is equal to the pK_a, because the log term is zero (log 1 = 0). Therefore, mixing 50 milliliters of solution A with 50 milliliters of solution B gives a buffer of pH 7.2. Since the concentration of each solution is 0.02 M, mixing equal volumes gives a buffer whose phosphate concentration is also 0.02 M. The reason why this is an effective buffer is that the final pH is at the pK_a value. This means that the buffer will resist changes in pH over a considerable range.

9. (a) The effective range of a buffer is from approximately one pH unit below to one pH unit above the pK_a. The buffering range for MOPS is therefore 6.2–8.2, and the buffering range for SHS is 4.5–6.5. Use the Henderson-Hasselbalch equation to calculate the ratios of basic to acidic species.

For MOPS: $\text{pH} = \text{p}K_a + \log\dfrac{[R_3N]}{[R_3NH^{\oplus}]}$

$$6.5 = 7.2 + \log\frac{[R_3N]}{[R_3NH^{\oplus}]}$$

$$\frac{[R_3N]}{[R_3NH^{\oplus}]} = \frac{1}{5}$$

For SHS:

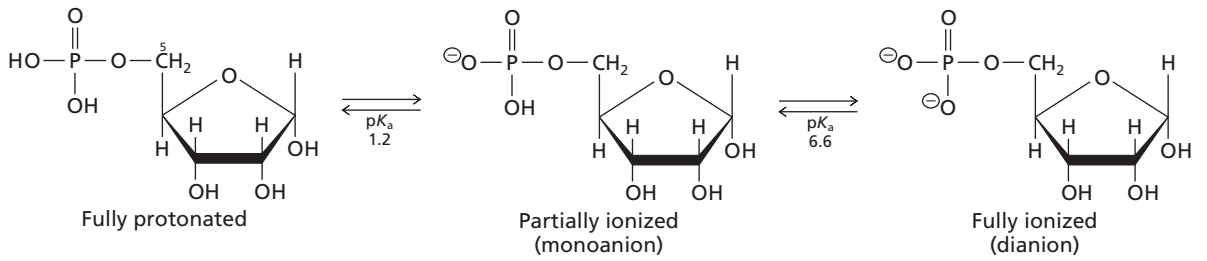
$$6.5 = 5.5 + \log\frac{[RCOO^{\ominus}]}{[RCOOH]}$$

$$\frac{[RCOO^{\ominus}]}{[RCOOH]} = \frac{10}{1}$$

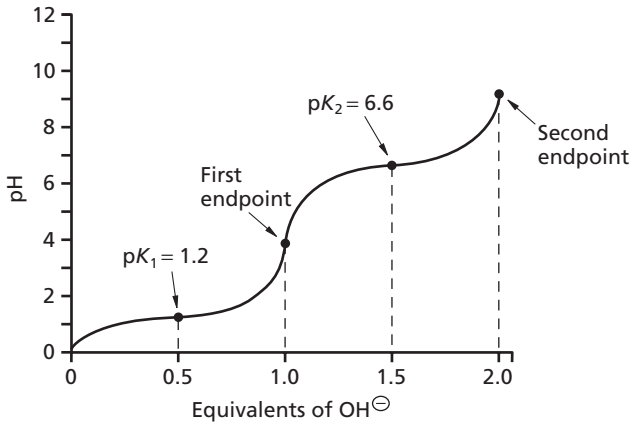
(b) An SHS buffer solution at pH 6.5 contains a much greater proportion of conjugate base relative to acid (10:1) than MOPS does (1:5). Therefore, an SHS buffer would more effectively maintain the pH upon addition of acid: $H^{\oplus} + RCOO^{\ominus} \rightleftharpoons RCOOH$. Conversely, a MOPS buffer at pH 6.5 contains a greater proportion of acid than SHS does; therefore, MOPS would more effectively maintain the pH upon addition of base: $R_3NH^{\oplus} + OH^{\ominus} \rightleftharpoons R_3N + H_2O$.

10.

(a)



(b)



11. Excess gaseous CO_2 rapidly equilibrates with aqueous CO_2 (Equation 2.25), leading to formation of carbonic acid (Equation 2.23). Carbonic acid ionizes to $H^{\oplus}$ and $HCO_3^{\ominus}$ (Equation 2.22). The excess acid, in the form of $H^{\oplus}$, can accumulate in bodily fluids, producing acidosis.

12. Although the metabolism of lactate and other organic acids in the diet can lead to production of CO_2 as shown, CO_2 is efficiently expired from the lungs (except during respiratory acidosis). Thus, the net product of the metabolic process is bicarbonate ($HCO_3^{\ominus}$), a base. Excess $H^{\oplus}$ present during metabolic acidosis can be removed when it combines with $HCO_3^{\ominus}$ to form H_2CO_3 (Equation 2.22), which then forms aqueous CO_2 and H_2O (Equation 2.23).

13. The acidic and conjugate base species of aspirin can be represented as RCOOH and $RCOO^{\ominus}$. Use the Henderson-Hasselbalch equation to calculate the ratio of the two species at pH 2.0 and pH 5.0. Then calculate the fraction of the total species that is unionized and available for absorption. In the stomach at pH 2.0,

$$\text{pH} = \text{p}K_a + \log\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]}$$

$$2.0 = 3.5 + \log\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]}$$

$$\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]} = \frac{0.03}{1}$$

The percentage of the uncharged species (RCOOH) is equal to the amount of RCOOH divided by the total of RCOOH and $RCOO^{\ominus}$, times 100%.

$$\frac{[\text{RCOOH}]}{[\text{RCOOH}] + [\text{RCOO}^{\ominus}]} \times 100\% = \frac{1}{1 + 0.03} \times 100\% = 97\%$$

Therefore, nearly all aspirin in the stomach is in a form available for absorption. In the upper intestine at pH 5.0, however, only a small percentage of aspirin is available for absorption.

$$5.0 = 3.5 + \log\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]}$$

$$\frac{[\text{RCOO}^{\ominus}]}{[\text{RCOOH}]} = \frac{32}{1}$$

$$\frac{[\text{RCOOH}]}{[\text{RCOOH}] + [\text{RCOO}^{\ominus}]} \times 100\% = \frac{1}{1 + 32} \times 100\% = 3\%$$

Note that aspirin must be in solution in order to be absorbed. For this reason, coated or slow-release forms of aspirin may alter the availability of aspirin in the stomach and intestine.

14. Use the Henderson-Hasselbach equation to calculate the ratio of the two species at each pH

At pH = 7.5

$$\text{pH} = \text{p}K_a + \log\frac{[\text{H}_2\text{NCH}_2\text{CONH}_2]}{[^{+}\text{H}_3\text{NCH}_2\text{CONH}_2]}$$

$$7.5 = 8.2 + \log\frac{[\text{H}_2\text{NCH}_2\text{CONH}_2]}{[^{+}\text{H}_3\text{NCH}_2\text{CONH}_2]}$$

$$\log\frac{[\text{H}_2\text{NCH}_2\text{CONH}_2]}{[^{+}\text{H}_3\text{NCH}_2\text{CONH}_2]} = 7.5 - 8.2 = -0.7$$

$$\frac{[\text{H}_2\text{NCH}_2\text{CONH}_2]}{[^{+}\text{H}_3\text{NCH}_2\text{CONH}_2]} = \frac{1}{5}$$

The ratio of $[H_2NCH_2CONH_2]$ to $[^{+}H_3NCH_2CONH_2]$ is 1 to 5. To determine the percent in the conjugate base form: 1/(1 + 5)*100 = 17%. Therefore, 17% is unprotonated at pH 7.5.

At pH = 8.2

$$\text{pH} = \text{p}K_a + \log\frac{[\text{H}_2\text{NCH}_2\text{CONH}_2]}{[^{+}\text{H}_3\text{NCH}_2\text{CONH}_2]}$$

$$8.2 = 8.2 + \log\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]}$$

$$\log\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]} = 8.2 - 8.2 = 0$$

$$\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]} = \frac{1}{1}$$

The ratio of $[H_2NCH_2CONH_2]$ to $[^+H_3NCH_2CONH_2]$ is 1.0 to 1.0. To determine the percent in the conjugate base form: 1/(1 + 1)*100 = 50%. Therefore, 50% is unprotonated at pH 8.2.

At pH 9.0:

$$pH = pK_a + \log\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]}$$

$$9.0 = 8.2 + \log\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]}$$

$$\log\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]} = 9.0 - 8.2 = 0.8$$

$$\frac{[H_2NCH_2CONH_2]}{[^+H_3NCH_2CONH_2]} = \frac{6.3}{1}$$

The ratio of $[H_2NCH_2CONH_2]$ to $[^+H_3NCH_2CONH_2]$ is 6.3 to 1. To determine the percent of the conjugate base: 6.3/(6.3 + 1)*100 = 86%. That is, 86% is unprotonated at pH 9.0.

15. This titration curve represents a compound with two pK_a values, shown by the two plateaus (near pH 2 and pH 10). Glycine has two pK_a values at 2.4 and at 9.8.

16. Only (a) vitamin C would be soluble in water. Vitamin C contains several hydroxyl groups, each of which can hydrogen-bond with water.

17. At 0°C the ion product for water is 1.14×10^{-15}. At neutral pH,

$$[H^{\oplus}] = [OH^{\ominus}] = \sqrt{1.14 \times 10^{-15}} = 3.38 \times 10^{-8}$$

$$pH = -\log(3.38 \times 10^{-8}) = 7.47$$

At 100°C

$$[H^{\oplus}] = [OH^{\ominus}] = \sqrt{4.0 \times 10^{-13}} = 6.32 \times 10^{-7}$$

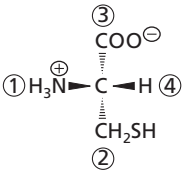
$$pH = -\log(6.32 \times 10^{-7}) = 6.2$$

Note that the density of water changes with temperature but this has very little effect on $[H^+]$.

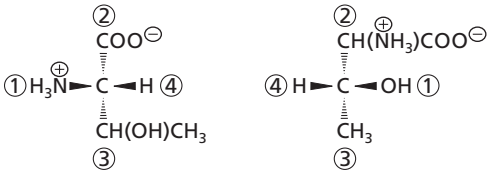
18. HCl dissociates completely in water. In 6 M HCl, $[H^{\oplus}] = 6$ M. The pH is $-\log(6) = -0.78$. The standard pH scale begins at zero ($[H^{\oplus}] = 1$ M) because it's very unusual to encounter more acidic solutions in biology.

Chapter 3 Amino Acids and the Primary Structures of Proteins

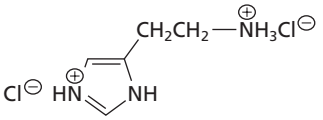
1. By comparing the priorities of L-cysteine (shown here) to those of L-serine (*S* configuration, page 57) you will find that their sequence is clockwise and therefore L-cysteine has the *R* configuration.



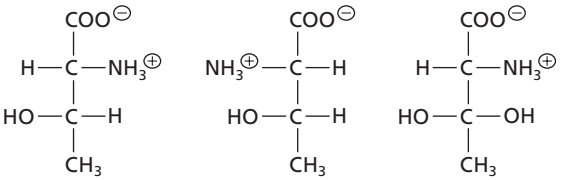
2. The stereochemistry of each chiral carbon must be examined to determine whether it has the *R* or *S* configuration.



C-2, *S*-configuration C-3, *R*-configuration



3. The other stereoisomers are:

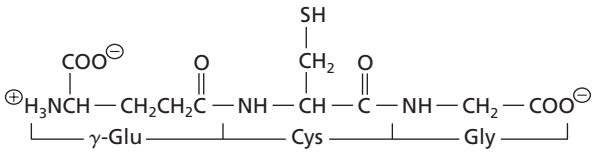


D-Threonine L-Allothreonine D-Allothreonine

4. Methionine.

5. (a) Serine; phosphorylation of the hydroxyl group.
(b) Glutamate; carboxylation of the γ-carbon.
(c) Lysine; acetylation of the ε-amino group.

6. By convention, peptides are designated from the N-terminus → C-terminus, therefore Glu is the N-terminus and Gly is the C-terminus.

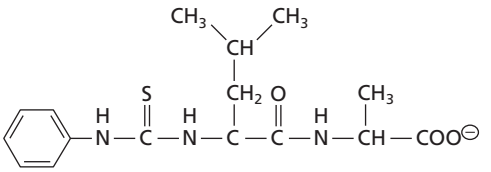


7. The 6 residues at the C-terminus of melittin are highly hydrophilic (Table 3.1). Of the remaining 20 amino acid residues, nearly all are hydrophobic, including 9 with highly hydrophobic side chains (leucine, isoleucine, valine). The hydrophilic portion of melittin is more soluble in aqueous solution, while the hydrophobic portion is more soluble in the membrane lipids.

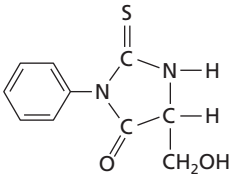
8. Use Table 3.2 to determine the net charge at each pK_a value. The pH at which the net charge is 0 lies midway between the two pK_a values at which the average charges are +0.5 and −0.5.
(a) At pH 9.0, the net charge of arginine is +0.5, and at pH 12.5, the net charge is −0.5. Therefore, $pI_{Arg} = (9.0 + 12.5) \div 2 = 10.8$.
(b) At pH 2.1, the net charge of glutamate is +0.5, and at pH 4.1, the net charge is −0.5. Therefore, $pI_{Glu} = (2.1 + 4.1) \div 2 = 3.1$.

9. The ionizable groups are the free amino group of the N-terminal cysteine residue (pK_a = 10.7), the glutamate side chain (pK_a = 4.1), and the histidine side chain (pK_a = 6.0).
(a) At pH 2.0, the N-terminus and the histidine side chain have positive charges and the glutamate side chain is uncharged. The net charge is +2.
(b) At pH 8.5, the N-terminus has a positive charge, the histidine side chain is uncharged, and the glutamate side chain has a negative charge. The net charge is 0.
(c) At pH 10.7, the charge of the N-terminus is +0.5, the histidine side chain is uncharged, and the glutamate side chain has a negative charge. The net charge is −0.5.

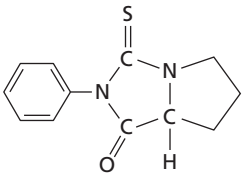
10. (a)



(b)

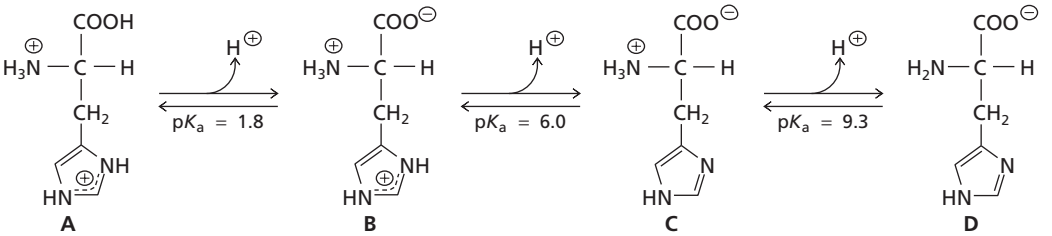


(c)



11. (a) Gly–Ala–Trp–Arg, Asp–Ala–Lys, Glu–Phe–Gly–Gln
(b) Gly–Ala–Trp, Arg–Asp–Ala–Lys–Glu–Phe, Gly–Gln
(c) Gly–Ala–Trp–Arg–Asp, Ala–Lys–Glu, Phe–Gly–Gln

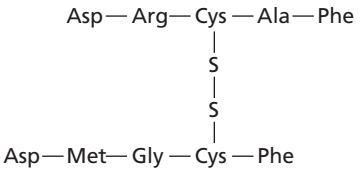
12. (a)



(b) A, 1; B, 3; C, 5; D, 7
(c) 1, 4, 5, 7
(d) 4
(e) 5
(f) Histidine would be a good buffer within one pH unit of any of its three pK_a values: 0.8–2.8, 5.0–7.0, and 8.3–10.3.

13. (a) Because there are two N-terminal groups, there must be two peptide chains, each having an N-terminal aspartate residue.
(b) 2-Mercaptoethanol reduces disulfide bonds, and trypsin catalyzes cleavage on the carboxyl side of arginine residues. Since aspartate is found at both N-termini of FP, the sequence of the dipeptide is Asp–Arg, and the sequence of the pentapeptide is Asp–(Cys, Gly, Met, Phe). The tripeptide has the sequence Cys–(Ala, Phe) and is derived from trypsin–catalyzed cleavage of a pentapeptide whose sequence is Asp–Arg–Cys–(Ala, Phe).
(c) The C-terminal residue of each peptide chain is phenylalanine. Now that the terminal residues are known, one peptide must have the sequence Asp–(Cys, Gly, Met)–Phe, and the other must have the sequence Asp–Arg–Cys–Ala–Phe.
(d) CNBr cleaves on the carbonyl side of methionine residues to produce C-terminal homoserine lactone residues. The peptides are therefore Asp–Met and (Cys, Gly)–Phe. Glycine is the N-terminal residue of the tripeptide, so that pentapeptide sequence is Asp–Met–Gly–Cys–Phe.

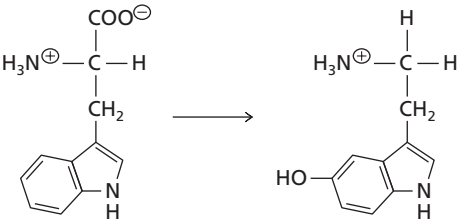
The complete FP structure is



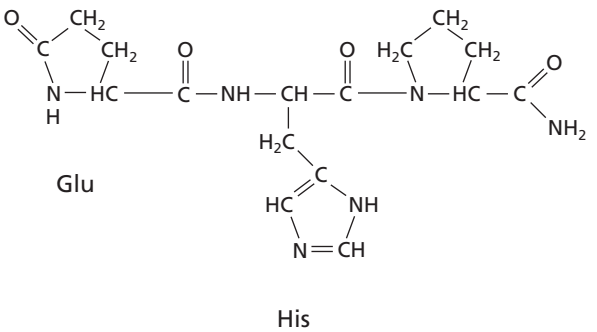
14. (a) The substitution of aspartate (D) for glutamate (E) at position 50 is an example of a conservative change. The amino acids aspartate and glutamate both contain acidic side chains that are negatively charged at physiological pH.
(b) The substitution of tyrosine (Y) for histidine (H) is an example of a nonconservative substitution since tyrosine contains an aromatic side chain and histidine contains a hydrophilic side chain consisting of an imidazole group.

15. The neurotransmitter serotonin is derived from the amino acid tryptophan.

In the conversion, the carboxyl group from tryptophan is removed and a hydroxyl group is added to the aromatic ring.



16. (a) There are two peptide bonds present in TRH. They are marked with the dashed lines.
(b) TRH is derived from the tripeptide Glu–His–Pro. The proline carboxyl group has been modified to an amide (marked with an *). The side chain carboxyl group of the amino terminal Glu forms an amide with the residue's α-amino group (marked with a **).
(c) The amino- and carboxyl-terminal groups have been modified to amide groups and thus are uncharged.



17. (a) L-Dopa is in the *S* configuration.
(b) They are both derived from the amino acid tyrosine.

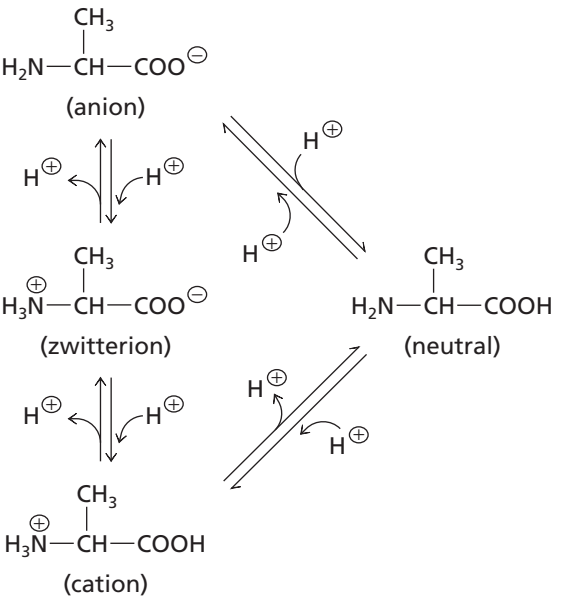
18. Although Figure 3.6 shows only three forms of alanine, there are actually four different forms in equilibrium (see next page). The neutral form will be present at very low concentrations because at any given pH the three other forms are much more stable. We can calculate the relative ratios of the four forms by assuming that the protonation/deprotonation of the two charged groups is independent.

For alanine at pH 2.4 the relative ratio of $R—COO^{\ominus}$ and R—COOH is

$$2.4 = 2.4 + \log\frac{[R—COO^{\ominus}]}{[R—COOH]} \text{ therefore } \frac{[R—COO^{\ominus}]}{[R—COOH]} = 1$$

and the ratio of $H_3N^{\oplus}—R$ to $H_2N—R$ is

$$2.4 = 9.9 + \log\frac{[H_2N—R]}{[H_3N^{\oplus}—R]} \text{ therefore } \frac{[H_2N—R]}{[H_3N^{\oplus}—R]} = 3.1 \times 10^{-8}$$



Thus the relative ratios of the four forms are approximately

$$\text{cation : zwitterion : anion : neutral} \quad 1 : 1 : 10^{-8} : 10^{-8}$$

and the concentration of the neutral form in a 0.01 M solution of alanine is about 10^{-10} M. Neutral molecules exist but their concentration is insignificant.

At pH 9.9 the ratios are

$$\text{anion : zwitterion : cation : neutral} \quad 1 : 1 : 10^{-8} : 10^{-8}$$

At pH 6.15 the relative ratio of $R—COO^{\ominus}$ and R—COOH is

$$6.15 = 2.4 + \log \frac{[R—COO^{\ominus}]}{[R—COOH]} \text{ therefore } \frac{[R—COO^{\ominus}]}{[R—COOH]} = 5.6 \times 10^3$$

and the relative ratio of $H_2N—R$ and $[H_3N^{\oplus}—R]$ is

$$6.15 = 9.9 + \log \frac{[H_2N—R]}{[H_3N^{\oplus}—R]} \qquad \frac{[H_2N—R]}{[H_3N^{\oplus}—R]} = 1.8 \times 10^{-4} \qquad \frac{[H_3N^{\oplus}—R]}{[H_2N—R]} = 5.6 \times 10^{-8}$$

The zwitterion is present in 5600-fold excess over the anion and cation forms and each of these forms is 5600-fold more likely than the neutral form. The ratios are

$$\text{zwitterion : anion : cation : neutral} \quad 3.1 \times 10^7 : 1 : 1.8 \times 10^{-4} : 1.8 \times 10^{-4} : 3.2 \times 10^{-8}$$

The concentration of the neutral form in a solution of 0.01 M alanine is insignificant.

19. The relative concentrations of the zwitterion and the cation are

$$2.4 = 2.4 + \log \frac{[H_3N^{\oplus}—\overset{CH_3}{\overset{|}{C}}—COO^{\ominus}]}{[H_3N^{\oplus}—\underset{CH_3}{\overset{|}{C}}—COOH]} \qquad \frac{[H_3N^{\oplus}—\overset{CH_3}{\overset{|}{C}}—COO^{\ominus}]}{[H_3N^{\oplus}—\underset{CH_3}{\overset{|}{C}}—COOH]} = 1$$

Thus the concentration of the zwitterion in a solution of 0.01 M alanine is 0.005 M. (We can ignore the concentrations of the anion and neutral forms—see previous question.)

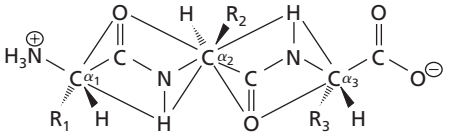
At pH 4.0

$$4.0 = 2.4 + \log \frac{[\text{zwitterion}]}{[\text{cation}]} \qquad \frac{[\text{zwitterion}]}{[\text{cation}]} = 40$$

The concentration of the zwitterion is 0.01 M $\times \frac{40}{41} = 0.00976$ M

Chapter 4 Proteins: Three-Dimensional Structure and Function

1. (a)

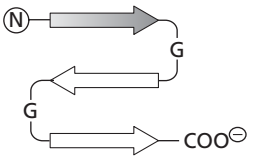


(b) The R groups represent the side chains of the amino acid residues.
(c) The partial double-bond character of the C—N amide bonds prevents free rotation.
(d) Both peptide groups in this tripeptide are in the *trans* conformation, since the α-carbon atoms are on opposite sides of the peptide bonds.
(e) The peptide groups may rotate around the N—C_α and C_α—C bonds.

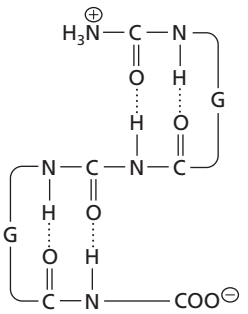
2. (a) (1) In an α helix, intrachain hydrogen bonds form between carbonyl oxygens of certain residues and amide hydrogens of other residues. The hydrogen bonds are approximately parallel to the helix axis (Figure 4.10).
(2) In a collagen triple helix, interchain hydrogen bonds form between amide hydrogens of the glycines in one chain and carbonyl oxygen atoms of residues (which are often proline) in an adjacent chain (Figure 4.41). There are no intrachain hydrogen bonds in a collagen helix.
(b) The side chains of an α helix point outward from the cylinder of the helix (Figure 4.11). In collagen, three chains coil around each other so that every third residue of a given chain makes contact with the other two chains along the central axis of the triple helix (Figure 4.42). Only the small side chain of glycine can fit at these positions. The other side chains point outward from the triple helical coil.

3. (1) The presence of glycine in an α helix destabilizes the helix due to the greater freedom of movement allowed by the small side chain. For this reason, many α helices begin or end with glycine.
(2) Proline tends to disrupt α helices because its rigid, cyclic side chain stereochemically interferes with the space that would normally be occupied by a neighboring residue in the α helix. In addition, proline lacks a hydrogen on its amide nitrogen and cannot participate in normal intrahelical hydrogen bonding.

4. (a) Due to the flexibility resulting from a small side chain (—H), glycine is often found in "hairpin loops" that connect sequential antiparallel β strands. The glycine residues (G) in positions 8 and 14 provide two hairpin-loop regions to connect the three β strands in Betanova.



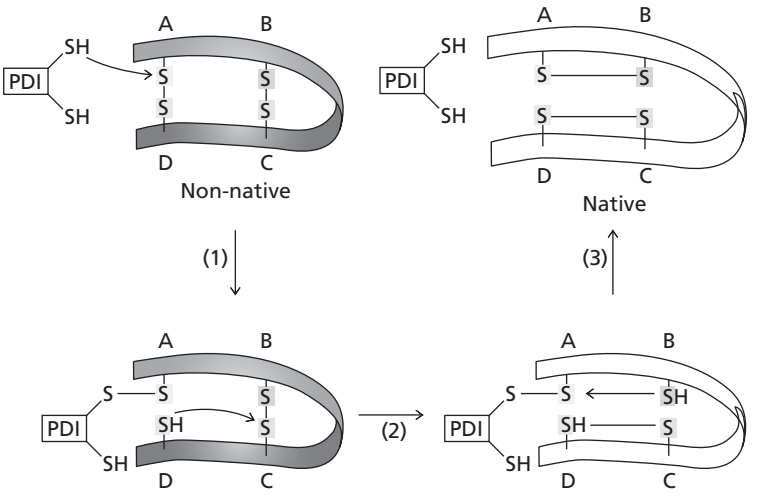
(b) β-sheet structures are stabilized by hydrogen bonds that form between a carbonyl oxygen of one strand and an amide nitrogen of an adjacent strand (Figure 4.15).



5. Helix-loop-helix (HLH) motif (Figure 4.19).

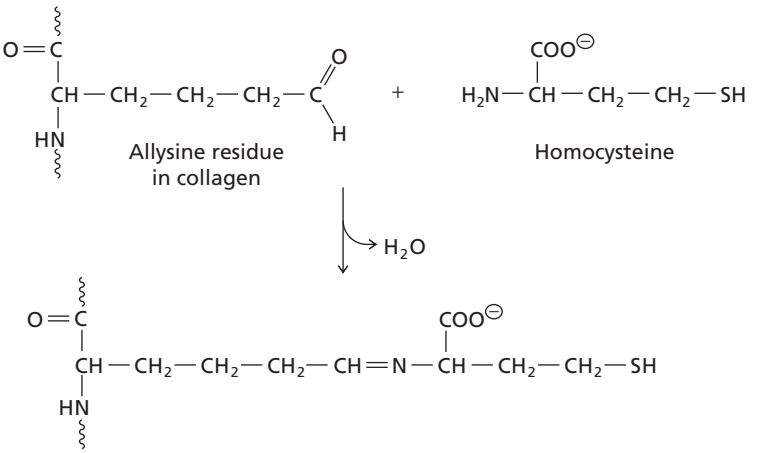
6. (a) α/β. Regions of α helix and β strand alternate in the polypeptide chain.
(b) α/β barrel. Parallel β strands are surrounded by a layer of α helices in a cylindrical shape.
(c) Yeast FMN oxidoreductase and *E. coli* enzyme required for tryptophan biosynthesis (Figure 4.24 (i) and (j) respectively).

7. Protein disulfide isomerase contains two reduced cysteine residues at the active site, and these participate in a reduction and disulfide exchange that allows the misfolded protein to refold into the lower energy native conformation.



8. The highly hydrophobic side chains of methionine, leucine, phenylalanine, and isoleucine are most likely to be on the side of the helix that faces the interior of the protein. Most of the other side chains are polar or charged and can interact with the aqueous solvent. Since the α helix is a repeating structure with approximately 3.6 residues per turn, the hydrophobic groups must be found every three or four residues along the sequence, so that one side of the helix is hydrophobic.

9. Covalent cross-linking contributes significantly to the strength and rigidity of collagen fibers. In one type of cross-link, allysine residues in a collagen molecule condense with lysine residues in an adjacent molecule, forming Schiff bases (Figure 4.38a). When an allysine residue reacts with homocysteine, it is unable to participate in the normal cross-linking of collagen molecules. High levels of homocysteine in blood probably lead to defective collagen structure and skeletal deformities.



10. The sequence –Gly–Pro–X–Y– occurs frequently in collagen, which is found throughout the body, including in the skin. Because the larval enzyme can catalyze cleavage of collagen chains, the parasite is able to enter the host.

11. The reaction of carbon dioxide with water explains why there is a concomitant lowering of pH when the concentration of CO_2 increases. Carbon dioxide produced by rapidly metabolizing tissue reacts with water to produce bicarbonate ions and $H^{\oplus}$.

(a) $CO_2 + H_2O \longleftrightarrow H_2CO_3 \longleftrightarrow HCO_3^{\ominus} + H^{\oplus}$

The $H^{\oplus}$ generated in this reaction decreases the pH of the blood and thus stabilizes the deoxy form (T conformation) of hemoglobin. The net effect is an increase in the P_{50}, that is, a lower affinity of hemoglobin for oxygen, so that more oxygen is released to the tissue (Figure 4.50). Carbon dioxide also lowers the affinity of hemoglobin for oxygen by forming

carbamate adducts with the N-termini of the four chains (Figure 4.51). These adducts contribute to the stability of the deoxy (T) conformation, thereby further increasing the P_{50} and promoting the release of oxygen to the tissue.

(b) Shock victims suffer a critical deficit of oxygen supply to their tissues. Bicarbonate administered intravenously provides a source of carbon dioxide to the tissues. By lowering the affinity of hemoglobin for oxygen, carbon dioxide facilitates a release of oxygen from oxyhemoglobin to the tissues.

12. (a) 2,3BPG binds to positively charged side chains in the central cavity of deoxyhemoglobin (Figure 4.49). Since Hb F lacks two positively charged groups (His-143 of each β chain), 2,3BPG binds less tightly to Hb F than to Hb A.

(b) 2,3BPG stabilizes the deoxy form of hemoglobin, increasing the fraction of molecules in the deoxy form. Since Hb F binds 2,3BPG less tightly than does Hb A, Hb F is less affected by 2,3BPG in the blood and has a greater fraction of molecules in the oxy form. Hb F therefore has a greater affinity than Hb A for oxygen at any oxygen pressure.

(c) At the oxygen pressure of tissues, 20–40 torr, Hb F has a greater affinity for oxygen than does Hb A. The difference in affinity allows efficient transfer of oxygen from maternal blood to the fetus.

13. The low P_{50} value of Hb_{Yakima} indicates a greater than normal affinity for oxygen even at the oxygen pressures found in working muscle. The increased affinity means that Hb_{Yakima} gives up less oxygen to the working muscle.

14. (a) Hydrophilic (italicized) and hydrophobic (underlined) residues are identified:
*E*CG*K*<u>F</u><u>M</u><u>W</u>*K*C*K*N*S*N*D*C*C*<u>*K*</u>*D*<u>L</u><u>V</u>C*S*S*R*<u>W</u>*K*<u>W</u><u>C</u><u>V</u><u>L</u>*A*S<u>P</u><u>F</u>

(b) In the three-dimensional structure of proteins, amino acids that are far from each other in the primary sequence can interact in the globular structure of the protein. Thus the hydrophobic amino acids can be very close to each other in the three-dimensional structure and provide a "hydrophobic" face for interaction with the membrane.

15. (a) The most effective binding of selenoprotein P to heparin is seen at a pH below 6. The binding of selenoprotein P to heparin decreases as the pH is increased to 7. There is very little binding of selenoprotein P to heparin at pH values greater than 7.

(b) Heparin is negatively charged. If selenoprotein P is positively charged, it can bind to heparin. Histidine residues are abundant in selenoprotein P. Histidine has an imidazole side chain that has a pK_a value of 6.0. That is, at a pH of 6.0, 50% of the histidine residues would be protonated and positively charged and 50% would be unprotonated and uncharged. Below a pH of 6.0, there would be a net positive charge on the histidine residues, resulting in effective electrostatic interactions with the heparin. At pH values above 7, almost all of the histidine residues would be unprotonated and uncharged and will not effectively interact with the negatively charged heparin molecule.

16. Collagen is protein consisting of three polypeptide chains that are wound together in a triple helix. The protease bromelin is an enzyme that cleaves some of the peptide bonds in the polypeptide chains. The polypeptide chains are necessary to trap the water molecules in a semisolid state when gelatin cools, and if these are cleaved, the gelatin will not set properly. The cleavage of the polypeptide chains in collagen by bromelin destroys the ability of the gelatin to harden. If the pineapple is first cooked, the heat will denature the protein and thus the enzyme activity will be destroyed. Therefore cooked pineapple can be added to slightly thickened gelatin, and the gelatin will proceed to the semisolid state as desired. (Assume that heat denaturation is irreversible.)

17. The replacement of lysine by methionine results in one less positive charge on each beta subunit in the central cavity (see Figure 4.49). 2,3BPG binds less tightly to HbH. This causes more of the mutant protein to be in the R state (oxyhemoglobin is stabilized). The curve is shifted towards the left (more like myoglobin). Since more is in the R state, the affinity for oxygen has increased.

Chapter 5 Properties of Enzymes

1. The initial velocities are approaching a constant value at the higher substrate concentrations, so we can estimate the V_{max} as 70 mM/min. Since K_m equals the concentration of substrate [S] required to reach half the maximum velocity, we can estimate the K_m to be 0.01 M since that's the concentration of substrate that yields a rate of 35 mM/min ($= V_{max}>2$).

2. (a) The ratio k_{cat}/K_m, or *specificity constant*, is a measure of the preference of an enzyme for different substrates. When two substrates at the same concentration compete for the

active site of an enzyme, the ratio of their rates of conversion to product is equal to the ratio of the k_{cat}/K_m values, since $v_0 = (k_{cat}/K_m)[E][S]$ for each substrate and [E] and [S] are the same.

$$\frac{v_0(S_1)}{v_0(S_2)} = \frac{(k_{cat}>K_m)^1>[E][S]}{(k_{cat}>K_m)^2>[E][S]}$$

(b) The upper limit of k_{cat}/K_m approaches 10^8 to 10^9 s^{-1}, the fastest rate at which two uncharged molecules can approach each other by diffusion at physiological temperatures.

(c) The catalytic efficiency of an enzyme cannot exceed the rate for the formation of ES from E and S. The most efficient enzymes have k_{cat}/K_m values approaching the rate at which they encounter a substrate. At this limiting velocity they have become as efficient catalysts as possible because every encounter produces a reaction. (Most enzymes don't need to catalyze reactions at the maximum possible rates so there's no selective pressure to evolve catalytically perfect enzymes.)

3. The catalytic constant (k_{cat}) is the first-order rate constant for the conversion of ES to E + P under saturating substrate concentrations (Equation 5.26), and CA has a much higher catalytic activity in converting substrate to product than does OMPD. However, the efficiency of an enzyme can also be measured by the *rate acceleration* provided by the enzyme over the corresponding uncatalyzed reaction (k_{cat}/k_n, Table 5.2). The reaction of the substrate for OMPD in the absence of enzyme is very slow ($k_n = 3 \times 10^{-16}$ s^{-1}) compared to the reaction for the CA substrate in the absence of enzyme ($k_n = 1 \times 10^{-1}$ s^{-1}). Therefore, while the OMPD reaction is much slower than the CA reaction in terms of k_{cat}, OMPD is one of the most efficient enzymes known and provides a much higher rate acceleration than does CA when the reactions of each enzyme are compared to the corresponding uncatalyzed reactions.

4. When [S] = 100 μM, [S] $\gg K_m$, so $v_0 = V_{max} = 0.1$ μM min^{-1}.

(a) For any substrate concentration greater than 100 μM, $v_0 = V_{max} = 0.1$ μM min^{-1}.

(b) When [S] = K_m, $v_0 = V_{max}/2$, or 0.05 μM min^{-1}.

(c) Since K_m and V_{max} are known, the Michaelis-Menten equation can be used to calculate v_0 at any substrate concentration. For [S] = 2 μM,

$$v_0 = \frac{V_{max}[S]}{K_m + [S]} = \frac{(0.1\ \mu M\ min^{-1})(2\ \mu M)}{(1\ \mu M + 2\ \mu M)} = \frac{0.2}{3}\mu M\ min^{-1} = 0.067\ \mu M\ min^{-1}$$

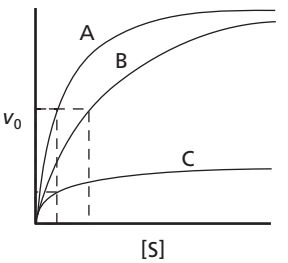
5. (a) Determine $[E]_{total}$ in moles per liter, then calculate V_{max}.

$$[E]_{total} = 0.2\ g\ l^{-1}\left(\frac{1\ mol}{21\ 500\ g}\right) = 9.3 \times 10^{-6}\ M$$

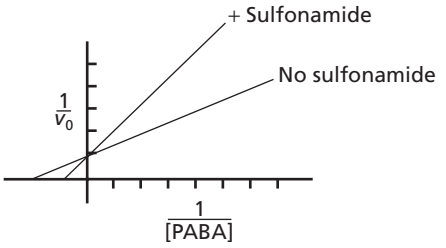
$$V_{max} = k_{cat}[E]_{total} = 1000\ s^{-1}(9.3 \times 10^{-6}\ M) = 9.3 \times 10^{-3}\ M\ s^{-1}$$

(b) Since V_{max} is unchanged in the presence of the inhibitor, competitive inhibition is occurring. Because the inhibitor closely resembles the heptapeptide substrate, competitive inhibition by binding to the enzyme active site is expected (i.e., classical competitive inhibition).

6. Curve A represents the reaction in the absence of inhibitors. In the presence of a competitive inhibitor (curve B), K_m increases and V_{max} is unchanged. In the presence of a noncompetitive inhibitor (curve C), V_{max} decreases and K_m is unchanged.

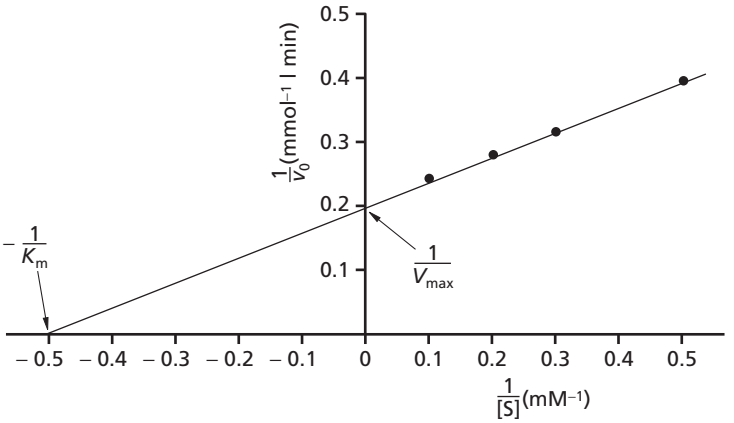


7. Since the inhibitor sulfonamides structurally resemble the PABA substrate we would predict that sulfonamides bind to the enzyme active site in place of PABA and act as competitive inhibitors (Figure 5.9).



8. (a) To plot the kinetic data for fumarase, first calculate the reciprocals of substrate concentrations and initial rates of product formation. (Note the importance of including correct units in calculating and plotting the data.)

Fumarate [S] (mM)	$\frac{1}{[S]}$ (mM^{-1})	Rate of product formation v_0 (mmol l^{-1} min^{-1})	$\frac{1}{v_0}$ (mmol^{-1} l min)
2.0	0.50	2.5	0.40
3.3	0.30	3.1	0.32
5.0	0.20	3.6	0.28
10.0	0.10	4.2	0.24



V_{max} is obtained by taking the reciprocal of $1/V_{max}$ from the *y* intercept (Figure 5.6).

$$1/V_{max} = 0.20 \text{ mmol}^{-1} \text{ l min, so } V_{max} = 5.0 \text{ mmol l}^{-1} \text{ min}^{-1}$$

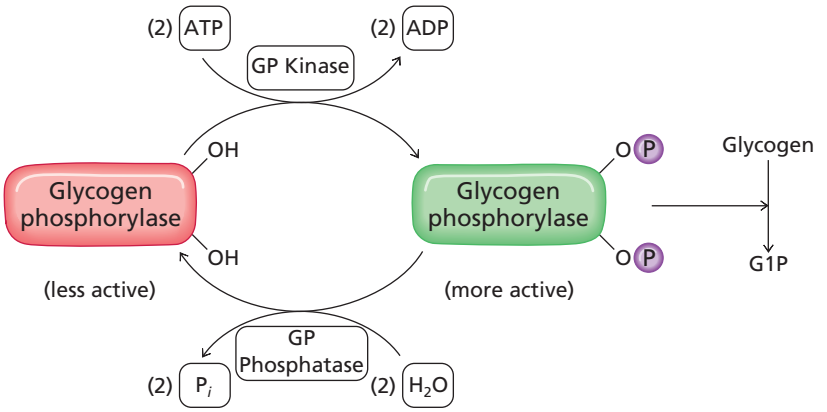
K_m is obtained by taking the reciprocal of $-1/K_m$ from the *x* intercept.

$$-1/K_m = -0.5 \text{ mM}^{-1}, \text{ so } K_m = 2.0 \text{ mM or } 2 \times 10^{-3} \text{ M}$$

(b) The value of k_{cat} represents the number of reactions per second that *one enzyme active site* can catalyze. Although the concentration of enzyme is 1×10^{-8} M, fumarase is a tetramer with four active sites per molecule so the total concentration of enzyme active sites [E_{total}] is 4×10^{-8} M. Using Equation 5.26:

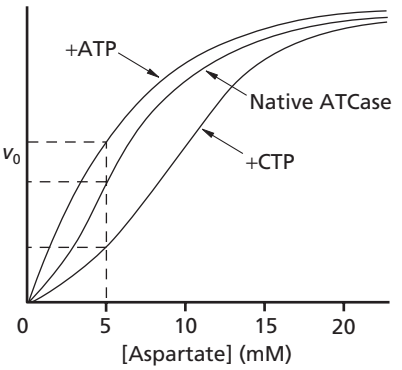
$$k_{cat} = \frac{V_{max}}{[E_{total}]} = \frac{5.0 \text{ mmol l}^{-1} \text{ min}^{-1}}{4 \times 10^{-5} \text{ mmol l}^{-1}} \times \frac{1 \text{ min}}{60 \text{ s}} = 2 \times 10^3 \text{ s}^{-1}$$

9. Like pyruvate dehydrogenase (PDH) (Figure 5.22), glycogen phosphorylase (GP) activity is regulated by alternate phosphorylation by a *kinase* and dephosphorylation by a *phosphatase*. However, unlike PDH, the active form of GP has two phosphorylated serine residues; in the inactive GP form, two serine residues are not phosphorylated.



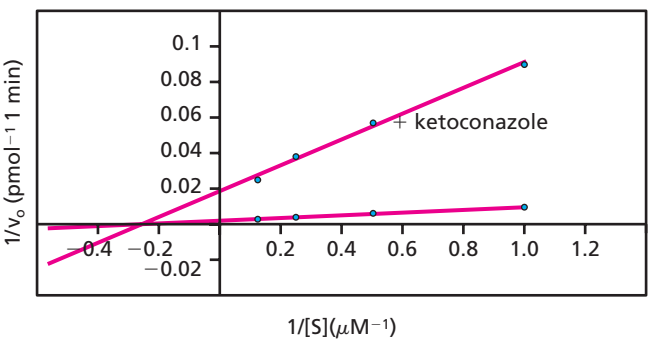
10. Inhibition of the first committed step of a multistep pathway allows the pathway to proceed only when the end product is needed. Since the first committed step is regulated, flux in the pathway is controlled. This type of regulation conserves raw material and energy.

11. When [aspartate] = 5 mM, $v_0 = V_{max}/2$. Therefore, in the absence of allosteric modulators, K_m = [S] = 5 mM. ATP increases v_0, and CTP decreases v_0.



12. (a) To plot the kinetic data for P450 3A4, first calculate the reciprocals of substrate concentrations and initial rates of product formation. The data are plotted in the double reciprocal plot and are shown with the dashed line.

Midazolam [S](μM)	1/[S] (μM^{-1})	Rate of product formation v_0 (pmol l^{-1} min^{-1})	$1/v_0$ ($pmol^{-1}$ l min)
1	1	100	0.01
2	0.5	156	0.0064
4	0.25	222	0.0045
8	0.125	323	0.0031



V_{max} is obtained by taking the reciprocal of $1/V_{max}$ from the y intercept (Figure 5.6).

$$1/V_{max} - 0.0025 \text{ pmol}^{-1} \text{ l min, so } V_{max} = 400 \text{ pmol l}^{-1} \text{ min}^{-1}$$

K_m is obtained by taking the reciprocal of $-1/K_m$ from the x intercept

$$-1/K_m = -0.3\ \mu M^{-1}, \text{ so } K_m = 3.3\ \mu M$$

(b) The reciprocals of the substrate concentration and activity in the presence of ketoconazole are given in the table.

Midazolam [S] (μM)	1/[S] (μM^{-1})	Rate of product formation in the presence of 0.1 μM ketoconazole/ v_0 (pmol l^{-1} min^{-1})	$1/v_0$ (pmol^{-1} l min)
1	1	11	0.091
2	0.5	18	0.056
4	0.25	27	0.037
8	0.125	40	0.025

The plot of the data (solid line) is given in the double reciprocal plot shown in (a). There is an increase in the y intercept and no apparent change in the x intercept. From the double reciprocal plot, it appears that ketoconazole is a noncompetitive inhibitor (see Figure 5.11). These inhibitors are characterized by an apparent decrease in V_{max} (increase in $1/V_{max}$) with no change in K_m.

13. (a) Bergamottin appears to inhibit the activity of P450 3A4 since the P450 activity measured in the presence of 0.1 and 5 μM bergamottin is less than that of the P450 activity in the absence of bergamottin.

(b) It might be dangerous for a patient to take their medication with grapefruit juice since there appears to be an inhibition of P450 activity in the presence of bergamottin. If the bergamottin decreases the P450 activity, and the P450 enzyme is known to metabolize the drug to an inactive form, the time it takes to convert the drug to its inactive form may be increased. This may prolong the effects of the drug, which may lead to adverse consequences for the patient.

14. (a) When $[S] \gg K_m$, then $K_m + [S] \approx [S]$. Substrate concentration has no effect on velocity, and $v_0 = V_{max}$, as shown in the upper part of the curve in Figure 5.4a.

$$v_0 = \frac{V_{max}[S]}{K_m + [S]} \approx \frac{V_{max}[S]}{[S]} = V_{max}$$

(b) When $[S] \ll K_m$, $K_m + [S] \approx K_m$, and the Michaelis-Menten equation simplifies to

$$v_0 = \frac{V_{max}[S]}{K_m + [S]} \approx \frac{V_{max}[S]}{K_m}$$

Velocity is related to [S] by a constant value, and the reaction is first order with respect to S, as shown in the lower part of the curve in Figure 5.4a.

(c) When $v_0 = V_{max}/2$, $K_m = [S]$.

$$v_0 = \frac{V_{max}}{2} = \frac{V_{max}[S]}{K_m + [S]}$$

$$K_m + [S] = 2[S]$$

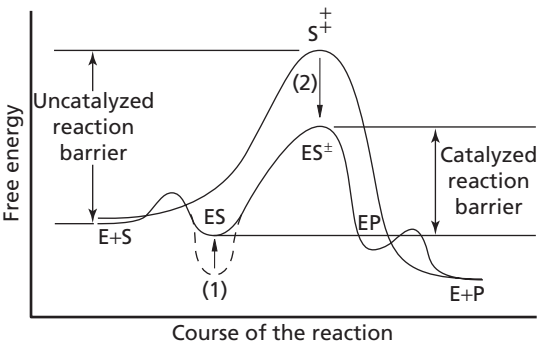
$$K_m = [S]$$

Chapter 6 Mechanisms of Enzymes

1. (a) The major binding forces in ES complexes include charge–charge interactions, hydrogen bonds, hydrophobic interactions, and van der Waals forces. (About 20% of enzymes bind a substrate molecule or part of it covalently.)

(b) Tight binding of a substrate would produce an ES complex that lies in a thermodynamic pit, effectively increasing the activation energy and thereby slowing down the reaction. Tight binding of the transition state, however, lowers the energy of the $ES^{\ddagger}$ complex, thereby decreasing the activation energy and increasing the rate of the reaction.

2. The activation barrier for the reaction is lowered by (1) raising the ground-state energy level (ES) and (2) lowering the transition-state energy level ($ES^{\ddagger}$), resulting in a reaction rate increase.



3. The rate determining step of a multistep reaction is the slowest step, which is the step with the highest activation energy. For Reaction 1, Step 2 is the rate determining step. For Reaction 2, Step 1 is the rate determining step.

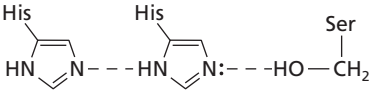
4. The reactive groups in Reaction 2 (—OH and —COOH) are held at close *proximity*. They are oriented in a manner suitable for catalysis by steric crowding of the bulky methyl groups of the ring. The reactive —COOH group cannot rotate away as freely as it can in Reaction 1. Model systems such as these are relevant because they indicate potential rate increases that might be obtained by enzymes that bring substrates and the enzyme's catalytic groups into positions that are optimal for reaction.

5. (1) Binding effects. Lysozyme binds the substrate so that the glycosidic bond to be cleaved is very close to both of the enzyme catalytic groups (Glu-35 and Asp-52). In addition, the energy of the ground-state sugar ring is raised because it is distorted into a half-chair conformation.
(2) Acid–base catalysis. Glu-35 first donates a proton to an oxygen of the leaving sugar (general acid catalysis), and then accepts a proton from the attacking water molecule (general base catalysis).
(3) Transition-state stabilization. Asp-52 stabilizes the developing positive charge on the oxocarbocation intermediate, and subsite D favors the half-chair sugar conformation of this intermediate. The structure proposed for the transition state includes both this charge and sugar conformation in addition to hydrogen bonding to several active-site residues.

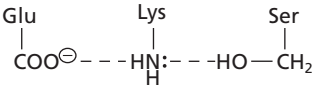
6. Serine 195 is the only serine residue in the enzyme that participates in the catalytic triad at the active site of α-chymotrypsin. The resulting increase in the nucleophilic character of Ser-195 oxygen allows it to react rapidly with DFP.

7. (a) The catalytic triad is composed of an aspartate, a histidine, and a serine residue. Histidine acts as a general acid–base catalyst, removing a proton from serine to make serine a more powerful nucleophile in the initial step. Aspartate forms a low-barrier hydrogen bond with histidine, stabilizing the transition state. An acid catalyst, histidine donates a proton to generate the leaving amine group.
(b) The oxyanion hole contains backbone —NH— groups that form hydrogen bonds with the negatively charged oxygen of the tetrahedral intermediate. The oxyanion hole mediates transition-state stabilization since it binds the transition state more tightly than it binds the substrate.
(c) During catalysis, aspartate forms a low-barrier hydrogen bond with the imidazolium form of histidine. Because asparagine lacks a carboxylate group to form the stabilizing hydrogen bond with histidine, enzyme activity is dramatically decreased.

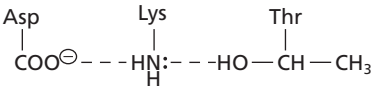
8. (a) Human cytomegalovirus protease: His, His, Ser



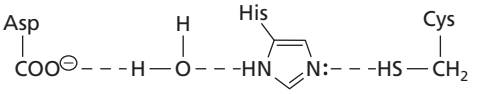
(b) β-Lactamase: Glu, Lys, Ser



(c) Asparaginase: Asp, Lys, Thr

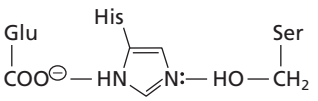


(d) Hepatitis A protease:

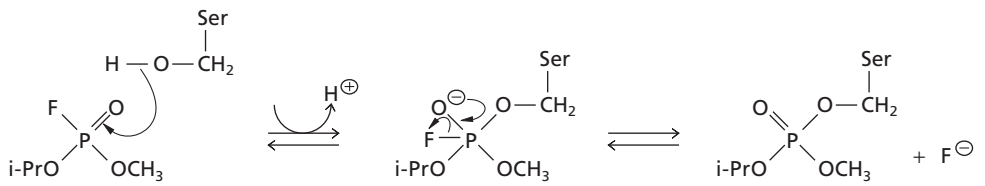


9. When tyrosine was mutated to phenylalanine, the activity of the mutant enzyme was less than 1% of the wild-type enzyme. Thus, the tyrosine residue is involved in the catalytic activity of DDP-IV. Tyrosine contains an -OH group on the aromatic ring of the side chain. As previously stated, this tyrosine is found in the oxyanion hole of the active site. Hydrogen bonds in the oxyanion hole of serine proteases are known to stabilize the tetrahedral intermediate. Tyrosine with an -OH group on the side chain can form a hydrogen bond and stabilize the tetrahedral intermediate. Phenylalanine does not have a side chain that can form a hydrogen bond. Therefore, the tetrahedral intermediate will not be stabilized resulting in a loss of enzyme activity.

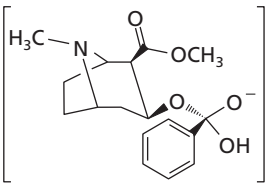
10. (a) Acetylcholinesterase catalytic triad: Glu-His-Ser



(b)



11. Transition-state analogs bound to carrier proteins are used as antigens to induce the formation of antibodies with catalytic activity. The tetrahedral phosphonate ester molecule is an analog of the tetrahedral intermediate structure in the transition state for hydrolysis of the benzyl ester moiety of cocaine. An antibody raised against the phosphonate structure that was able to stabilize the transition state of the cocaine benzyl ester hydrolysis could effectively catalyze this reaction.



Transition state

12. (a) Wild-type α1-proteinase inhibitor is given as treatment to individuals who produce an α1-proteinase inhibitor with substitutions in the amino acid sequence. These changes result in a protein that does not effectively inhibit the protease elastase. Uncontrolled elastase activity leads to increased breakdown of elastin, leading to destructive lung disease. Therefore, these patients are given a functional elastase inhibitor.

(b) The treatment for α1-proteinase inhibitor deficiency is to administer the wild-type protein intravenously. If the protein is given orally, the enzymes present in the digestive tract will cleave the peptide bonds in the α1-proteinase inhibitor. By administering the drug directly into the bloodstream, the protein can circulate to the lungs to act at the site of the neutrophil elastase.

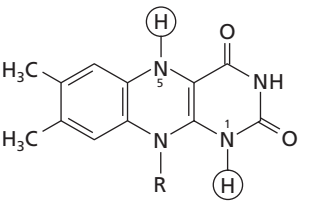
Chapter 7 Coenzymes and Vitamins

1. (a) Oxidation; $NAD^{\oplus}$, FAD, or FMN. (The coenzyme for the reaction shown is $NAD^{\oplus}$.)
(b) Decarboxylation of an α-keto acid; thiamine pyrophosphate.
(c) Carboxylation reaction requiring bicarbonate and ATP; biotin.
(d) Molecular rearrangement; adenosylcobalamin.
(e) Transfer of a hydroxyethyl group from TDP to CoA as an acyl group; lipoic acid.

2. (a) $NAD^{\oplus}$, $NADP^{\oplus}$, FAD, FMN, lipoamide, ubiquinone. Protein coenzymes such as thioredoxin and the cytochromes.
(b) Coenzyme A, lipoamide.
(c) Tetrahydrofolate, *S*-adenosylmethionine, methylcobalamin
(d) Pyridoxal phosphate
(e) Biotin, thiamine pyrophosphate, vitamin K

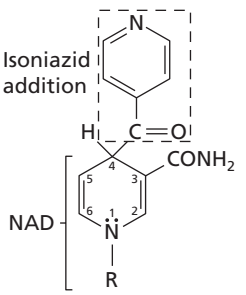
3. No. $NAD^{\oplus}$ acquires two electrons but only one proton. The second proton is released into solution and is reutilized by other proton-requiring reactions.

4.



5. $NAD^{\oplus}$, FAD, and coenzyme A all contain an ADP group (or ADP with 3′-phosphate for coenzyme A).

6.



7. Vitamin B_6 is converted to pyridoxal phosphate, which is the coenzyme for a large number of reactions involving amino acids, including the decarboxylation reactions in the pathways that produce serotonin and norepinephrine from tryptophan and tyrosine, respectively. Insufficient vitamin B_6 can lead to decreased levels of PLP and a decrease in the synthesis of the neurotransmitters.

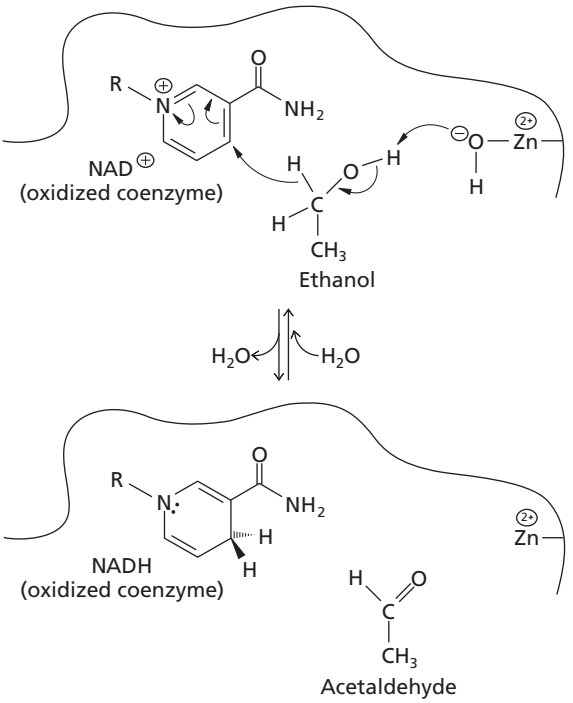
8. The synthesis of thymidylate (dTMP) requires a tetrahydrofolate (folic acid) derivative. Deficiency of folic acid decreases the amount of dTMP available for the synthesis of DNA. Decreased DNA synthesis in red blood cell precursors results in slower cell division, producing macrocytic red blood cells. The loss of cells by rupturing causes anemia.

9. (a) Cobalamin.
(b) The cobalamin derivative adenosylcobalamin is a coenzyme for the intramolecular rearrangement of methylmalonyl CoA to succinyl CoA (Figure 7.28). A deficiency of adenosylcobalamin results in increased levels of methylmalonyl CoA and its hydrolysis product, methylmalonic acid. Another cobalamin derivative, methylcobalamin, is a coen-

zyme for the synthesis of methionine from homocysteine (Reaction 7.5), and a deficiency of cobalamin results in an excess of homocysteine and a deficiency of methionine.

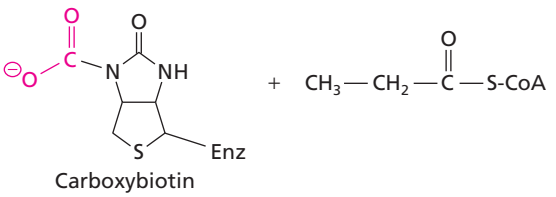
(c) Plants do not synthesize cobalamin and are therefore not a source of this vitamin.

10. (a) In one proposed mechanism, a water molecule bound to the zinc ion of alcohol dehydrogenase forms $OH^{\ominus}$, in the same manner as the water bound to carbonic anhydrase (Figure 7.2). The basic hydroxide ion abstracts the proton from the hydroxyl group of ethanol to form H_2O. (Another mechanism proposes that the zinc also binds to the alcoholic oxygen of the ethanol, polarizing it.)

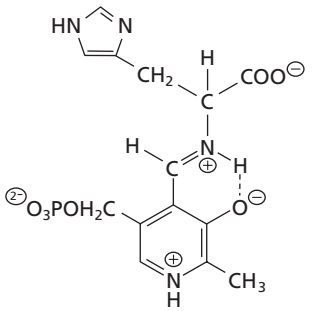


(b) No, a residue such as arginine is not required. Ethanol, unlike lactate, lacks a carboxylate group that can bind electrostatically to the arginine side chain.

11. A carboxyl group is transferred from methylmalonyl CoA to biotin to form carboxybiotin and propionyl CoA.



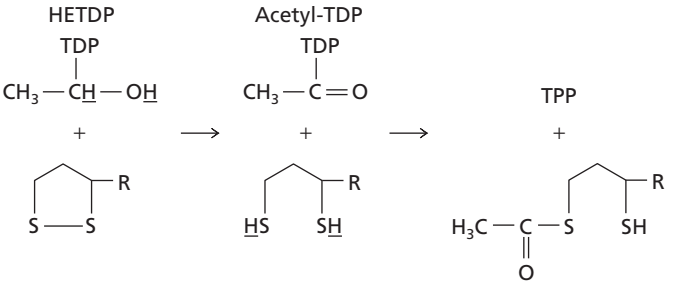
12. (a)



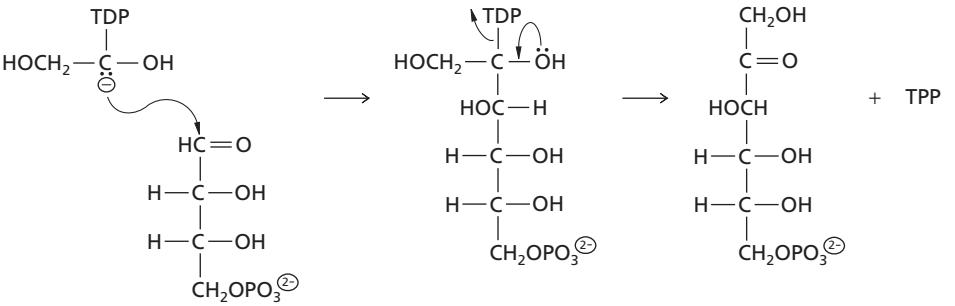
(b) Racemization would not occur. Although a Schiff base forms during decarboxylation as well as racemization, the reactive groups in the histidine decarboxylase active site specifically catalyze decarboxylation, not racemization, of histidine.

13. (a) See Reactions 13.2–13.4 on pages 412 and 413.

(b)



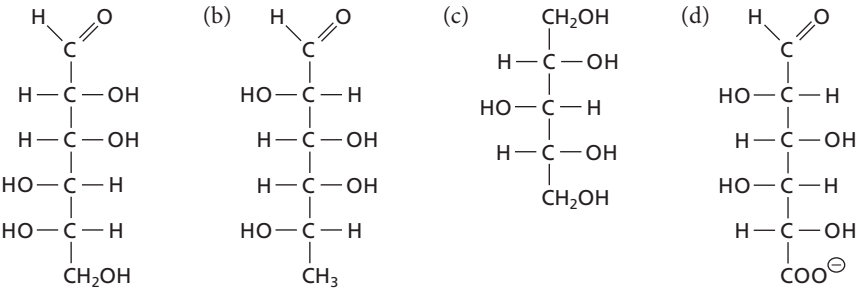
(c)



Chapter 8 Carbohydrates

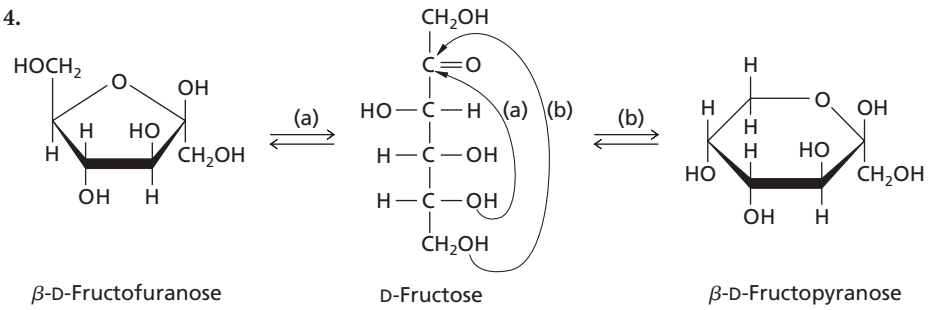
1. (a) D-Glucose and D-mannose
(b) L-Galactose
(c) D-Glucose or D-talose
(d) Dihydroxyacetone
(e) Erythrulose (either D or L)
(f) D-Glucose
(g) *N*-Acetylglucosamine

2. (a) (b) (c) (d)



3. Glycosaminoglycans are unbranched heteroglycans of repeating disaccharide units. One component of the disaccharide is an amino sugar and the other component is usually an al-duronic acid. Specific hydroxyl and amino groups of many glycosaminoglycans are sulfated

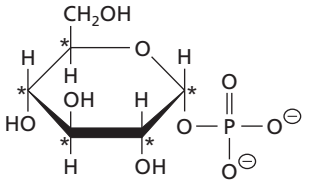
4.



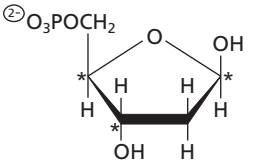
5. (a) α-Anomer
(b) Yes, it will mutorotate.
(c) Yes, it is a deoxy sugar.
(d) A ketone
(e) Four chiral carbons

6. Glucopyranose has five chiral carbons and 2^5, or 32, possible stereoisomers; 16 are D sugars and 16 are L sugars. Fructofuranose has four chiral carbons and 2^4, or 16, possible stereoisomers; 8 are D sugars and 8 are L sugars.

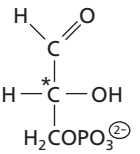
7. (a)



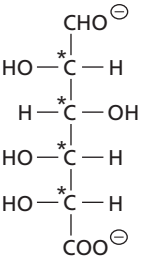
(b)



(c)



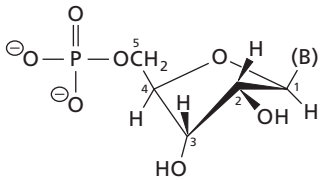
(d)



8. Only the open-chain forms of aldoses have free aldehyde groups that can form Schiff bases with amino groups of proteins. Because relatively few molecules of D-glucose are found in the open-chain form, D-glucose is less likely than other aldoses to react with proteins.

9. A pyranose is most stable when the bulkiest ring substituents are equatorial, minimizing steric repulsion. In the most stable conformer of β-D-glucopyranose, all the hydroxyl groups and the $-CH_2OH$ group are equatorial; in the most stable conformer of α-D-glucopyranose, the C-1 hydroxyl group is axial.

10.

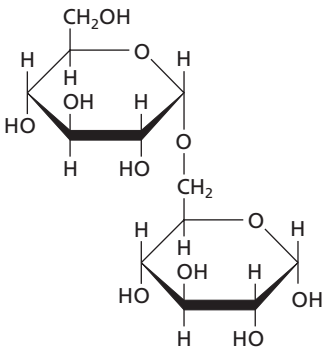


Envelope conformation

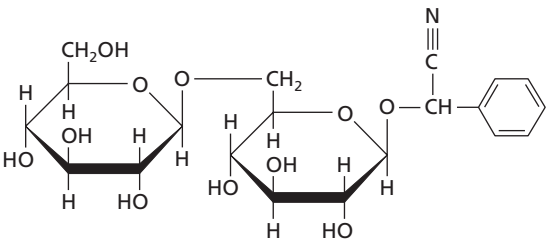
11. The α and β anomers of glucose are in rapid equilibrium. As β-D-glucose is depleted by the glucose oxidase reaction, more β anomer is formed from the α anomer until all the glucose has been converted to gluconolactone.

12. Sucralose is a derivative of the disaccharide sucrose (see Figure 8.20). The two hydroxyl groups on C-1 and C-6 of the fructose molecule have been replaced with chlorine. The hydroxyl group on C-4 of the glucose molecule was removed and then chlorine added. In the chemical synthesis of sucralose from sugar, the configuration of the C-4 substituent of the glucose moiety is reversed.

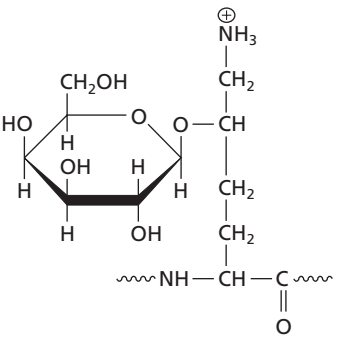
13. (a)



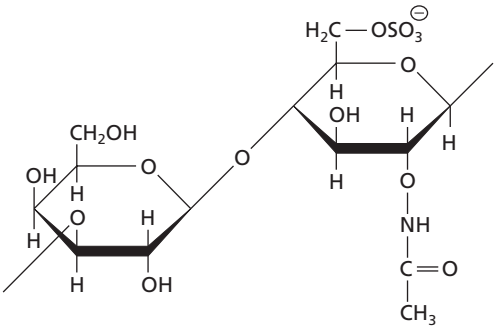
(b)



(c)

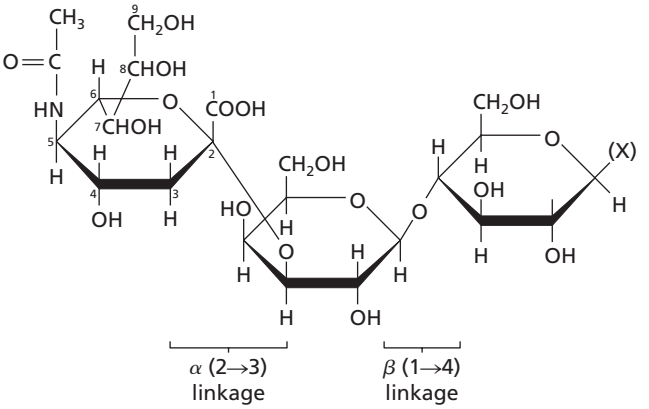


14.



15. (a) a, b, and c; these oligosaccharides contain GlcNAc — Asn bonds.
(b) b and c; these oligosaccharides contain β-galactosidic bonds.
(c) b; this oligosaccharide contains sialic acid.
(d) None, since none of the oligosaccharides shown contains fucose.

16.



17. Paper is made of cellulose and β-glucosidases break down cellulose to glucose residues. If you took a pill, this book would still taste like chewed up paper that tastes like paste (ugh!). That's because your taste buds are in your mouth and the enzyme is in your stomach. If you marinate the book in an enzyme solution, it would taste much sweeter.
Publishers would not print textbooks using flavored ink because they, and the authors, want students to keep their textbooks as valuable resources for future reference in the many advanced courses that you are planning to take. On the other hand, encouraging students to eat their textbooks, instead of selling them, might be a good thing because it promotes better health and nutrition.

Chapter 9 Lipids and Membranes

1.

(a) $CH_3(CH_2)_7-C(H)=C(H)-(CH_2)_{13}COO^{\ominus}$

(b) $CH_3(CH_2)_5-C(H)=C(H)-(CH_2)_9COO^{\ominus}$

(c) $CH_3CH_2(C(H)=C(H)CH_2)_5(CH_2)_2COO^{\ominus}$

2.

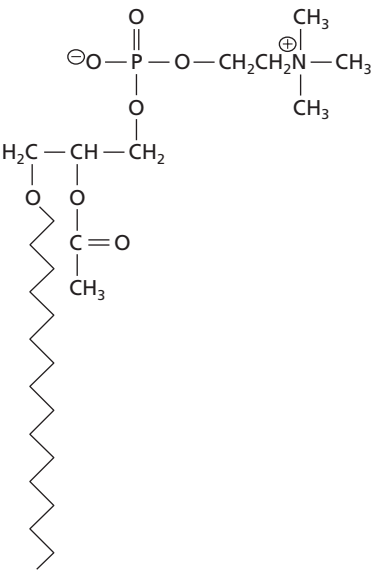
(a) $CH_3CH_2CH_2-O-(CH_2)_9COO^{\ominus}$

(b) $CH_3(CHCH_2CH_2CH_2)_3CHCH_2COO^{\ominus}$ (with CH_3 branches)

(c) $CH_3(CH_2)_5CH-CH(CH_2)_9COO^{\ominus}$ (cyclopropane ring)

3. (a) ω-3; (b) ω-6; (c) ω-6; (d) neither (ω-9); (e) ω-6.

4.

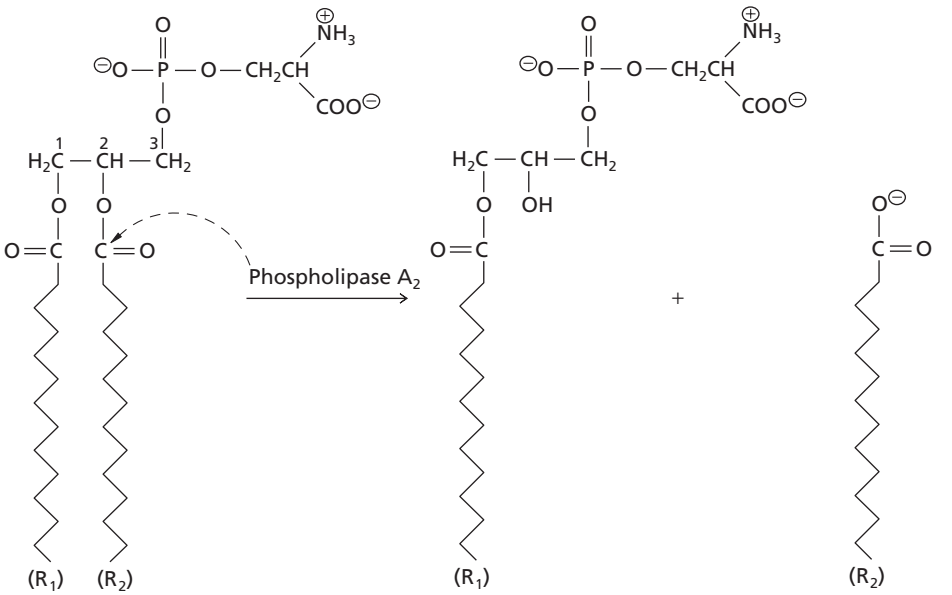


5. (a)

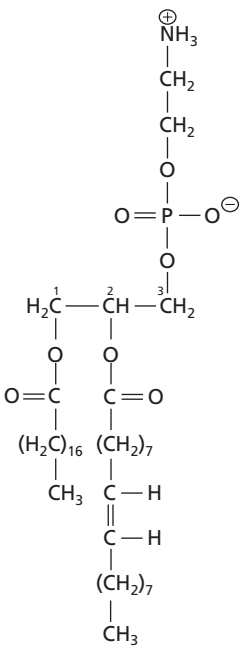


(b) Docosahexaenoic acid is classified as an ω-3 fatty acid

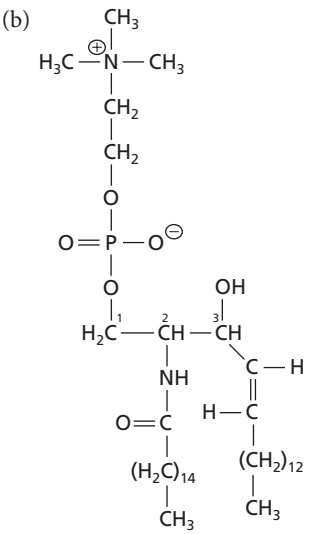
6.



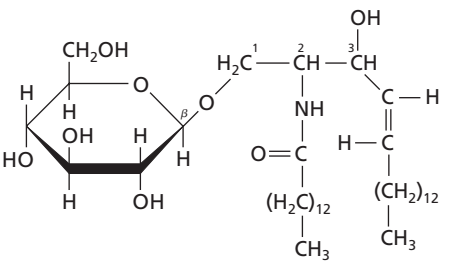
7. (a)



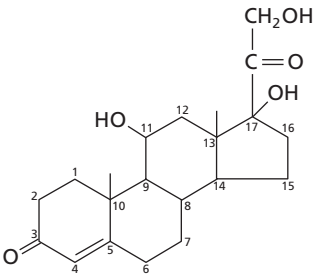
(b)



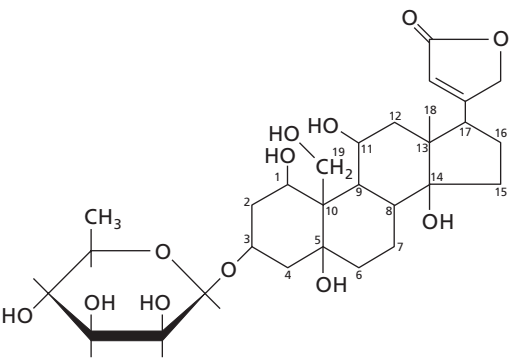
(c)



8. (a)



(b)

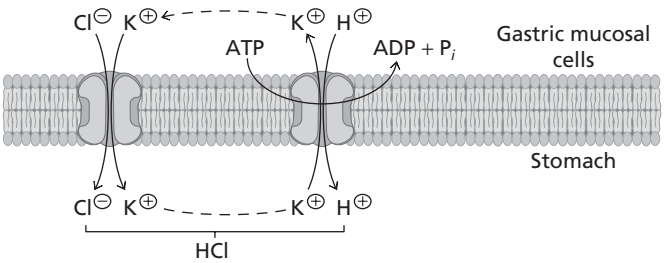


9. PE contains docosahexaenoic acid at position C-2 on the glycerol-3-phosphate backbone at both temperatures. At lower temperatures, the percent of the monounsaturated fatty acyl groups at position C-1 increased from 14% at 30°C to 39% at 10°C. The membrane fluidity must be maintained for the organism, and this is accomplished by changing the composition of the membrane lipids. The increase in the unsaturated lipids at the lower temperature will allow for the proper membrane fluidity.

10. Farnesyl transferase adds a farnesyl or "prenyl" group to a cysteine side chain of the *ras* protein (Figure 9.23b). The *ras* protein is subsequently anchored to the plasma and endoplasmic reticulum membranes and is active in cell signaling processes. Farnesyl transferase is a chemotherapy target because inhibition of this enzyme in tumor cells would disrupt the signaling activity of the mutated *ras* protein. In fact, farnesyl transferase (FT) inhibitors are potent suppressors of tumor growth in mice.

11. Line A represents diffusion of glucose through a channel or pore, and line B represents passive transport. Diffusion through a channel or pore is generally not saturable, with the rate increasing linearly with the concentration of the solute. Transport via a transport protein is saturable at high solute concentrations, much like an enzyme is saturated at high substrate concentrations (Section 9.10C).

12.



13. Theobromine is structurally related to caffeine and theophylline (Figure 9.45). The methylated purines, including the obromine, inhibit cAMP phophodiesterase, a soluble enzyme that catalyzes the hydrolysis of cAMP to AMP (Figure 9.43). These methylated purines inhibit the breakdown of the intracellular messenger cAMP to AMP. Therefore, the effects of the cAMP are prolonged. For dogs, this is combined with the fact that they have slower clearance of the ingested theobromine from their system. Both of these result in the toxicity associated with ingesting the chocolate.

14. The two second messengers IP_3 and DAG are complementary in that they both promote the activation of cellular kinases, which then activate intracellular target proteins by causing their phosphorylation. Diacylglycerol activates protein kinase C directly, whereas IP_3 elevates Ca^{2+} levels by opening a Ca^{2+} channel in the membrane of the endoplasmic reticulum, releasing stored Ca^{2+} into the cytosol (Figure 9.48). The increased Ca^{2+} levels activate other kinase leading to a phosphorylation and activation of certain target proteins.

15. Insulin can still bind normally to the α subunits of the insulin receptor, but due to the mutation, the β subunits lack tyrosine-kinase activity and cannot catalyze autophosphorylation or other phosphorylation reactions. Therefore, insulin does not elicit an intracellular response. The presence of more insulin will have no effect.

16. G proteins are molecular switches with two interconvertible forms, an active GTP-bound form and an inactive GDP-bound form (Figure 9.42). In normal G proteins, GTPase activity converts the active G protein to the inactive form. Because the *ras* protein lacks GTPase activity,

it cannot be inactivated. The result is continuous activation of adenylyl cyclase and prolonged responses to certain extracellular signals.

17. The surface of a sphere is $4\pi r^2$. The surface area of the oocyte is $4\pi(50)^2\mu m$, or $3.9 \times 10^5 \mu m^2$. The surface area of a lipid molecule is $10^{-14} cm^2 = 10^{-6}\mu m^2$. Since only 75% of the membrane is lipid, the total number of lipid molecules is

$$\frac{3.9 \times 10^5}{10^{-6}} \times 0.75 = 2.9 \times 10^{11} \text{ molecules}$$

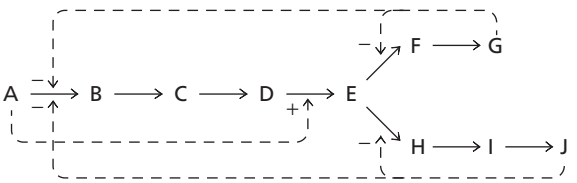
18. Assuming that the lipid molecules made by your grandmother are equally divided between daughter cells at each cell division, then after 30 cell divisions the oocyte (egg cell) produced by your mother will have $1/2^{30}$ of the original lipid molecules. Since the number of lipid molecules she inherited from her mother (your grandmother) was 2.9×10^{11} (see previous question), then the number remaining in each oocyte was

$$1/2^{30} \times 2.9 \times 10^{11} = 270$$

You inherited 270 lipid molecules from your grandmother.

Chapter 10 Introduction to Metabolism

1. (a)



(b) Inhibition of the first step in the common pathway by either G or J prevents the needless accumulation of intermediates in the pathway. When there is ample G or J, fewer molecules of A enter the pathway. By regulating an enzyme after the branch point, G or J inhibits its own production without inhibiting production of the other.

2. Compartmentalizing metabolic processes allows optimal concentrations of substrates and products for each pathway to exist independently in each compartment. In addition, separation of pathway enzymes also permits independent regulation of each pathway without interference by regulators from the other pathway.

3. Bacteria are much smaller than most eukaryotic cells so having separate compartments may not be as much of an advantage. It's also possible that localizing the citric acid cycle in mitochondria may be an historical accident rather than a selective advantage in eukaryotes.

4. In a multistep enzymatic pathway, the product from one enzyme will be the substrate for the next enzyme in the pathway. For independent soluble enzymes, the product of each enzyme must find the next enzyme by random diffusion in solution. By having sequential enzymes located in close proximity to each other, either in a multienzyme complex or on a membrane, the product of each enzyme can be passed directly on to the next enzyme without losing the substrate by diffusion into solution.

5. (a) $\Delta G^{\circ\prime} = RT \ln K_{eq}$

$$\ln K_{eq} = -\frac{\Delta G^{\circ\prime}}{RT} = -\frac{-9000 \text{ J mol}^{-1}}{(8.315 \text{ J K}^{-1} \text{ mol}^{-1})(298 \text{ K})} = 3.63$$

$$K_{eq} = 38$$

(b) $\Delta G^{\circ\prime} = -RT \ln K_{eq}$

$$K_{eq} = \frac{[\text{Glucose}][P_i]}{[\text{Glucose 6-P}][H_2O]} = \frac{(0.1 \text{ M})(0.1 \text{ M})}{(3.5 \times 10^{-5} \text{ M})(1)} = 286$$

$$\Delta G^{\circ\prime} = -(8.315 \text{ JK}^{-1} \text{ mol}^{-1})(298 \text{ K}) \ln 286$$

$$\Delta G^{\circ\prime} = -14\,000 \text{ J mol}^{-1} = -14 \text{ kJ mol}^{-1}$$

6. (a) $\Delta G = \Delta G^{\circ\prime} + RT \ln \frac{[\text{Arginine}][P_i]}{[\text{Phosphoarginine}][H_2O]}$

$$\Delta G = -32\,000 \text{ J mol}^{-1} + (8.315 \text{ J K}^{-1} \text{ mol}^{-1})(298 \text{ K}) \ln \frac{(2.6 \times 10^{-3})(5 \times 10^{-3})}{(6.8 \times 10^{-3})(1)}$$

$$\Delta G = -48 \text{ kJ mol}^{-1}$$

(b) $\Delta G^{\circ\prime}$ is defined under standard conditions of 1 M concentrations of reactants and products. (The concentration of water is assigned a value of 1.) ΔG depends on the actual concentrations of the reactants and products.

(c) Molecules with high free energies of hydrolysis, such as phosphoarginine and acetyl CoA, are thermodynamically unstable but may be kinetically stable. These molecules are hydrolyzed very slowly in the absence of an appropriate catalyst.

7.

	$\Delta G^{\circ\prime}$(kJ mol^{-1})
Glucose 1-phosphate + UTP $\longrightarrow$ UDP-glucose + PP_i	0
$PP_i + H_2O \longrightarrow 2\ P_i$	−29
	$\Delta G^{\circ\prime}$ −29

8. (a) Although ATP is rapidly utilized for energy purposes such as muscle contraction and membrane transport, it is also rapidly resynthesized from ADP and P_i through intermediary metabolic routes. Energy for this process is supplied from the degradation of carbohydrates, fats, and amino acids or from energy storage molecules such as muscle creatine phosphate (CP + ADP → ATP + C). With this rapid recycling, 50 grams total of ATP and ADP is sufficient for the chemical energy needs of the body.

(b) The role of ATP is that of a free energy transmitter rather than an energy storage molecule. As indicated in part (a), ATP is not stored, but is rapidly utilized in energy-requiring reactions.

9. $\Delta G^{\circ\prime}$ for the reaction of ATP and creatine is calculated as

	$\Delta G^{\circ\prime}$(kJ mol^{-1})
Creatine + $P_i \longleftrightarrow$ Phosphocreatine + H_2O	+43
ATP + $H_2O \longleftrightarrow$ ADP + P_i	−32
Creatine + ATP $\longleftrightarrow$ Phosphocreatine + ADP	+11

The ratio of ATP to ADP needed to maintain a 20:1 ratio of phosphocreatine to creatine is calculated from Equation 10.13. At equilibrium, $\Delta G = 0$, so

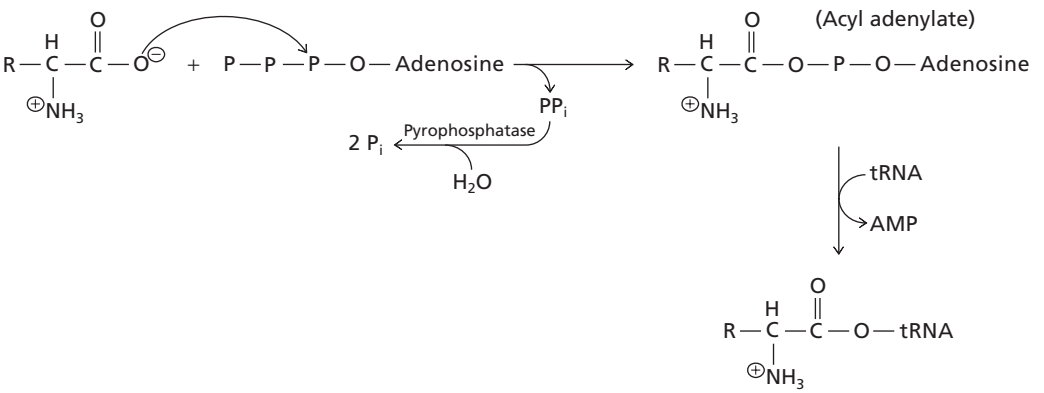
$$\Delta G^{\circ\prime} = -RT \ln \frac{[\text{Phosphocreatine}][\text{ADP}]}{[\text{Creatine}][\text{ATP}]}$$

$$\ln \frac{(20)[\text{ADP}]}{(1)[\text{ATP}]} = -\frac{\Delta G^{\circ\prime}}{RT} = -\frac{(11\ 000\ \text{J mol}^{-1})}{(8.315\ \text{J K}^{-1}\ \text{mol}^{-1})(298\ \text{K})} = -4.44$$

$$\frac{(20)[\text{ADP}]}{(1)[\text{ATP}]} = 1.2 \times 10^{-2}$$

$$\frac{[\text{ATP}]}{[\text{ADP}]} = 1667{:}1$$

10.



11. $\Delta G^{\circ\prime} = -RT \ln K_{eq}$

$$K_{eq} = \frac{[\text{fructose-6-phosphate}]}{[\text{glucose-6-phosphate}]} = \frac{2}{1}$$

$$\Delta G^{\circ\prime} = -(8.315\ \text{J K}^{-1}\ \text{mol}^{-1})(298\ \text{K}) \ln 2$$

$$\Delta G^{\circ\prime} = -1.7\ \text{kJ mol}^{-1}$$

12. (a) $\ln K_{eq} = -\dfrac{\Delta G^{o\prime}}{RT} = -\dfrac{(25\,000\ \text{J mol}^{-1})}{(8.315\ \text{J K}^{-1}\ \text{mol}^{-1})(298\ \text{K})} = -10.1$

$K_{eq} = 4.1 \times 10^{-5}$

(b) $\Delta G^{o\prime}$ for the coupled reaction is calculated as

	$\Delta G^{o\prime}$(kJ mol^{-1})
$A \longrightarrow B$	$+25$
$ATP + H_2O \longleftrightarrow ADP + P_i$	-32
$A + ATP + H_2O \longleftrightarrow B + ADP + P_i$	-7

$\ln K_{eq} = -\dfrac{\Delta G^{o\prime}}{RT} = 2.8$

$K_{eq} = 17$

K_{eq} for the coupled reaction is about 180,000 times larger than K_{eq} in part (a).

(c) $K_{eq} = 17 = \dfrac{[B][ADP][P_i]}{[A][ATP][H_2O]} = \dfrac{[B][ADP]}{[A][ATP]} = \dfrac{[B](1)}{[A](400)}$

$\dfrac{[B]}{[A]} = 6800:1$

Coupling the reaction to ATP hydrolysis increases the ratio of [B] to [A] by a factor of about 166 million ($6800 \div (4.1 \times 10^{-5}) = 1.6 \times 10^{8}$).

13. Electrons flow from the molecule with a more negative standard reduction potential to the molecule with a more positive standard reduction potential.

(a) Cytochrome $b_5(Fe^{2+})$ + Cytochrome $f(Fe^{3+}) \rightarrow$
Cytochrome $b_5(Fe^{3+})$ + Cytochrome $f(Fe^{2+})$

(b) Succinate + Q → Fumarate + QH_2

(c) Isocitrate + $NAD^{\oplus}$ → α-Ketoglutarate + NADH

14. The standard reduction potentials in Table 10.4 refer to half-reactions that are written as $S_{ox} + n\,e^{\ominus} \rightarrow S_{red}$. Two half-reactions can be added to obtain the coupled oxidation–reduction reaction by reversing the direction of the half-reaction involving the reduced species and reversing the sign of its reduction potential.

(a) 2 Cyt $c(Fe^{3+}) + 2e^{\ominus} \longrightarrow$ 2 Cyt $c(Fe^{2+})$
$QH_2 \longrightarrow Q + 2\,H^{\oplus} + 2e^{\ominus}$
2 Cyt $c(Fe^{3+}) + QH_2 \longrightarrow$ 2 Cyt $c(Fe^{2+}) + Q + 2\,H^{\oplus}$

$E^{o\prime}$(V)
$+0.23$
-0.04
$\Delta E^{o\prime} = 0.19$ V

$\Delta G^{o\prime} = -nF\Delta E^{o\prime} = -(2)(96.48\ \text{kJ V}^{-1}\ \text{mol}^{-1})(0.19\ \text{V})$
$\Delta G^{o\prime} = -37\ \text{kJ mol}^{-1}$

(b) $\frac{1}{2}\,O_2 + 2\,H^{\oplus} + 2\,e^{\ominus} \longrightarrow H_2O$
Succinate → Fumarate $2\,H^{\oplus} + 2\,e^{\ominus}$
$\frac{1}{2}\,O_2$ + Succinate → H_2O + Fumarate

E^{o}(V)
$+0.82$
-0.03
$\Delta E^{o\prime} = 0.79$ V

$\Delta G^{o\prime} = -(2)(96.48\ \text{kJ V}^{-1}\ \text{mol}^{-1})(0.79\ \text{V})$
$\Delta G^{o\prime} = -150\ \text{kJ mol}^{-1}$

15. The expected results are as shown in the bottom graph. As NADH is formed in the reaction mixture, the absorbance at 340 nm will increase (see Box 10.1).

16. $Q + 2\,H^{\oplus} + 2\,e^{\ominus} \longrightarrow QH_2$
$FADH_2 \longrightarrow FAD + 2\,H^{\oplus} + 2\,e^{\oplus}$
$Q + FADH_2 \longrightarrow QH_2 + FAD$

$E^{o\prime}$(V)
$+0.04$
$+0.22$
$\Delta E^{o\prime} = 0.26$ V

$$\Delta E = \Delta E^{o\prime} - \frac{RT}{nF}\ln\frac{[QH_2][FAD]}{[Q][FADH_2]}$$

$$\Delta E = 0.26\ \text{V} - \frac{0.026\ \text{V}}{2}\ln\frac{(5 \times 10^{-5})(2 \times 10^{-4})}{(1 \times 10^{-4})(5 \times 10^{-3})}$$

$$\Delta E = 0.26\ \text{V} - 0.013(-3.9) = 0.31\ \text{V}$$

$$\Delta G = -nF\Delta E = -(2)(96.48\ \text{kJ V}^{-1}\ \text{mol}^{-1})(0.31\ \text{V})$$

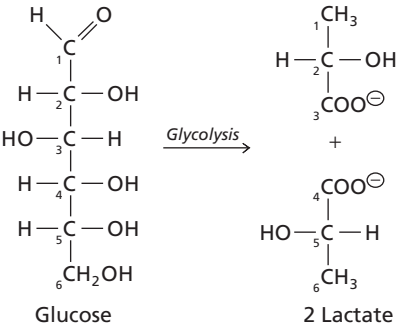
$$\Delta G = -60\ \text{kJ mol}^{-1}$$

Theoretically, the oxidation of $FADH_2$ by ubiquinone liberates more than enough free energy to drive ATP synthesis from ADP and P_i.

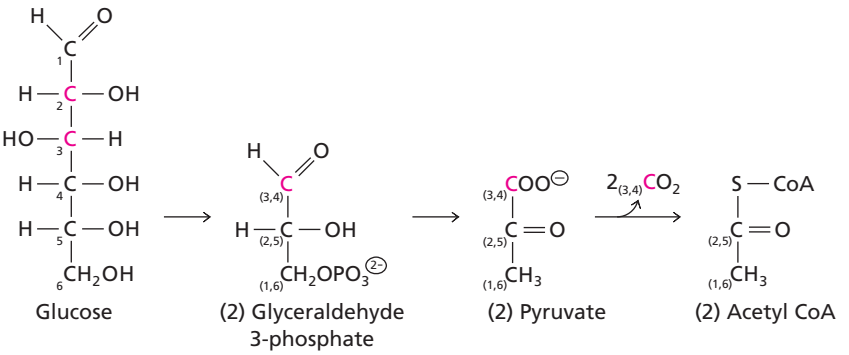
Chapter 11 Glycolysis

1. (a) 2 (see Figure 11.2 and Reaction 11.12)
(b) 2 (1 ATP is consumed by the fructokinase reaction, 1 ATP is consumed by the triose kinase reaction, and 4 ATP are generated by the triose stage of glycolysis)
(c) 2 (2 ATP are consumed in the hexose stage, and 4 ATP are generated by the triose stage)
(d) 5 (2 ATP are obtained from fructose, as in part (b), and 3 ATP—rather than 2—are obtained from the glucose moiety since glucose 1-phosphate, not glucose, is formed when sucrose is cleaved)

2. (a)



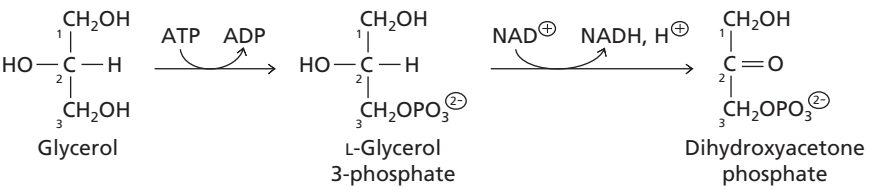
(b) Glucose labeled at either C-3 or C-4 yields $^{14}CO_2$ from the decarboxylation of pyruvate.



3. Inorganic phosphate ($^{32}P_i$) will be incorporated into 1,3-*bis*phosphoglycerate (1,3 BPG) at the C-1 carbon in the glyceraldehyde 3-phosphate dehydrogenase (GADPH) reaction—glyceraldehyde 3-phosphate + $NAD^{\oplus}$ + $P_i \rightarrow$ 1,3 BPG—and then transferred to the γ-position of ATP in the next step: 1,3 BPG + ADP $\rightarrow$ ATP + 3-phosphoglycerate.

4. Since the brain relies almost solely on glucose for energy, it is dependent on glycolysis as the major pathway for glucose catabolism. Since the Huntington protein binds tightly to GAPDH, this suggests that it might inhibit this crucial glycolytic enzyme and thereby impair the production of ATP. Decreased ATP levels would be detrimental to neuronal cells in the brain.

5. (a)



(b) C-2 and C-3 of glycerol 3-phosphate must be labeled. Once dihydroxyacetone phosphate is converted to glyceraldehyde 3-phosphate, C-1 is oxidized to an aldehyde and subsequently lost as CO_2 (Problem 2).

6. Cells that metabolize glucose to lactate by anaerobic glycolysis produce far less ATP per glucose than do cells that metabolize glucose aerobically to CO_2 via glycolysis and the citric acid cycle (Figure 11.1). More glucose must be utilized via anaerobic glycolysis to produce a sufficient amount of ATP for cellular needs, and the rate of conversion of glucose to lactate is

much higher than under aerobic conditions. Cancer cells in an anaerobic environment take up far more glucose and may overproduce some glycolytic enzymes to compensate for the increase in the activity of this pathway of carbohydrate metabolism.

7. No. The conversion of pyruvate to lactate, catalyzed by lactate dehydrogenase, oxidizes NADH to $NAD^{\oplus}$, which is required for the glyceraldehyde 3-phosphate dehydrogenase reaction of glycolysis.

8. In the reactions catalyzed by these enzymes, the bond between the γ-phosphorus atom and the oxygen of the β-phosphoryl group is cleaved when the γ-phosphoryl group of ATP is transferred (Figure 11.3). The analog cannot be cleaved in this way and therefore inhibits the enzymes by competing with ATP for the active site.

9. The free energy change for the aldolase reaction under standard conditions ($\Delta G^{\circ\prime}$) is $+22.8$ kJ mol^{-1}. The concentrations of fructose 1,6-*bis*phosphate, dihydroxyacetone phosphate, and glyceraldehyde 3-phosphate in heart muscle, however, are much different than the 1 M concentrations assumed under standard conditions. The actual free energy change under cellular concentrations ($\Delta G^{\circ\prime} = -5.9$ kJ mol^{-1}) is much different than $\Delta G^{\circ\prime}$, and the aldolase reaction readily proceeds in the direction necessary for glycolysis: Fructose 1, 6-*bis*phosphate ⟶ glyceraldehyde 3-phosphate + dihydroxyacetone phosphate.

10. The standard Gibbs free energy change is + 28 kJ mol^{-1}. The equilibrium constant is

$$28 = RT \ln K_{eq} \approx 10^{-5} \quad \text{(Equation 1.12)}$$

(a) $$\frac{[\text{DHAP}][\text{G3P}]}{[\text{FBP}]} = 10^{-5} \quad \frac{[5 \times 10^{-6}][5 \times 10^{-6}]}{[\text{FBP}]} = 10^{-5} \quad \text{FBP} = 2.5\ \mu\text{M}$$

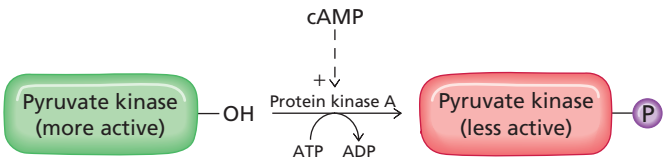
(b) 250μM

(c) 25,000 μM = 25 *m*M

11. (a) ATP is both a substrate and an allosteric inhibitor for PFK-1. Higher concentrations of ATP result in a decrease in the activity of PFK-1 due to an increase in the K_m. AMP is an allosteric activator that acts by relieving the inhibition caused by ATP, thus raising the curve when AMP is present with ATP.

(b) F2,6P is an allosteric activator of PFK-1. In the presence of F2,6P the activity of PFK-1 is increased due to a decrease in the apparent K_m for fructose 6-phosphate.

12. Increased [cAMP] activates protein kinase A, which catalyzes the phosphorylation and inactivation of pyruvate kinase.



13. (a) A decrease in glycolysis in the liver makes more glucose available for export to other tissues.

(b) Decreased activity of the glucagon transducer system decreases the amount of cAMP formed. As existing cAMP is hydrolyzed by the activity of a phosphodiesterase, cAMP-dependent protein kinase A becomes less active. Under these conditions, PFK-2 activity increases and fructose 2,6-*bis*phosphatase activity decreases (Figure 11.18). The resulting increase in fructose 2,6-*bis*phosphate activates PFK-1, increasing the overall rate of glycolysis. A decrease in cAMP also leads to the activation of pyruvate kinase (Problem 12).

14. Chemoautotrophs use glycolysis to generate energy from stored glucose residues in glycogen as described in Chapter 12.

Chapter 12 Gluconeogenesis, The Pentose Phosphate Pathway, and Glycogen Metabolism

1. 2 pyruvate + 2NADH + 4 ATP + 2 GTP + 6 H_2O + 2 $H^{\oplus}$ ⟶
   glucose + 2 $NAD^{\oplus}$ + 4 ADP + 2 GDP + 6 P_i

   2 NADH ≡ 5 ATP equivalents

   4 ATP = 4 ATP

   $$2\text{GTP} \equiv \frac{2\ \text{ATP}}{11\ \text{ATP}}$$

The energy required to synthesize one molecule of glucose 6-phosphate from CO_2 can be calculated from Reaction 12.7.

$$12\ \text{NADPH} \equiv 30\ \text{ATP}$$

The conversion of G6P to glucose does not require or produce ATP equivalents. The synthesis of glucose from pyruvate via the gluconeogenesis pathway is only about one third (11/30) as expensive as the synthesis of glucose from CO_2.

2. Reducing power in the form of NADH (2), and ATP (4) and GTP (2) are required for the synthesis of glucose from pyruvate (Equation 12.1). The NADH and GTP are direct products of the citric acid cycle, and ATP can be generated from NADH and QH_2($FADH_2$) during the oxidative phosphorylation process.

3. Epinephrine interacts with the liver β-adrenergic receptors and activates the adenylyl cyclase signaling pathway, leading to cAMP production and activation of protein kinase A (Figure 12.15). Protein kinase A activates phosphorylase kinase, which in turn activates glycogen phosphorylase (GP), leading to glycogen degradation (Figure 12.16). Glucose can then be transported out of the liver and into the bloodstream, where it is taken up by muscles for needed energy production.

$$\text{Liver[Glycogen} \xrightarrow{\text{(GP)}} \text{GIP} \longrightarrow \text{G6P} \longrightarrow \text{Glucose]} \longrightarrow \text{Bloodstream} \longrightarrow \text{Muscles}$$

4. (a) Protein phosphatase-1 activated by insulin catalyzes the hydrolysis of the phosphate ester bonds on glycogen synthase (activating it) and on glycogen phosphorylase and phosphorylase kinase (inactivating them), as shown in Figure 12.17. Therefore, insulin stimulates glycogen synthesis and inhibits glycogen degradation in muscle cells.

(b) Only liver cells are rich in glucagon receptors, so glucagon selectively exerts its effects on liver enzymes.

(c) The binding of glucose to the glycogen phosphorylase–protein phosphatase-1 complex in liver cells relieves the inhibition of protein phosphatase-1 and makes glycogen phosphorylase more susceptible to dephosphorylation (inactivation) by protein phosphatase-1 (Figure 12.18). Protein phosphatase-1 also catalyzes the dephosphorylation of glycogen synthase, making it more active. Therefore, glucose stimulates glycogen synthesis and inhibits glycogen degradation in the liver.

5. Decreased concentrations of fructose 2,6-*bis*phosphate (F2,6BP) lead to a decreased rate of glycolysis and an increased rate of gluconeogenesis. F2,6BP is an activator of the glycolytic enzyme phosphofructokinase-1 (PFK-1), and lower F2,6BP levels will result in decreased rates of glycolysis. In addition, F2,6BP is an inhibitor of the gluconeogenic enzyme fructose 1,6-*bis*phosphatase, and therefore decreased levels of F2,6BP will decrease the inhibition and increase the rate of gluconeogenesis (Figure 12.4).

6. When glucagon binds to its receptor, it activates adenylyl cyclase. Adenylyl cyclase catalyzes the synthesis of cAMP from ATP. The cAMP activates protein kinase A. Protein kinase A catalyzes the phosphorylation of PFK-2, which inactivates the kinase activity and activates the phosphatase activity. Fructose 2,6-*bis*phosphatase catalyzes the hydrolytic dephosphorylation of fructose 2,6-*bis*phosphate to form fructose 6-phosphate. The resulting decrease in the concentration of fructose 2,6-*bis*phosphate relieves the inhibition of fructose 1,6-*bis*phosphatase, thereby activating gluconeogenesis. Thus, the kinase activity of PFK-2 is decreased.

7. (a) Yes. The synthesis of glycogen from glucose 6-phosphate requires the energy of one phosphoanhydride bond (in the hydrolysis of PP_i; Figure 12.10). However, when glycogen is degraded to glucose 6-phosphate, inorganic phosphate (P_i) is used in the phosphorolysis reaction. No "high energy" phosphate bond is used.

(b) One fewer ATP molecule is available for use in the muscle when liver glycogen is the source of the glucose utilized. Liver glycogen is degraded to glucose phosphates and then to glucose without consuming ATP. After transport to muscle cells, the glucose is converted to glucose 6-phosphate by the action of hexokinase in a reaction that consumes one molecule of ATP. Muscle glycogen, however, is converted directly to glucose 1-phosphate by the action of glycogen phosphorylase, which does not consume ATP. Glucose 1-phosphate is isomerized to glucose 6-phosphate by the action of phosphoglucomutase.

8. A deficiency of glycogen phosphorylase in the muscle prevents the mobilization of glycogen to glucose. Insufficient glucose prevents the production of ATP by glycolysis. Existing ATP used for muscle contraction is not replenished, thus increasing the levels of ADP and P_i. Since no glucose is available from glycogen in the muscle, no lactate is produced.

9. Converting glucose 1-phosphate to two molecules of lactate yields 3 ATP equivalents (1 ATP expended in the phosphofructokinase-1 reaction, 2 ATP produced in the phosphoglycerate

kinase reaction, and 2 ATP produced in the pyruvate kinase reaction). Converting two molecules of lactate to one molecule of glucose 1-phosphate requires 6 ATP equivalents (2 ATP in the pyruvate carboxylase reaction, 2 GTP in the PEP carboxykinase reaction, and 2 ATP in the phosphoglycerate kinase reaction).

10. (a) Muscle pyruvate from glycolysis or amino acid catabolism is converted to alanine by transamination. Alanine travels to the liver, where it is reconverted to pyruvate by transamination with α-ketoglutarate. Gluconeogenesis converts pyruvate to glucose, which can be returned to muscles.

(b) NADH is required to reduce pyruvate to lactate in the Cori cycle, but it is not required to convert pyruvate to alanine in the glucose-alanine cycle. Thus, the glucose-alanine cycle makes more NADH available in muscles for the production of ATP by oxidative phosphorylation.

11. (a) Inadequate glucose 6-phosphatase activity (G6P $\rightarrow$ glucose + P_i) leads to accumulation of intracellular G6P, which inhibits glycogen phosphorylase and activates glycogen synthase. This prevents liver glycogen from being mobilized. This results in increased glycogen storage (and enlargement of the liver) and low blood glucose levels (hypoglycemia).

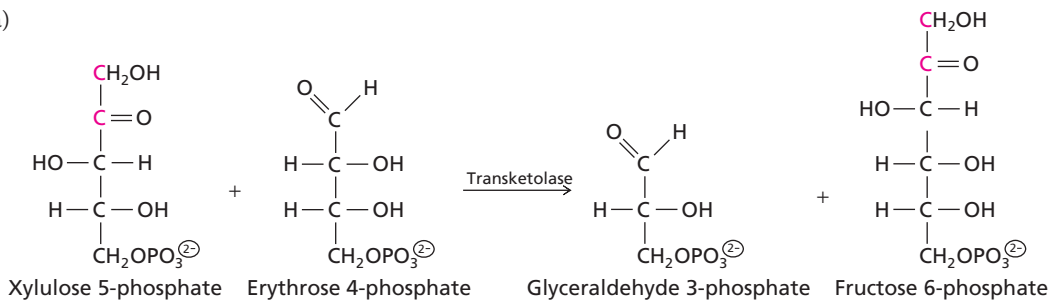
(b) Yes. A defective branching enzyme leads to accumulation of glycogen molecules with defective, short outer branches. These molecules cannot be degraded, so there will be much less efficient glycogen degradation for glucose formation. Low blood glucose levels result due to the impaired glycogen degradation.

(c) Inadequate liver phosphorylase activity leads to an accumulation of liver glycogen since the enzyme cleaves a glucose molecule from the nonreducing end of a glycogen chain. Low blood glucose levels result, due to the impaired degradation of glycogen.

12. Glucose 6-phosphate, glyceraldehyde 3-phosphate, and fructose 6-phosphate.

13. The repair of tissue injury requires cell proliferation and synthesis of scar tissue. NADPH is needed for the synthesis of cholesterol and fatty acids (components of cellular membranes), and ribose 5-phosphate is needed for the synthesis of DNA and RNA. Since the pentose phosphate pathway is the primary source of NADPH and ribose 5-phosphate, injured tissue responds to the increased demands for these products by increasing the level of synthesis of the enzymes in the pentose phosphate pathway.

14. (a)



(b) C-2 of glucose 6-phosphate becomes C-1 of xylulose 5-phosphate. After C-1 and C-2 of xylulose 5-phosphate are transferred to erythrose 4-phosphate, the label appears at C-1 of fructose 6-phosphate, as shown in part (a).

Chapter 13 The Citric Acid Cycle

1. (a) No net synthesis is possible since two carbons from acetyl CoA enter the cycle in the citrate synthase reaction and two carbons leave as CO_2 in the isocitrate dehydrogenase and α-ketoglutarate dehydrogenase reactions.

(b) Oxaloacetate can be replenished by the pyruvate carboxylase reaction, which carries out a net synthesis of OAA,

$$\text{Pyruvate} + CO_2 + \text{ATP} + H_2O \longrightarrow \text{Oxaloacetate} + \text{ADP} + P_i$$

This is the major anaplerotic reaction in some mammalian tissues. Many plants and some bacteria supply oxaloacetate via the phosphoenolpyruvate carboxykinase reaction,

$$\text{Phosphoenolpyruvate} + HCO_3^{\ominus} \longrightarrow \text{Oxaloacetate} + P_i$$

In most species, acetyl CoA can be converted to malate and oxaloacetate via the glyoxylate pathway.

2. Aconitase would be inhibited by fluorocitrate formed from fluoroacetate, leading to increased levels of citric acid and decreased levels of all subsequent citric acid cycle intermediates from

isocitrate to oxaloacetate. Since fluorocitrate is a competitive inhibitor, very high levels of citrate would at least partially overcome the inhibition of aconitase by fluorocitrate and permit the cycle to continue at some level.

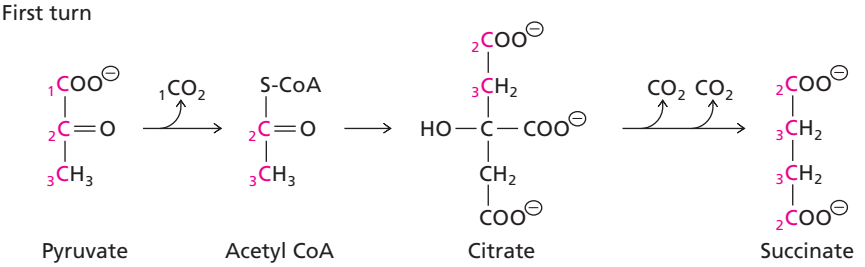
3. (a) 12.5; 10.0 from the cycle and 2.5 from the pyruvate dehydrogenase reaction.
(b) 10.0; 7.5 from oxidation of 3 NADH, 1.5 from oxidation of 1 QH_2, and 1.0 from the substrate-level phosphorylation catalyzed by CoA synthetase.

4. 87.5% (28 of 32) of the ATP is produced by oxidative phosphorylation, and 12.5% (4 of 32) is produced by substrate-level phosphorylation.

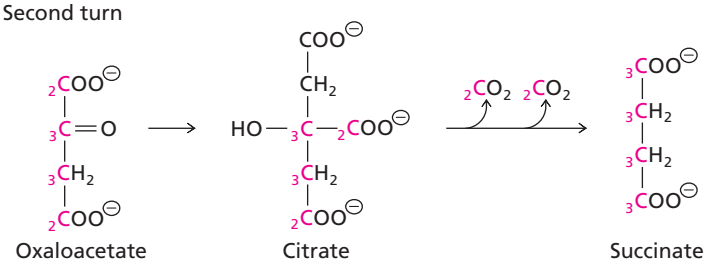
5. Thiamine is the precursor of the coenzyme thiamine pyrophosphate (TPP), which is found in two enzyme complexes associated with the citric acid cycle: the pyruvate dehydrogenase complex and the α-ketoglutarate dehydrogenase complex. A deficiency of TPP decreases the activities of these enzyme complexes. Decreasing the conversion of pyruvate to acetyl CoA and of α-ketoglutarate to succinyl CoA causes accumulation of pyruvate and α-ketoglutarate.

6. Since C-1 of pyruvate is converted to CO_2 in the reaction catalyzed by the pyruvate dehydrogenase complex, 1-[^{14}C]-pyruvate is the first to yield $^{14}CO_2$. Neither of the two acetyl carbon atoms of acetyl CoA is converted to CO_2 during the first turn of the citric acid cycle (Figure 13.5). However, the carboxylate carbon atoms of oxaloacetate, which arise from C-2 of pyruvate, become the two carboxylates of citrate that are removed as CO_2 during a second turn of the cycle. Therefore, 2-[^{14}C]-pyruvate is the second labeled molecule to yield $^{14}CO_2$. 3-[^{14}C]-Pyruvate is the last to yield $^{14}CO_2$, in the third turn of the cycle.

First turn



Second turn



Half of the ^{14}C is eliminated by the third turn of the cycle. An additional one-fourth is eliminated in the fourth turn, then one-eighth in the fifth turn, etc. It will take a very long time to eliminate all of the ^{14}C from the citric acid cycle intermediates.

7. (a) The NADH produced by the oxidative reactions of the citric acid cycle must be recycled back to $NAD^{\oplus}$, which is required for the pyruvate dehydrogenase reaction. When O_2 levels are low, fewer NADH molecules are reoxidized by O_2 (via the process of oxidative phosphorylation), so the activity of the pyruvate dehydrogenase complex decreases.
(b) Pyruvate dehydrogenase kinase catalyzes phosphorylation of the pyruvate dehydrogenase complex, thereby inactivating it (Figure 13.12). Inhibiting the kinase shifts the pyruvate dehydrogenase complex to its more active form.

8. A deficiency in the citric acid cycle enzyme fumarase would result in abnormally high concentrations of fumarate and prior cycle intermediates including succinate and α-ketoglutarate, which could lead to excretion of these molecules.

9. The different actions of acetyl CoA on two components of the pyruvate dehydrogenase (PDH) complex both lead to an inhibition of the pyruvate to acetyl CoA reaction. Acetyl CoA inhibits the E_2 component of the PDH complex directly (Figure 13.11). Acetyl CoA causes inhibition of the E_1 component indirectly by activating the pyruvate kinase (PK) component of

the PDH complex, and PK phosphorylates the E_1 component of the PDH complex, thus inactivating it (Figure 13.12).

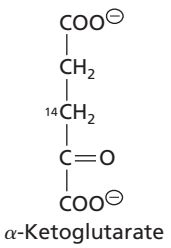
10. The pyruvate dehydrogenase complex catalyzes the oxidation of pyruvate to form acetyl CoA and CO_2. If there is reduced activity of this complex, then the pyruvate concentration will increase. Pyruvate will be converted to lactate through the action of lactate dehydrogenase. Lactate builds up since glycolytic metabolism is increased to synthesize ATP since oxidation of pyruvate to acetyl CoA is impaired. In addition, pyruvate is converted to alanine, as shown in Reaction 12.6.

11. Calcium activates both isocitrate dehydrogenase and α-ketoglutarate dehydrogenase in the citric acid cycle, thereby increasing this catabolic process and producing more ATP. In addition, Ca^{2+} activates the pyruvate dehydrogenase phosphatase enzyme of the PDH complex, which activates the E_1 component (Figure 13.12). Activation of the PDH complex converts more pyruvate into acetyl CoA for entry into the citric acid cycle, resulting in an increased production of ATP.

12. (a) Alanine degradation replenishes citric acid cycle intermediates, since pyruvate can be converted to oxaloacetate via the pyruvate carboxylase reaction, the major anaplerotic reaction in mammals (Reaction 13.19). Leucine degradation cannot replenish intermediates of the citric acid cycle, since for every molecule of acetyl CoA that enters the cycle, two molecules of CO_2 are lost.

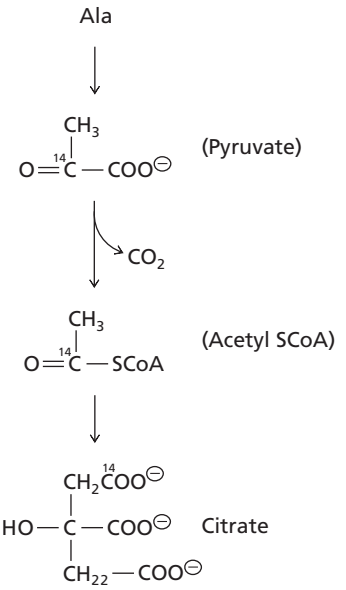
(b) By activating pyruvate carboxylase, acetyl CoA increases the amount of oxaloacetate produced directly from pyruvate. The oxaloacetate can react with the acetyl CoA produced by the degradation of fatty acids. As a result, flux through the citric acid cycle increases to recover the energy stored in the fatty acids.

13. (a)

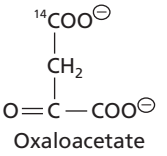


α-Ketoglutarate

(b)



(c)



Oxaloacetate

14. (a) Two molecules of acetyl CoA yield 20 ATP molecules via the citric acid cycle (Figure 13.10) or 6.5 ATP molecules via the glyoxylate cycle (from the oxidation of two molecules of NADH and one molecule of QH_2; Reaction 13.22).

(b) The primary function of the citric acid cycle is to oxidize acetyl CoA to provide the reduced coenzymes necessary for the generation of energy-rich molecules such as ATP. The primary function of the glyoxylate cycle is not to produce ATP, but to convert acetyl groups to four-carbon molecules that can be used to produce glucose.

15. The protein that controls the activity of isocitrate dehydrogenase in *E. coli* is a bifunctional enzyme with kinase and phosphatase activities in the same protein molecule. The kinase activity phosphorylates isocitrate dehydrogenase to inhibit the activity of isocitrate dehydrogenase, and the phosphatase activity dephosphorylates isocitrate dehydrogenase to activate isocitrate dehydrogenase. When concentrations of glycolytic and citric acid cycle intermediates are high, isocitrate dehydrogenase is not phosphorylated and is active. When phosphorylation decreases the activity of isocitrate dehydrogenase, isocitrate is diverted to the glyoxylate cycle.

Chapter 14 Electron Transport and Oxidative Phosphorylation

1. The formula for calculating protonmotive force is

$$\Delta G = F\,\Delta\psi - 2.303\,RT\,\Delta\ \text{pH}$$

If $G = -21{,}000$ kJ and $\Delta\psi = -0.15$ V, then at 25°C

$$-21{,}200 = (96485 \times -0.15) - 2.303(8.315 \times 298)\,\Delta\text{pH}$$

$$5707\ \Delta\ \text{pH} = 6727$$

$$\Delta\ \text{pH} = 1.2$$

Since the outside pH is 6.35 and the inside is negative (higher pH), then the cytoplasmic pH is 6.35 + 1.2 = 7.55.

2. The reduction potential of an iron atom in a heme group depends on the surrounding protein environment, which differs for each cytochrome. The differences in reduction potentials allow electrons to pass through a series of cytochromes.

3. Refer to Figure 14.6.
(a) Complex III. The absence of cytochrome *c* prevents further electron flow.
(b) No reaction occurs since Complex I, which accepts electrons from NADH, is missing.
(c) O_2
(d) Cytochrome *c*. The absence of Complex IV prevents further electron flow.

4. UCP-2 leaks protons back into the mitochondria, thereby decreasing the protonmotive force. The metabolism of foodstuffs provides the energy for electron transport, which in turn creates the protonmotive gradient used to produce ATP. An increase in UCP-2 levels would make the tissue less metabolically efficient (i.e., less ATP would be produced per gram of foodstuff metabolized). As a result, more carbohydrates, fats, and proteins would have to be metabolized in order to satisfy the basic metabolic needs, and this could "burn off" more calories and potentially cause weight loss.

5. (a) Demerol interacts with Complex I and prevents electron transfer from NADH to Q. The concentration of NADH increases since it cannot be reoxidized to $NAD^{\oplus}$. The concentration of Q increases since electrons from QH_2 are transferred to O_2 but Q is not reduced back to QH_2.

(b) Myxothiazole inhibits electron transfer from QH_2 to cytochrome c_1 and from QH_2 (via $\cdot Q^-$) to cytochrome b_{566} in Complex III (Figure 14.14). The oxidized forms of both cytochromes predominate since Fe^{3+} cannot be reduced by electrons from QH_2.

6. (a) Oxygen (O_2) must bind to the Fe^{3+} of cytochrome a_3 in order to accept electrons (Figure 14.19), and it is prevented from doing so by the binding of $CN^{\ominus}$ to the iron atom.
(b) The methemoglobin (Fe^{3+}) generated from nitrite treatment competes with cytochrome a_3 for the $CN^{\ominus}$ ions. This competition effectively lowers the concentration of cyanide available to inhibit cytochrome a_3 in Complex IV, and decreases the inhibition of the electron transport chains in the presence of $CN^{\ominus}$.

7. A substrate is usually oxidized by a compound with a more positive reduction potential. Since $E^{\circ\prime}$ for the fatty acid is close to $E^{\circ\prime}$ for FAD in Complex II (0.0 V, as shown in Table 14.1), electron transfer from the fatty acid to FAD is energetically favorable.

$$\Delta E^{\circ\prime} = 0.0\ \text{V} - (-0.05\ \text{V}) = +0.05\ \text{V}$$

$$\Delta G^{\circ\prime} = -nF\Delta E^{\circ\prime}$$

$$\Delta G^{\circ\prime} = -(2)(96.48\ \text{kJ V}^{-1})(0.05\ \text{V}) = -9.6\ \text{kJ mol}^{-1}$$

Since $E^{\circ\prime}$ for NADH in Complex I is −0.32 V, the transfer of electrons from the fatty acid to NADH is unfavorable.

$$\Delta E^{\circ\prime} = -0.32\ \text{V} - (-0.05\ \text{V}) = -0.27\ \text{V}$$

$$\Delta G^{\circ\prime} = -(2)(96.48\ \text{kJ V}^{-1}\ \text{mol}^{-1})(-0.27\ \text{V}) = 52\ \text{kJ mol}^{-1}$$

8. (a) 10 protons; 2.5 ATP; P : O = 2.5.
(b) 6 protons; 1.5 ATP; P : O = 1.5.
(c) 2 protons; 0.5 ATP; P : O = 0.5.

9. (a) The inner mitochondrial membrane has a net positive charge on the cytosolic side (outside). The exchange of one ATP^{4-} transferred out for one ADP^{3-} transferred in yields a

net movement of one negative charge from the inner matrix side to the positive cytosolic side. The membrane potential thereby assures that outward transport of a negatively charged ATP is favored by the outside positive charge.

(b) Yes. The electrochemical potential with a net positive charge outside the membrane is a result of proton pumping, which is driven by the electron transport chain. This in turn requires oxidation of metabolites to generate NADH and QH_2 as electron donors.

10. ATP synthesis is normally associated with electron transport. Unless ADP can continue to be translocated into the mitochondrial matrix for the ATP synthesis reaction ($ADP + P_i \rightarrow ATP$), ATP synthesis will not occur and the proton gradient will not be dissipated. Electron transport will be inhibited as the proton concentration increases in the intermembrane space.

11. (a) $\Delta G = F\Delta\Psi - 2.303\ RT\ \Delta$ pH (Equation 14.6)

$\Delta G = ((96485)(-0.18)) - ((2.303)(8.315)(0.7))$

$\Delta G = -17367 - 3995$

$\Delta G = -2136 = 21\ \text{kJ mol}^{-1}$

(b) $\Delta G_{\text{total}} = 21.36\ \text{kJ mol}^{-1}$

Charge gradient contribution is $17.367\ \text{kJ mol}^{-1}$, or $17.367 \div 21.36 \times 100 = 81.3\%$

pH gradient contribution is $3.995\ \text{kJ mol}^{-1}$, or $3.995 \div 21.36 \times 10 = 18.7\%$

12. (a) In the malate-aspartate shuttle, the reduction of oxaloacetate in the cytosol consumes a proton that is released in the matrix by the oxidation of malate (Figure 14.27). Therefore, one fewer proton is contributed to the proton concentration gradient for every cytosolic NADH oxidized (9 versus 10 for mitochondrial NADH). The ATP yield from two molecules of cytoplasmic NADH is about 4.5 rather than 5.0.

(b) *Cytoplasmic reactions*

Glucose $\longrightarrow$ 2 Pyruvate	2.0 ATP
2 NADH $\longrightarrow$	4.5 ATP

Mitochondrial reactions

2 Pyruvate $\longrightarrow$ 2 Acetyl CoA + 2 CO_2	2 NADH $\longrightarrow$ 5.0 ATP
2 Acetyl CoA $\longrightarrow$ 4 CO_2	2.0 GTP
	6 NADH $\longrightarrow$ 15.0 ATP
	2 QH_2 $\longrightarrow$ 3.0 ATP
	Total 31.5 ATP

Chapter 15 Photosynthesis

1. Because in photosynthesis there are *two* steps where light energy is absorbed to produce "high energy" electrons, thus PS II transfers 6 $H^{\oplus}$ insead of 10 $H^{\oplus}$ in respiration but PSI produces 2.5 ATP equivalents—the same as respiration.

2. Plant chlorophylls absorb energy in the red region of the spectrum (Figure 15.2). The dragonfish chlorophyll derivatives absorb the red light energy (667 nm), and pass the signals on to the visual pigments in much the same manner that plant antenna chlorophylls and related molecules capture light energy and transfer it to a reaction center where electrons are promoted into excited states for transfer to acceptors of the electron transport chain.

3. (a) Rubisco is the world's most abundant protein and the principal catalyst for photosynthesis, the basic means by which living organisms acquire the carbon necessary for life. Its importance in the process of providing food for all living things can be well justified.

(b) Photorespiration is a process that wastes ribulose 1,5-*bis*phosphate, consumes the NADPH and ATP generated by the light reactions, and can greatly reduce crop yields. As much as 20% to 30% of the carbon fixed in photosynthesis can be lost to photorespiration. This process results from the lack of specificity of Rubisco, which can use O_2 instead of CO_2 (Figure 15.8) to produce phosphoglycolate and 3-phosphoglycerate (Figure 15.18) instead of two triose phosphate molecules. In addition, Rubisco has low catalytic activity ($K_{\text{cat}} \approx 3\ \text{s}^{-1}$). This lack of specificity and low activity earns Rubisco the title of a relatively incompetent, inefficient enzyme.

4. $6\ CO_2 + 6\ H_2S \rightarrow C_6H_{12}O_6 + 3\ O_2 + 6\ S$

$6\ CO_2 + 12\ H^{\oplus} \rightarrow C_6H_{12}O_6 + 3\ O_2$

5. (a) $CO_2 + 2\,H_2S \xrightarrow{\text{Light}} (CH_2O) + H_2O + 2\,S$

$CO_2 + 2\,CH_3CH_2OH \xrightarrow{\text{Light}} (CH_2O) + H_2O + 2\,CH_3CHO$

Ethanol Acetaldehyde

(b) When H_2O is the proton donor, O_2 is the product, but when other proton donors such as H_2S and ethanol are used, oxygen cannot be produced. Most photosynthetic bacteria do not produce O_2 and are obligate anaerobes that are poisoned by O_2.

(c) $CO_2 + 2\,H_2A \xrightarrow{\text{Light}} (CH_2O) + H_2O + 2A$

6. Rubisco is not active in the dark because it requires alkaline conditions. Those conditions only occur when photosynthesis is active so there's nothing (except light) that can be added to the chloroplast suspension in the dark that will activate the calvin cycle.

7. (a) Two H_2O molecules provide the oxygens for one O_2 during the photosynthetic process. A total of four electrons must be removed from two H_2O and passed through an electron transport system to two NADPH. One quantum of light is required to transfer one electron through PSI and one quantum for PSII. Therefore, a total of eight photons will be required to move four electrons through both reaction centers (four photons for PHI and four photons for PHII).

(b) Six NADPH are required for the synthesis of one triose phosphate by the Calvin cycle (Figure 15.21). Therefore, 12 electrons must be transferred through the two reaction centers of the electron transport system and this will require the absorption of 24 hν.

8. (a) Yes. (Refer to the Z-scheme, Figure 15.14). When DCMU blocks electron flow, PSII in the P680* state will not be reoxidized to the $P680^{\oplus}$ state, which is required as an acceptor of electrons from H_2O. If H_2O is not oxidized by $P680^{\oplus}$, then no O_2 will be produced. In the absence of electron flow through the cytochrome *bf* complex, no protons will be translocated across the membrane. Without a pH gradient no photophosphorylation (ATP synthesis) will be possible.

(b) External electron acceptors for PSII will permit P680 to be reoxidized to $P680^{\oplus}$ and will restore O_2 evolution. No electrons will flow through the cytochrome *bf* complex, however, so no photophosphorylation will occur.

9. (a) When the external pH rises to 8.0, the stromal pH also rises quickly, but the luminal pH remains low initially because the thylakoid membrane is relatively impermeable to protons. The pH gradient across the thylakoid membrane drives the production of ATP via proton translocation through chloroplast ATP synthase (Figure 15.16).

(b) Protons are transferred from the lumen to the stroma by ATP synthase, driving ATP synthesis. The pH gradient across the membrane decreases until it is insufficient to drive the phosphorylation of ADP, and ATP synthesis stops.

10. During cyclic electron transport, reduced ferredoxin donates its electrons back to P700 via the cytochrome *bf* complex (Figure 15.11). As these electrons cycle again through photosystem I, the proton concentration gradient generated by the cytochrome *bf* complex drives ATP synthesis. However, no NADPH is produced because there is no *net* flow of electrons from H_2O to ferredoxin. No O_2 is produced because photosystem II, the site of O_2 production, is not involved in cyclic electron transport.

11. The light absorbing complexes, electron transport chain, and chloroplast ATP synthase all reside in the thylakoid membranes, and the structure and interactions of any of these photosynthetic components could be affected by a change in the physical nature of the membrane lipids.

12. The compound is acting as an uncoupler. The electron transfer is occurring without the synthesis of ATP. The compound destroys the proton gradient that is produced through electron transfer.

13. (a) The synthesis of one triose phosphate from CO_2 requires 9 molecules of ATP and 6 molecules of NADPH (Equation 15.5). Since two molecules of triose phosphate can be converted to glucose, glucose synthesis requires 18 molecules of ATP and 12 molecules of NADPH.

(b) Incorporating glucose 1-phosphate into starch requires one ATP equivalent during the conversion of glucose 1-phosphate to ADP-glucose (Figure 15.24), bringing the total requirement to 19 molecules of ATP and 12 molecules of NADPH.

14. Refer to Figure 15.21. (a) C-1. (b) C-3 and C-4. (c) C-1 and C-2. C-1 and C-2 of fructose 6-phosphate are transferred to glyceraldehyde 3-phosphate to form xylulose 5-phosphate. C-3 and C-4 of fructose 6-phosphate become C-1 and C-2 of erythrose 4-phosphate.

15. (a) In the C_4 pathway (Figure 15.29), the pyruvate-phosphate dikinase reaction consumes two ATP equivalents for each CO_2 fixed (since PP_i is hydrolyzed to 2 P_i). Therefore, C_4

plants require 12 more molecules of ATP per molecule of glucose synthesized than C_3 plants require.

(b) Because C_4 plants minimize photorespiration, they are more efficient than C_3 plants in using light energy to fix CO_2 into carbohydrates, even though the chemical reactions for fixing CO_2 in C_4 plants require more ATP.

16. (a) An increase in stromal pH increases the rate of the Calvin cycle in two ways.

(1) An increase in stromal pH increases the activity of ribulose 1,5-*bis*phosphate carboxylase-oxygenase (Rubisco), the central regulatory enzyme of the Calvin cycle, and the activities of fructose 1,6-*bis*phosphatase and sedoheptulose 1,7-*bis*phosphatase. It also increases the activity of phosphoribulokinase. Phosphoribulokinase is inhibited by 3-phosphoglycerate (3PG) in the $3PG^{2-}$ ionization state but not in the $3PG^{3-}$ ionization state, which predominates at higher pH.

(2) An increase in stromal pH also increases the proton gradient that drives the synthesis of ATP in chloroplasts. Since the reactions of the Calvin cycle are driven by ATP, an increase in ATP production increases the rate of the Calvin cycle.

(b) A decrease in the stromal concentration of Mg^{2+} decreases the rate of the Calvin cycle by decreasing the activity of Rubisco, fructose 1,6-*bis*phosphatase, and sedoheptulose 1,7-*bis*phosphatase.

Chapter 16 Lipid Metabolism

1. (a) LDLs are rich in cholesterol and cholesterol esters and transport these lipids to peripheral tissues. Delivery of cholesterol to tissues is moderated by LDL receptors on the cell membranes. When LDL receptors are defective, receptor-mediated uptake of cholesterol does not occur (Section 16.10B). Because cholesterol is not cleared from the blood it accumulates and contributes to the formation of atherosclerotic plaques.

(b) Increased cholesterol levels normally repress transcription of HMG-CoA reductase and stimulate the proteolysis of this enzyme as well. With defective LDL, however, cholesterol synthesis continues in spite of high blood cholesterol levels because the extracellular cholesterol cannot enter the cells to regulate intracellular synthesis.

(c) HDLs remove cholesterol from plasma and cells of nonhepatic tissues and transport it to the liver where it can be converted into bile salts for disposal. In Tangier patients, defective cholesterol-poor HDLs cannot absorb cholesterol, and the normal transport process to the liver is disrupted.

2. (a) Carnitine is required to transport fatty acyl CoA into the mitochondrial matrix for β-oxidation (Figure 16.24). The inhibition of fatty acid transport caused by a deficiency in carnitine diminishes energy production from fats for muscular work. Excess fatty acyl CoA can be converted to triacylglycerols in the muscle cells.

(b) Since carnitine is not required to transport pyruvate, a product of glycolysis, into mitochondria for oxidation, muscle glycogen metabolism is not affected in individuals with a carnitine deficiency.

3. (a) Activation of the C_{12} fatty acid to a fatty acyl CoA consumes 2 ATP. Five rounds of β-oxidation generate 6 acetyl CoA, 5 QH_2 (which yield 7.5 ATP via oxidative phosphorylation), and 5 NADH (which yield 12.5 ATP). Oxidation of the 6 acetyl CoA by the citric acid cycle yields 60 ATP. Therefore, the net yield is 78 ATP equivalents.

(b) Activation of the C_{16} monounsaturated fatty acid to a fatty acyl CoA consumes 2 ATP. Seven rounds of β-oxidation generate 8 acetyl CoA, 6 QH_2 (which yield 9 ATP via oxidative phosphorylation), and 7 NADH (which yield 17.5 ATP). The fatty acid contains a *cis*-β,γ double bond that is converted to a *trans*-α,β double bond, so the acyl-CoA dehydrogenase-catalyzed reaction, which generates QH_2, is bypassed in the fifth round. Oxidation of the 8 acetyl CoA by the citric acid cycle yields 80 ATP. Therefore, the net yield is 104.5 ATP equivalents.

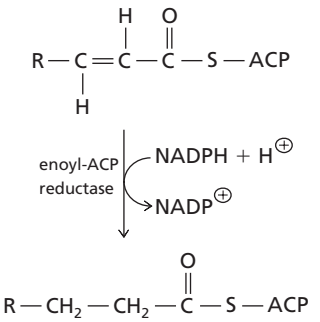
4. When triacylglycerols are ingested in our diets, the hydrolysis of the dietary lipids occurs mainly in the small intestine. Pancreatic lipase catalyzes the hydrolysis at the C-1 and C-3 positions of triacylglycerol, producing free fatty acids and 2-monoacylglycerol. These molecules are transported in bile-salt micelles to the intestine, where they are absorbed by intestinal cells. Within these cells, the fatty acids are converted to fatty acyl CoA molecules, which eventually form a triacylglycerol that is incorporated into chylomicrons for transport to other tissues. If the pancreatic lipase is inhibited, the ingested dietary triglyceride cannot be absorbed. The triglyceride will move through the digestive tract and will be excreted without absorption.

5. (a) Oleate has a *cis*-Δ^9 double bond, so oxidation requires enoyl-CoA isomerase (as in Step 2 of Figure 16.26).
(b) Arachidonate has *cis* double bonds at both odd (Δ^5, Δ^{11}) and even (Δ^8, Δ^{14}) carbons, so oxidation requires both enoyl-CoA isomerase and 2,4-dienoyl-CoA reductase (as in Step 5 of Figure 16.26).
(c) This C_{17} fatty acid contains a *cis* double bond at an even-numbered carbon (Δ^6), so oxidation requires 2,4-dienoyl-CoA reductase. In addition, three enzymes are required to convert the propionyl CoA product into succinyl CoA: propionyl-CoA carboxylase, methylmalonyl-CoA racemase, and methylmalonyl-CoA mutase (Figure 16.25).

6. Even-chain fatty acids are degraded to acetyl CoA, which is not a gluconeogenic precursor. Acetyl CoA cannot be converted directly to pyruvate because for every two carbons of acetyl CoA that enter the citric acid cycle, two carbons in the form of two CO_2 molecules leave as products. The last three carbons of odd-chain fatty acids, on the other hand, yield a molecule of propionyl CoA upon degradation in the fatty acid oxidation cycle. Propionyl CoA can be carboxylated and converted to succinyl CoA in three steps (Figure 16.25). Succinyl CoA can be converted to oxaloacetate by citric acid cycle enzymes, and oxaloacetate can be a gluconeogenic precursor for glucose synthesis.

7. (a) The labeled carbon remains in $H^{14}CO_3^{\ominus}$; none is incorporated into palmitate. Although $H^{14}CO_3^{\ominus}$ is incorporated into malonyl CoA (Figure 16.2), the same carbon is lost as CO_2 during the ketoacyl-ACP synthase reaction in *each* turn of the cycle (Figure 16.5).
(b) All the even-numbered carbons are labeled. Except for the acetyl CoA that becomes C-15 and C-16 of palmitate, the acetyl CoA is converted to malonyl CoA and then to malonyl-ACP before being incorporated into a growing fatty acid chain with the loss of CO_2.

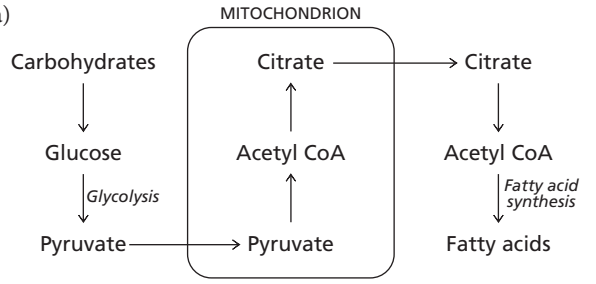
8. (a) Enoyl ACP reductase catalyzes the second reductive step in the fatty acid biosynthesis pathway, converting a *trans*-2,3 enoyl moiety into a saturated acyl chain, and uses NADPH as cofactor.



(b) Fatty acids are essential for membranes in bacteria. If fatty acid synthesis is inhibited, there will be no new membranes and no growth of the bacteria.
(c) The fatty acid synthesis systems are different in animals and bacteria. Animals contain a type I fatty acid synthesis system (FAS I) where the various enzymatic activities are localized to individual domains in a large, multifunctional enzyme. In bacteria, each reaction in fatty acid synthesis is catalyzed by a separate monofunctional enzyme. Understanding some of the differences in these two systems, would allow for the design of specific inhibitors of the bacterial FAS II.

9. Eating stimulates the production of acetyl CoA from the metabolism of carbohydrates (glycolysis and pyruvate dehydrogenase) and fats (FA oxidation). Normally, increased acetyl CoA results in the elevation of malonyl CoA levels (acetyl CoA carboxylase reaction, Figure 16.2), which may act to inhibit appetite. By blocking fatty acid synthase enzyme, C75 prevents the removal of malonyl CoA for the synthesis of fatty acids, thereby elevating the levels of malonyl-CoA and further suppressing appetite.

10. (a)



(b) The NADH generated by glycolysis can be transformed into NADPH by a variety of different reactions and pathways.

11. (a) Plentiful citrate and ATP levels promote fatty acid synthesis. High citrate levels activate ACC by preferential binding and stabilization of the active dephosphorylated filamentous form. On the other hand, high levels of fatty acyl CoAs indicate that there is no further need for more fatty acid synthesis. Palmitoyl CoA inactivates ACC by preferential binding to the inactive protomeric dephosphorylated form.

(b) Glucagon and epinephrine inhibit fatty acid synthesis by inhibiting the activity of acetyl CoA carboxylase. Both hormones bind to cell receptors and activate cAMP synthesis, which in turn activates protein kinases. Phosphorylation of ACC by protein kinases converts it to the inactive form, thus inhibiting fatty acid synthesis. On the other hand, the active protein kinases catalyze phosphorylation and activation of triacylglycerol lipases that catalyze hydrolysis of triacylglycerols, releasing fatty acids for β-oxidation.

12. (a) An inhibitor of acetyl-CoA acetylase will affect a key regulatory reaction for fatty acid synthesis. The concentration of malonyl CoA, the product of the acetyl-CoA carboxylase-catalyzed reaction, will be decreased in the presence of the inhibitor. The decrease in the concentration of malonyl CoA will relieve the inhibition of carnitine acyltransferase I, which is a key regulatory site for the oxidation of fatty acids. Thus, with an active carrier system, fatty acids will be translocated to the mitochondrial matrix where the reactions of β-oxidation occur. In the presence of an inhibitor of acetyl-CoA carboxylase, fatty acid synthesis will decrease and β-oxidation will increase.

(b) CABI is a structural analog of biotin. Acetyl-CoA carboxylase is a biotin-dependent enzyme. A biotin analog may bind in place of biotin and inhibit the activity of acetyl-CoA carboxylase.

13. The overall reaction for the synthesis of palmitate from acetyl CoA is the sum of two processes: (1) the formation of seven malonyl CoA by the action of acetyl-CoA carboxylase and (2) seven cycles of the fatty acid biosynthetic pathway.

$$7\ \text{Acetyl CoA} + 7\ CO_2 + 7\ \text{ATP} \longrightarrow 7\ \text{Malonyl CoA} + 7\ \text{ADP} + 7\ P_i$$

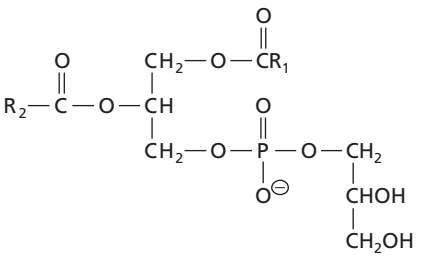
$$\text{Acetyl CoA} + 7\ \text{Malonyl CoA} + 14\ \text{NADPH} + 14\ H^{\oplus} \longrightarrow \text{Palmitate} + 7\ CO_2 + 14\ \text{NADP}^{\oplus} + 8\ \text{HS} - \text{CoA} + 6\ H_2O$$

$$8\ \text{Acetyl CoA} + 7\ \text{ATP} + 14\ \text{NADPH} + 14\ H^{\oplus} \longrightarrow \text{Palmitate} + 7\ \text{ADP} + 7\ P_i + 14\ \text{NADP}^{\oplus} + 8\ \text{HS} - \text{CoA} + 6\ H_2O$$

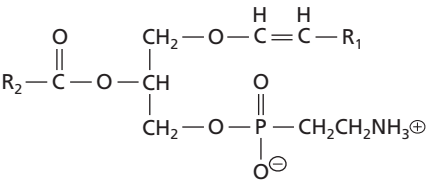
14. (a) Arachidonic acid is a precursor for synthesis of eicosanoids including "local regulators" such as prostaglandins, thromboxanes, and leukotrienes (Figure 16.14). These regulators are involved in mediation of pain, inflammation, and swelling responses resulting from injured tissues.

(b) Both prostaglandins and leukotrienes are derived from arachidonate, which is released from membrane phospholipids by the action of phospholipases. By inhibiting a phospholipase, steroidal drugs block the biosynthesis of *both* prostaglandins and leukotrienes. Aspirin-like drugs block the conversion of arachidonate to prostaglandin precursors by inhibiting cyclooxygenase but do not affect leukotriene synthesis.

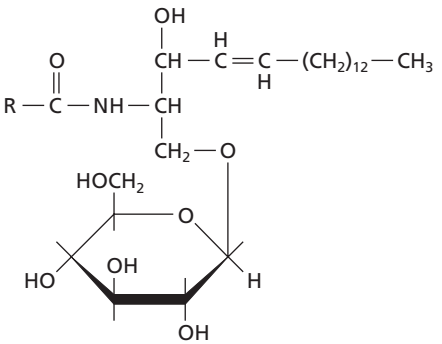
15. (a)



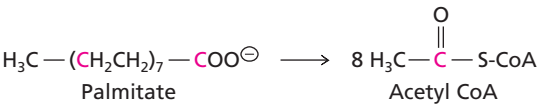
(b)



(c)

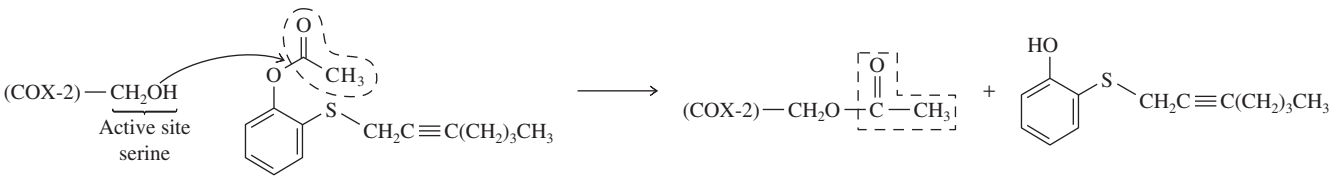


16. Palmitate is converted to eight molecules of acetyl CoA labeled at C-1. Three acetyl CoA molecules are used to synthesize one molecule of mevalonate (Figure 16.17).





17. Both APHS and aspirin transfer an acetyl group to a serine residue on COX enzymes. Since APHS is an irreversible inhibitor, it does not exhibit competitive inhibition kinetics even though it acts at the active site of COX enzymes.



Chapter 17 Amino Acid Metabolism

1. PSII contains the oxygen evolving complex and oxygen is produced during photosynthesis. Since oxygen inhibits nitrogenase, the synthesis of O_2 in hetocysts must be avoided. PSI is retained because it can still generate a light-induced proton gradient by cyclic electron transport and it is not involved in the production of O_2.

2. (a) Glutamate dehydrogenase + glutamine synthetase

$$\begin{aligned} NH_4^{\oplus} + \alpha\text{-Ketoglutarate} + NAD(P)H + H^{\oplus} &\longrightarrow \text{Glutamate} + NAD(P)^{\oplus} + H_2O \\ NH_3 + \text{Glutamate} + ATP &\longrightarrow \text{Glutamine} + ADP + P_i \\ \hline 2\ NH_4^{\oplus} + \alpha\text{-Ketoglutarate} + NAD(P)H + ATP &\longrightarrow \text{Glutamine} + NAD(P)^{\oplus} + ADP + P_i + H_2O \end{aligned}$$

(b) Glutamine synthetase + glutamate synthase

$$\begin{aligned} 2\ NH_3 + 2\ \text{Glutamate} + 2\ ATP &\longrightarrow 2\ \text{Glutamine} + 2\ ADP + 2\ P_i \\ \text{Glutamine} + \alpha\text{-Ketoglutarate} + NAD(P)H + H^{\oplus} &\longrightarrow 2\ \text{Glutamate} + NAD(P)^{\oplus} \\ \hline 2\ NH_3 + \alpha\text{-Ketoglutarate} + NAD(P)H + 2\ ATP + H^{\oplus} &\longrightarrow \text{Glutamine} + NAP(P)^{\oplus} + 2\ ADP + 2\ P_i \end{aligned}$$

The coupled reactions in (b) consume one more ATP molecule than the coupled reactions in (a). Because the K_m of glutamine synthetase for NH_3 is much lower than the K_m of glutamate dehydrogenase for $NH_4^{\oplus}$, the coupled reactions in (b) predominate when $NH_4^{\oplus}$ levels are low. Thus, more energy is spent to assimilate ammonia when its concentration is low.

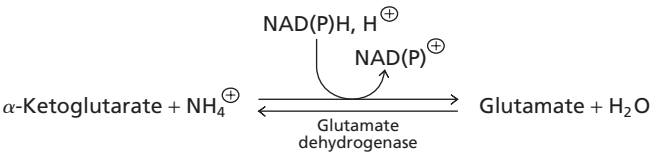
3. The ^{15}N-labeled amino group is transferred from aspartate to α-ketoglutarate, producing glutamate in a reaction catalyzed by aspartate transaminase (Figure 17.10). Since transaminases catalyze near-equilibrium reactions and many transaminases use glutamate as the α-amino group donor, the labeled nitrogen is quickly distributed among the other amino acids that are substrates of glutamate-dependent transaminases.

4. (a) α-Ketoglutarate + Amino acid $\rightleftharpoons$ Glutamate + α-Keto acid

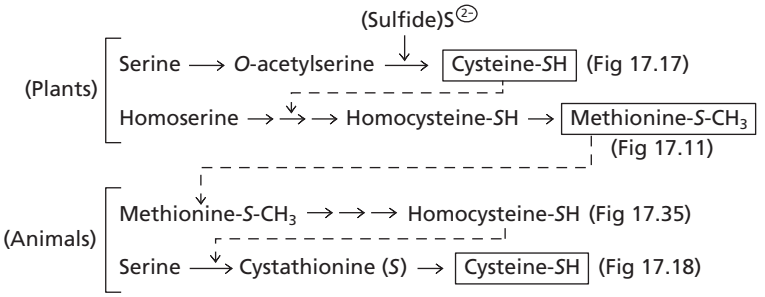
Oxaloacetate + Amino acid $\rightleftharpoons$ Aspartate + α-Keto acid

Pyruvate + Amino acid $\rightleftharpoons$ Alanine + α-Keto acid

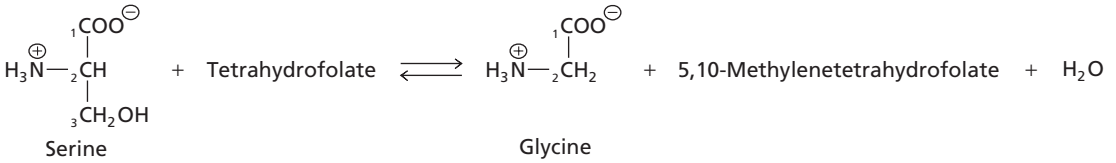
(b)

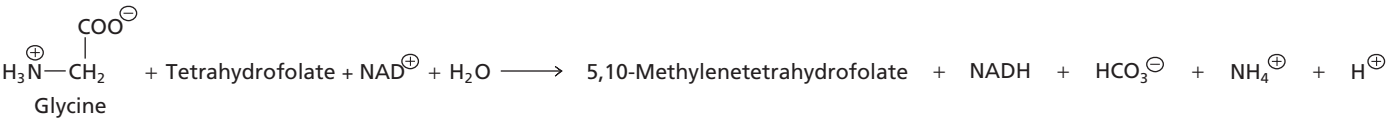


5.



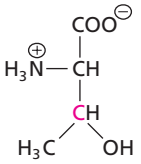
6. (a) C-3 of serine is transferred to tetrahydrofolate during the synthesis of glycine, and C-2 is transferred to tetrahydrofolate when glycine is cleaved to produce ammonia and bicarbonate.



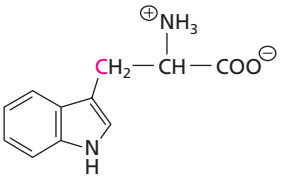


(b) Serine is synthesized from 3-phosphoglycerate (Figure 17.15), an intermediate of glycolysis. C-3 of both 3-phosphoglycerate and serine is derived from either C-1 or C-6 of glucose, and C-2 of both 3-phosphoglycerate and serine is derived from either C-2 or C-5 of glucose.

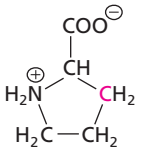
7. (a)



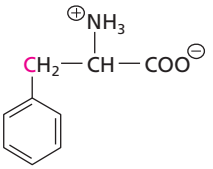
(b)



(c)



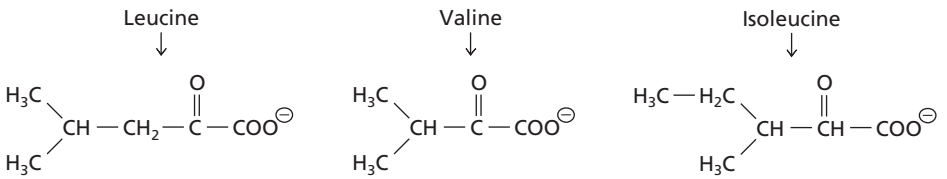
(d)



8. (a) Glutamic acid. PPI inhibits glutamine synthetase.
(b) Histidine biosynthesis pathway (Figure 17.23).

9. Aspartame is a dipeptide consisting of an asparate and a phenylalanine residue joined by a peptide bond. This bond is eventually hydrolyzed inside the cell producing aspartate and phenylalanine. Phenylketonuria patients must avoid any excess phenylalanine.

10. (a)



(b) Lysine degradation pathway. α-Aminoadipate δ-semialdehyde synthase is deficient (Figure 17.39).
(c) Urea cycle. Argininosuccinate synthetase is deficient (Figure 17.43).

11. (a) Alanine (c) Glycine
(b) Aspartate (d) Cysteine

12. The urea cycle does not operate in muscle, so ammonia from the deamination of amino acids cannot be converted to urea. Because high concentrations of ammonia are toxic, ammonia is converted to other products for disposal. In the first pathway, ammonia is incorporated into glutamine by the action of glutamine synthetase (Figure 17.5). Glutamine can then be

transported to the liver or kidneys. The second pathway is the glucose-alanine cycle (Figure 17.45). Pyruvate accepts the amino group of amino acids by transamination, and the alanine produced is transported to the liver where it can be deaminated back to pyruvate. The amino group is used for urea synthesis, and the pyruvate can be converted to glucose.

13. Inhibition of nitric oxide synthase (NOS) can prevent excess amounts of nitric oxide from being produced in cells lining the blood vessels. Nitric oxide causes relaxation of the vessels and in excess amounts can cause reduced blood pressure leading to shock. Thiocitrulline and *S*-methylthocitrulline inhibit NOS because they are unreactive analogs of the NOS reaction product citrulline (Figure 17.25).

14. There are two reasons. Firstly, many of the amino acid biosynthesis pathways aren't found in humans, so there won't be any metabolic diseases of nonexistent essential amino acid pathways. Secondly, the remaining pathways are probably crucial pathways during development so that any defects in these pathways are likely to be lethal. This is the same reasoning that we used to explain the lack of metabolic diseases in the sphingolipid biosynthesis pathways (Box 16.2).

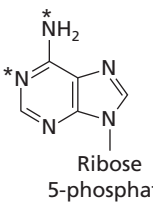
15. The 21st, 22nd, and 23rd amino acids are *N*-formylmethionine, selenocysteine, and pyrrolysine. *N*-formylmethionine and selenocysteine are synthesized during translation on aminoacylated tRNA and not by the standard metabolic pathways covered in this chapter. Pyrrolysine may also be synthesized on aminoacylated tRNA. The precursors are methionine, serine, and lysine.

16. The precursor in the serine biosynthesis pathway is 3-phosphoglycerate. This precursor can be derived from glyceraldehyde-3-phosphate (G3P) in the glycolytic pathway, where the conversion is associated with the *gain* of 1 ATP + 1 NADH. This gain must be subtracted from the total cost of G3P synthesis. Therefore, the cost of making 3-phosphoglycerate is 24 − 3.5 = 20.5 ATP equivalents, assuming that each NADH is equivalent to 2.5 ATPs. (The same cost can be derived from the Calvin cycle pathway.) The serine biosynthesis pathway produces one NADH when 3-phosphoglycerate is oxidized to 3-phosphohydroxypyruvate, so the next cost of making serine is 20.5 − 2.5 = 18 ATP equivalents. This value is identical to the value given in Box 17.3. (Note that the transamination reaction in the serine biosynthesis pathway is cost-free.)

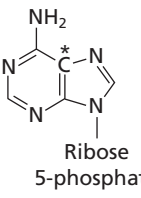
Alanine is made from pyruvate in a simple, cost-free, transamination reaction. The cost of making pyruvate can be estimated from the conversion of 3-phosphoglycerate to pyruvate in the glycolytic pathway. This conversion is associated with a gain of 1 ATP, so the cost of pyruvate is 20.5 − 1 = 19.5 ATP equivalents. Thus, the cost of synthesizing alanine is 19.5 ATP equivalents, or 20 ATP equivalents when rounded to two significant figures. This value is the same as that given in Box 17.3.

Chapter 18 Nucleotide Metabolism

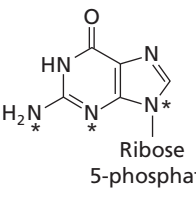
1. (a)



(b)



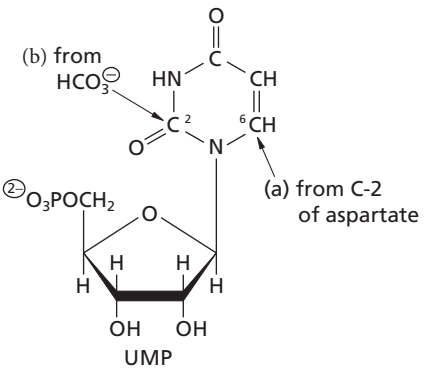
(c)



See Figure 18.10 for the reactions in the pathway of UMP synthesis.

(d) Labeled C-2 from aspartate, which is incorporated into carbamoyl aspartate, appears at C-6 of the uracil of UMP.

(e) The labeled carbon from $HCO_3^{\ominus}$, which is incorporated into carbamoyl phosphate, appears at C-2 of the pyrimidine ring of UMP.



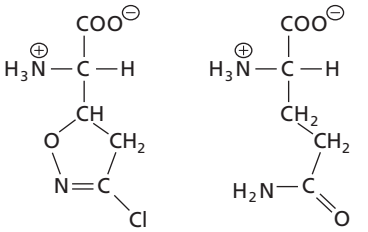
2. Seven ATP equivalents are required. One ATP is cleaved to AMP when PRPP is synthesized (Figure 18.3). The pyrophosphoryl group of PRPP is released in step 1 of the IMP biosynthetic pathway and subsequently hydrolyzed to 2 P_i (Figure 18.5), accounting for the second ATP equivalent. Five ATP molecules are consumed in steps 2, 4, 5, 6, and 7.

3. Purines: Reaction 3: GAR transformylase 10-formyl-THF, C-8 position.

Reaction 9: AICAR transformylase, 10-formyl-THF, C-2 position.

Pyrimidines: Thymidylate synthase, 5,10-methylene-THF, 5-CH_3 of thymidylate.

4. (a)



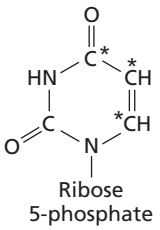
Acivicin Glutamine

(b) Acivicin inhibits glutamine-PRPP amidotransferase, the first enzyme in the purine biosynthetic pathway, so PRPP accumulates.

(c) Acivicin inhibits the carbamoyl phosphate synthetase II activity of dihydroorotate synthase that catalyzes the first step in the pyrimidine biosynthetic pathway.

5. (a) When β-alanine is used instead of aspartate, no decarboxylation reaction (step 6 of the *E. coli* pathway) would be required.

(b)



Ribose 5-phosphate

6. (a) dUMP + $NH_4^{\oplus}$

(b) Synthesis of DNA requires certain ratios of A, T, G and C. If dTTP levels are higher than necessary, dTTP will act to decrease its own synthesis pathway by inhibiting the conversion of dCMP to dUMP by dCMP deaminase. dUMP is the precursor to dTMP (thymidylate synthase, Figure 18.16), and the subsequent conversion to dTDP and dTTP (needed for DNA synthesis). On the other hand, if dCTP levels are high, activation of dCMP deaminase will lead to an increased conversion of dCMP to dUMP and this diverts any dCMP that might have been converted to more dCTP by phosphorylation (Figure 18.20).

7. Four ATP equivalents are required. One ATP equivalent is required for the synthesis of PRPP from ribose 5-phosphate (Figure 18.3). Carbamoyl phosphate synthesis requires 2 ATP (Figure 18.10, step 1). One ATP equivalent is consumed in step 5, when PP_i is hydrolyzed to 2 P_i.

8. In the absence of adenosine deaminase, adenosine and deoxyadenosine are not degraded via inosine and hypoxanthine to uric acid (Figure 18.19 and 18.21). This leads to an increase in the concentration of deoxyadenosine, which can be converted to dATP. High concentrations of dATP inhibit ribonucleotide reductase (Table 18.1). The inhibition of ribonucleotide reductase results in decreased production of all deoxynucleotides and therefore inhibits DNA synthesis.

9. Glutamine-PRPP amidotransferase is the first enzyme and the principal site of regulation in the *de novo* pathway to IMP (Figure 18.5). In humans, PRPP is both a substrate and a positive effector of this enzyme. An increase in the cellular levels of PRPP due to increased PRPP synthetase activity will therefore enhance the activity of the amidotransferase. This will result in an increased synthesis of IMP and other purine nucleosides and nucleotides. Overproduction of purine nucleotides and subsequent degradation can lead to elevated uric acid levels characteristic of gout.

10. (a) ATP (b) ATP (c) ATP (d) GTP (e) UTP (f) GTP (g) C TP (h) UTP (i) ATP (j) IMP (k) IMP

11. Purines and pyrimidines are not significant sources of energy. The carbon atoms of fatty acids and carbohydrates can be oxidized to yield ATP, but there are no comparable energy-yielding pathways for nitrogen-containing purines and pyrimidines. However, the NADH produced when hypoxanthine is converted to uric acid may indirectly generate ATP via

oxidative phosphorylation. The degradation of uracil and thymine yields acetyl CoA and succinyl CoA, respectively, which can be metabolized via the citric acid cycle to generate ATP.

12. The sugar D-ribose exists as an equilibrium mixture of α-D-ribopyranose, α-D-ribofuranose, β-D-ribopyranose, and β-D-ribofuranose. These forms freely interconvert with each through the open-chain form (Section 8.2).

13. Xanthine is 2,6-dioxopurine; hypoxanthine is 6-oxopurine; orotate is 2,4-dioxo-6-carboxylpyrimidine.

14. SAICAR synthetase + adenylosuccinate lyase in the IMP biosynthesis pathway (Figure 18.5) and argininosuccinate synthetase + argininosuccinate lyase in the arginine biosynthesis pathway (urea cycle: Figure 17.43).

Chapter 19 Nucleic Acids

1. In the α helix, hydrogen bonds form between the carbonyl oxygen of one residue and the amine hydrogen four residues, or one turn, away. These hydrogen bonds between atoms in the backbone are roughly parallel to the axis of the helix. The amino acid side chains, which point away from the backbone, do not participate in intrahelical hydrogen bonding. In double-stranded DNA, the sugar-phosphate backbone is not involved in hydrogen bonding. Instead, two or three hydrogen bonds, which are roughly perpendicular to the helix axis, form between complementary bases in opposite strands.

In the α helix, the individual hydrogen bonds are weak, but the cumulative forces of these bonds stabilize the helical structure, especially within the hydrophobic interior of a protein where water does not compete for hydrogen bonding. In DNA, the principal role of hydrogen bonding is to allow each strand to act as a template for the other. Although the hydrogen bonds between complementary bases help stabilize the helix, stacking interactions between base pairs in the hydrophobic interior make a greater contribution to helix stability.

2. If 58% of the residues are (G + C), 42% of the residues must be (A + T). Since every A pairs with a T on the opposite strand, the number of adenine residues equals the number of thymine residues. Therefore, 21%, or 420, of the residues are thymine ($2000 \times 0.21 = 420$).

3. (a) The base compositions of complementary strands of DNA are usually quite different. For example, if one strand is poly dA (100% A), the other strand must be poly dT (100% T). However, since the two strands are complementary, the amount of (A + T) must be the same for each strand, and the amount of (G + C) must be the same for each strand.
(b) (A + G) = (T + C). Complementarity dictates that for every purine (A or G) on one strand, there must be a pyrimidine (T or C) on the complementary strand.

4. Since the DNA strands are anti-parallel, the complementary strand runs in the opposite direction. The sequence of the double-stranded DNA is

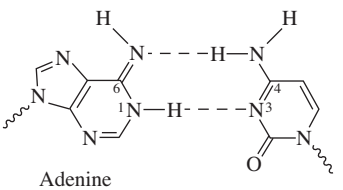
ATCGCGTAACATGGATTCGG
TAGCGCATTGTACCTAAGCC

By convention, DNA sequences are written in the 5′ → 3′ direction. Therefore, the sequence of the complementary strand is

CCGAATCCATGTTACGCGAT

5. The stability of the single-stranded helix is largely due to stacking interactions between adjacent purines. Hydrophobic effects also contribute, since the stacked bases form an environment that is partially shielded from water molecules.

6.



Adenine (imino form) Cytosine

7. There will be two discrete melting points separated by a plateau. When the extra strand of poly dT is released, the absorbance of the solution at 260 nm will increase as the stacked bases leave the largely hydrophobic interior of the triple helix. A second increase in the absorbance occurs when the remaining two DNA strands denature.

8. The sequence is

5′ ACGCACGUAUAUGUACUUAUACGUGGCU 3′

The underlined sequences are palindromic.

9. The main products will be a mixture of mononucleotides *and* pieces of single-stranded DNA approximately 500 bp in length. A piece of DNA with an enzyme molecule bound at each end will be degraded until the two strands can no longer base-pair; at that point the single strands cease to be a substrate for the enzyme.

10. In the 30 nm fiber, DNA is packaged in nucleosomes, each containing about 200 bp of DNA; therefore, the DNA in a nucleosome has a molecular weight of 130,000 (200 × 650 = 130,000). Assuming there is one molecule of histone H1 per nucleosome, the molecular weight of the protein component of the nucleosome would be 129,800.

Histone H1	21,000
Histone H2A (×2)	28,000
Histone H2 B (×2)	27,600
Histone H3 (×2)	30,600
Histone H4 (×2)	22,600
Total	129,800

Thus, the ratio by weight of protein to DNA is 129,800:130,000, or approximately 1:1.

11. Nucleosomes are composed of histones plus 200 base pairs of DNA. Since you inherited half your chromosomes from your mother, the oocyte contained

$$(3.2 \times 10^9 \text{ bp}) \times \frac{1 \text{ nucleosome}}{200 \text{ bp}} = 8 \times 10^6 \text{ nucleosomes}$$

(You inherited no nucleosomes from your father since nucleosomes are replaced by small, positively charged polypeptides during spermatogenesis.)

12. (a) pdApdGpdT + pdC
(b) pdAp + dGpdTpdC
(c) pdA + pdGpdTpdC

13. Since the supercoiled plasmid DNA is in equilibrium with relaxed DNA containing short unwound regions, the *Aspergillus* enzyme will slowly convert the DNA into nicked circles. Eventually the enzyme will convert the relaxed circles into unit-length linear fragments of double-stranded DNA.

14. Yes. The sugar-phosphate backbone in both RNA and DNA contains phosphodiester bonds that link the sugar residues.

15. pppApCpUpCpApUpApGp + CpUpApUpGp + ApGp + U

16. Bacteriophages have evolved several mechanisms to protect their DNA from restriction endonucleases. In general, bacteriophage DNA contains few restriction sites. Restriction endonuclease recognition sites are strongly selected and any mutations that alter these sites will be favored. In addition, restriction sites are often methylated, as in the bacterial chromosome. This is presumably due to a fortuitous event in the distant past when the phage DNA became methylated before it could be cleaved.

Some bacteriophages incorporate modified nucleotides into their DNA. The modified nucleotides (e.g., 5-hydroxymethylcytosine in bacteriophage T4) are not recognized by restriction endonucleases.

Phage genomes may also encode an enzyme that inactivates restriction endonucleases, or they may encode proteins that bind to restriction sites to prevent cleavage.

17. (a) The probability can be estimated from the probability of each nucleotide in the *Hind* III restriction site. (G = C = 0.18 and A = T = 0.32)
For the sequence AAGCTT there will be, on average, one *Hind*III site every
1/(0.32)(0.32)(0.18)(0.18)(0.32)(0.32) = 2943 bp
Thus, in a 100 Mb genome there will be, on average,
100,000/2943 = 33,070 sites
(b) 24,414

18. Although the recognition sites for *Bgl*II and *Bam*H1 differ, the enzymes produce fragments with identical sticky ends. These fragments can be ligated as easily as fragments produced by a single enzyme.

*Bgl*II
~~~A GATCT~~~
~~~TCTAG A~~~

*Bam*HI
~~~G GATCC~~~
~~~CCTAG G~~~

19. Restriction enzymes present in normal host cells might cleave newly introduced recombinant molecules, making it impossible to clone certain fragments of DNA. Using a host strain that does not make restriction endonucleases avoids this problem.

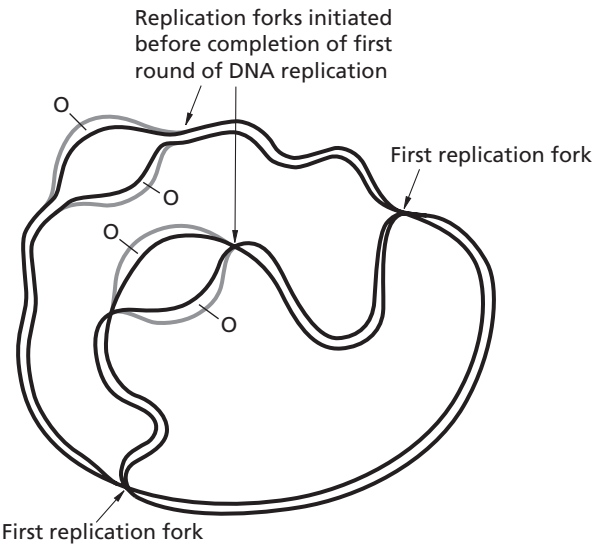
A mutation in RecA reduces recombination, thereby preventing the rearrangement of recombinant DNA molecules during propagation in the host cells. Rearrangement is often a problem, particularly when the cloned fragment of DNA contains repetitive sequences that can serve as sites for homologous recombination.

Chapter 20 DNA Replication, Repair, and Recombination

1. (a) Two replication forks form at the origin of replication and move in opposite directions until they meet at a point opposite the origin. Therefore, each replisome replicates half the genome (2.6×10^6 base pairs). The time required to replicate the entire chromosome is

$$\frac{2.6 \times 10^6 \text{ base pairs}}{1\,000 \text{ base pairs s}^{-1}} = 2600 \text{ s} = 43 \text{ min and } 20 \text{ s}$$

(b) Although there is only one origin (O), replication can be reinitiated before the previous replication forks have reached the termination site. Thus, the chromosome can contain more than two replication forks. Replication of a single chromosome still requires approximately 43 minutes, but completed copies of each chromosome can appear at shorter intervals, depending on the rate of initiation.



2. T4 DNA polymerase should be an early gene product because it is required for replication of the viral genome.

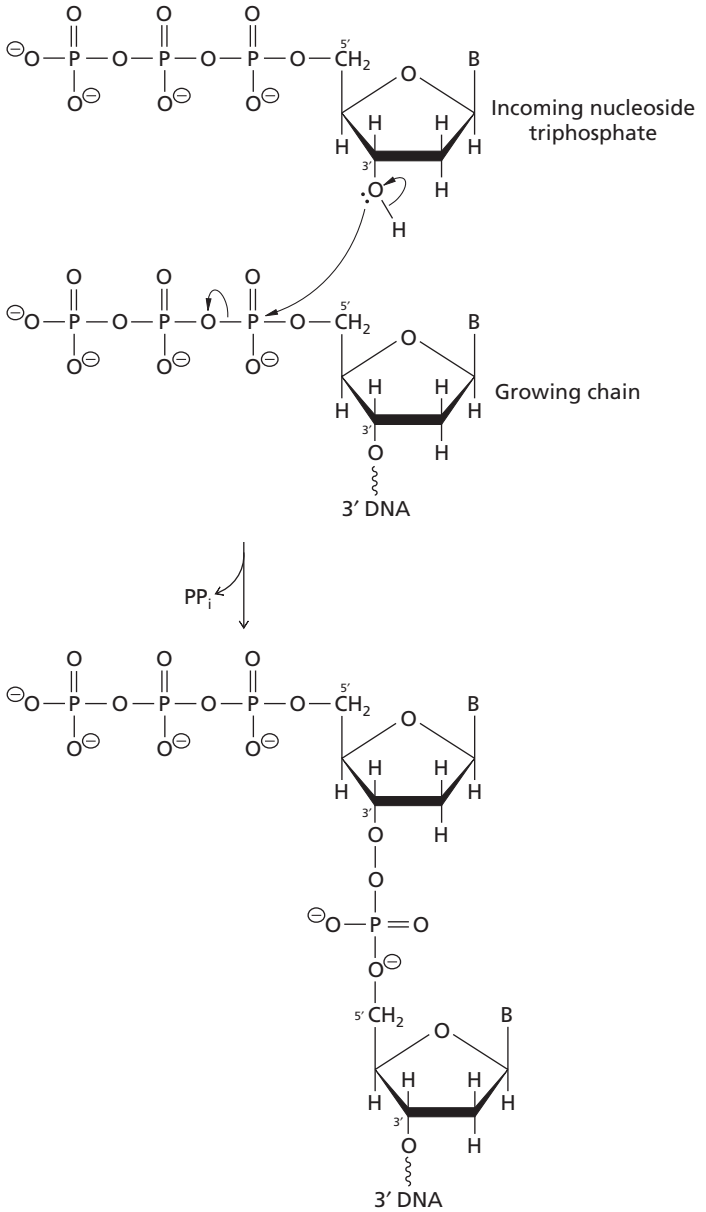
3. (a) The single-stranded DNA template used for DNA synthesis *in vitro* can form secondary structures such as hairpins. SSB prevents the formation of double-stranded structure by binding to the single-stranded template. SSB thus renders the DNA a better substrate for DNA polymerase.

(b) The yield of DNA *in vitro* is improved at higher temperatures because formation of secondary structure in the template is less likely. A temperature of 65°C is high enough to prevent formation of secondary structure but not high enough to denature the newly synthesized DNA. DNA polymerases from bacteria that grow at high temperatures are used because they are active at 65°C, a temperature at which DNA polymerases from other bacteria would be inactive.

4. Extremely accurate DNA replication requires a proofreading mechanism to remove errors introduced during the polymerization reaction. Synthesis of an RNA primer by a primase, which does not have proofreading activity, is more error prone than DNA synthesis. However,

because the primer is RNA, it can be removed by the 5′ → 3′ exonuclease activity of DNA polymerase I and replaced with accurately synthesized DNA when Okazaki fragments are joined. If the primer were composed of DNA made by a primase without proofreading activity, it would not be removed by DNA polymerase I and the error rate of DNA replication would be higher at sites of primer synthesis.

5. (a) In the hypothetical nucleotidyl group transfer reaction, the nucleophilic 3′-hydroxyl group of the *incoming* nucleotide would attack the triphosphate group of the growing chain. Pyrophosphate would be released when a new phosphodiester linkage was formed.



(b) If the hypothetical enzyme had 5′ → 3′ proofreading activity, removal of a mismatched nucleotide would leave a 5′-monophosphate group at the end of the growing chain. Further DNA synthesis, which would require a terminal triphosphate group, could not occur.

6. Topisomerase II or gyrase relieves supercoiling ahead of and behind the replication fork. If this enzyme is inhibited, the unwinding of the parental DNA cannot occur. Therefore, the DNA of the *E. coli* cannot be replicated.

7. (a) Assume that the genome is one large linear molecule of DNA and that the origin of replication is at the midpoint of this chromosome. Since the replication forks move in opposite directions, 60 base pairs can be replicated per second. The time required to replicate the entire genome would be

$$\frac{1.65 \times 10^8 \text{ base pairs}}{60 \text{ base pairs s}^{-1}} = 2.75 \times 10^6 \text{ s} = 764 \text{ h} = 32 \text{ days}$$

(b) Assuming that the 2000 bidirectional origins are equally spaced along the DNA molecule and that initiation occurs simultaneously at all origins, the rate would be $2000 \times 2 \times 30$ base pairs per second, or 1.2×10^5 base pairs per second. The time required to replicate the entire genome would be

$$\frac{1.65 \times 10^8 \text{ base pairs}}{1.2 \times 10^5 \text{ base pairs s}^{-1}} = 1375 \text{ s} = 23 \text{ min}$$

(c) Assume that the origins are equally spaced and that initiation at all origins is simultaneous. The required rate of replication is

$$\frac{1.65 \times 10^8 \text{ base pairs}}{300 \text{ s}} = 5.5 \times 10^5 \text{ base pairs s}^{-1}$$

Bidirectional replication from each fork proceeds at an overall rate of 60 base pairs per second. The minimum number of origins would be

$$\frac{5.5 \times 10^5 \text{ base pairs s}^{-1}}{60 \text{ base pairs s}^{-1} \text{ origin}^{-1}} = 9170 \text{ origins}$$

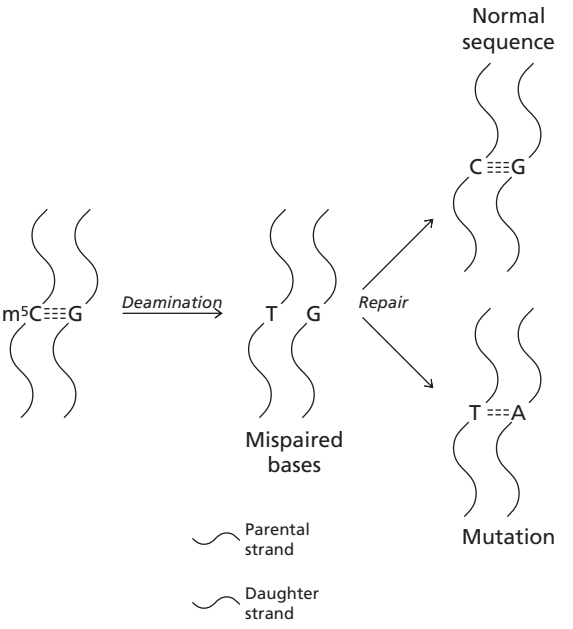
8. The modified G can no longer form a productive Watson-Crick base pair with C but can now base-pair with T. Therefore, one of the daughter strands of DNA will contain a T across from the modified base. After further rounds of replication, the T will base-pair with A and what was originally a G/C base pair will have mutated into an A/T base pair.

9. Ultraviolet light can damage DNA by causing dimerization of thymidylate residues. One mechanism for repairing thymine dimers is enzymatic photoreactivation, catalyzed by DNA photolyase. This enzyme uses energy from visible light to cleave the dimer and repair the DNA. Thus, cells that are exposed to visible light following ultraviolet irradiation are better able to repair DNA than cells kept in the dark.

10. (a) DNA from a *dut*$^-$ strain will appear normal because the Ung enzyme will remove any uracil that gets incorporated.
(b) DNA from a *dut*$^-$, *ung*$^-$ strain will contain dU residues in the place of some dT residues.

11. The DNA repair enzyme uracil *N*-glycosylase removes uracil formed by the hydrolytic deamination of cytosine. Because the enzyme does not recognize thymine or the other three bases normally found in DNA, it cannot repair the damage when 5-methylcytosine is deaminated to thymine.

12. High mutation rates occur at methylcytosine-containing regions because the product of deamination of 5-methylcytosine is thymine, which cannot be recognized as abnormal. When the mismatched T/G base pair that results from deamination of methylcytosine is repaired, the repair enzymes may delete either the incorrect thymine or the correct guanine. When the guanine is replaced by adenine, the resulting A/T base pair is a mutation.



13. Proofreading during replication results in excision of 99% of misincorporated nucleotides, thus reducing the overall error rate to 10^{-7}. Of those errors that escape the proofreading step, a further 99% are corrected by repair enzymes. The overall mutation rate is therefore 10^{-9}.

14. Yes. The *E. coli* enzyme DNA ligase is required to seal the nicks left in the DNA strands following DNA repair. This enzyme has a strict requirement for $NAD^{\oplus}$.

15. The dimers can be removed by excision repair. UvrABC endonuclease removes a 12–13 residue segment containing the pyrimidine dimer. The DNA oligonucleotide is removed with the help of a helicase. The gap is filled by the action of DNA polymerase I, and the nick sealed by the action of DNA ligase. The dimers can also be repaired through direct repair. DNA photolyase binds to the distorted double helix at the site of the dimer. As the DNA–enzyme complex absorbs light, the dimerization reaction is reversed.

16. The repair enzymes need an undamaged template in order to repair mutations in DNA. If both strands of the DNA molecule have been damaged, there is not a template to use for repair.

17. The proteins that catalyze strand exchange recognize regions of high sequence similarity and promote formation of a triple-stranded intermediate in which the invading strand base-pairs with a complementary strand. This pairing would not be possible if the sequences of the two DNA molecules were different.

18. DNA polymerase III is a component of the replisome that synthesizes the leading strand and the lagging strand during replication of the *E. coli* chromosome. DNA polymerase I is required to remove the short RNA primers on the lagging strand.

Chapter 21 Transcription and RNA Processing

1. (a) Since the rate of transcription is 70 nucleotides per second and each transcription complex covers 70 base pairs of DNA, an RNA polymerase completes a transcript and leaves the DNA template each second (assuming that the complexes are densely packed). Therefore, when the gene is loaded with transcription complexes, 60 molecules of RNA are produced per minute.

(b) Since each transcription complex covers 70 base pairs, the maximum number of complexes is

$$\frac{6000 \text{ base pairs}}{70 \text{ base pairs transcription complex}} = 86 \text{ transcription complexes}$$

2. (a) Since the average *E. coli* gene is 1 kb (1000 bp) long, 4000 genes account for 4000 kb of DNA. The percentage of DNA that is not transcribed is

$$\frac{500 \text{ kb}}{4600 \text{ kb}} \times 100\% = 10.9\%$$

Most of the nontranscribed DNA consists of promoters and regions that regulate transcription initiation.

(b) Since the gene products in mammals and bacteria are similar in size, the amount of DNA in the exons of a typical mammalian gene must also be 1000 bp. The total amount of DNA in exons is

$$5 \times 10^4 \text{ genes} \times 1.0 \text{ kb gene}^{-1} = 5 \times 10^4 \text{ kb}$$

This DNA represents about 1.7% of the mammalian genome.

$$\frac{5 \times 10^4 \text{ kb}}{3 \times 10^6 \text{ kb}} \times 100\% = 1.7\%$$

The remaining 97.5% of DNA consists of introns and other sequences.

3. No. It is extremely unlikely that the eukaryotic gene's promoter will contain the correct sequences in the correct location to permit accurate initiation by the prokaryotic RNA polymerases. Likewise, it is extremely unlikely that the prokaryotic gene's promoter will contain the correct sequence in the correct location to permit accurate initiation by RNA polymerase II.

4. No. A typical eukaryotic triose phosphate isomerase gene contains introns. The prokaryotic cell contains no spliceosomes and therefore will not be able to correctly process the primary transcript. Therefore, translation of the RNA will yield an aberrant protein fragment.

5. (a) In the presence of both lactose and glucose, the *lac* operon is transcribed at a low level because *lac* repressor forms a complex with allolactose (an isomer of lactose). Because the

allolactose–repressor complex cannot bind to the promoter region of the *lac* operon, the repressor does not prevent initiation of transcription.

(b) In the absence of lactose, no allolactose is formed. Thus, *lac* repressor binds near the *lac* operon promoter and prevents transcription.

(c) When lactose is the sole carbon source, the *lac* operon is transcribed at the maximum rate. In the presence of allolactose, transcription is allowed since *lac* repressor does not bind to the promoter region of the *lac* operon. Also, in the absence of glucose, the transcription rate increases because cAMP production increases, making more CRP-cAMP available to bind to the promoter region of the *lac* operon. The absence of the repressor and the enhancement of transcription initiation by CRP-cAMP allow the cell to synthesize the quantities of enzymes required to support growth when lactose is the only carbon source.

6. Since the wild-type *lac* promoter is relatively weak, maximal transcription requires the activator CRP. The UV5 mutations alters the −10 region such that it now resembles the consensus −10 sequence, making it a much stronger promoter. In the absence of the *lac* repressor, the promoter is independent of CRP.

7. ^{32}P appears only at the 5′ end of mRNA molecules that have ATP as the first residue. It does not appear in any other residues because pyrophosphate, which includes the β-phosphoryl group, is released when nucleoside triphosphates are added to the 3′ end of a growing RNA chain (Figure 21.3).

When the 5′ end of mRNA is capped, only the γ-phosphoryl group of the initial residue is removed when the cap forms. The β-phosphoryl group, which contains the label, is retained and receives the GMP group from GTP (Figure 21.26).

8. The lack of proofreading activity in RNA polymerase makes the error rate of transcription greater than the error rate of DNA replication. However, the defective RNA molecules produced are not likely to affect cell viability because most copies of RNA synthesized from a given gene are normal. In the case of defective mRNA, the number of defective proteins is only a small percentage of the total number of proteins synthesized. Also, mistakes made during transcription are quickly eliminated since most mRNA molecules have a short half-life.

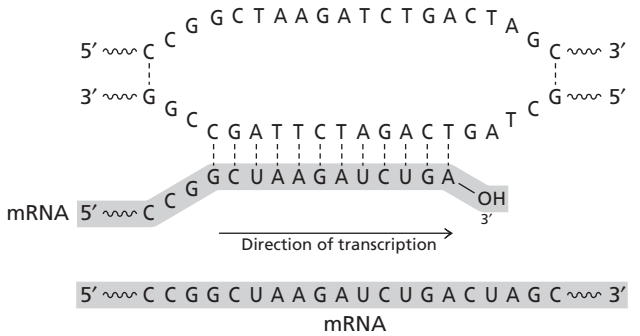
9. During maturation, eukaryotic mRNA precursors are modified at their 3′ ends by the addition of a poly A tail. When a mixture of components from a cell extract is passed over the column, the poly A tail will hybridize with oligo dT on the column. The other components in the cell extract will pass through the column. The bound mature mRNA with the poly A tail is removed from the column by changing the pH or the ionic strength of the buffer. This will disrupt the hydrogen bonds between the A and T nucleotides.

10. (a) A much lower concentration of rifampicin stopped the growth of the wild-type *E. coli* ($<5\ \mu$g/mL) as compared to the concentration of rifampicin that stopped the growth of the mutant ($>50\ \mu$g/mL).

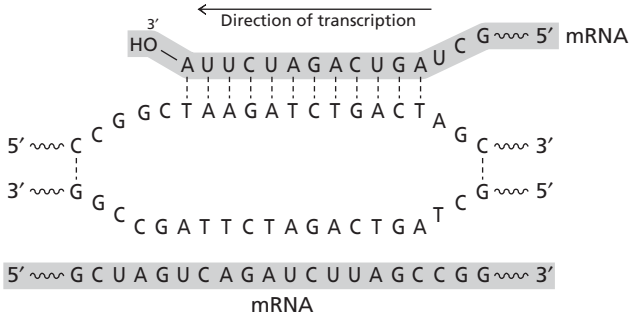
(b) RNA polymerase consists of a core enzyme with a stoichiometry of $\alpha_2\beta\beta'\omega$ that participates in many of the transcription reactions. The large β and β' subunits make up the active site of the enzyme.

(c) The rifampicin-resistant bacteria could arise from mutations that occur in the gene for the β subunit of RNA polymerase.

11. Since either strand can serve as a template, two mRNA molecules can be transcribed from this DNA segment. When the bottom strand is the template, the mRNA sequence is complementary to the bottom strand.



When the top strand is the template, the mRNA sequence is complementary to the top strand.



12. A gene was defined as a DNA sequence that is transcribed. By this definition, the entire ribosomal RNA operon is a gene. However, it is sometimes more convenient to restrict the term *gene* to the segment of RNA that encodes a functional product, for example, one of the enzymes encoded by the *lac* operon. The operon in Figure 21.25 therefore contains tRNA and 16S, 23S, and 5S rRNA genes. The DNA sequences between these genes, although transcribed, are not considered part of any gene.

13. The genomic DNA sequence provides an accurate rendition of the primary RNA sequence as expected. However, sequencing a purified tRNA reveals that many of the nucleotides have been specifically modified post-transcriptionally. The same is true for eukaryotes.

14. The gene for triose phosphate isomerase in maize contains about 3400 base pairs. If the spliceosome assembles at the first intron, then 2900 base pairs remain to be transcribed. The time required to transcribe 2900 base pairs is 97 seconds (2900 nucleotides ÷ 30 nucleotides per second). If the spliceosome assembles immediately after transcription of the first intron, and if splicing cannot begin until transcription of the entire gene is complete, the spliceosome must be stable for at least 97 seconds.

15. The CRP-cAMP binding site probably overlaps the promoter of the gene. When CRP-cAMP binds, the promoter is blocked and transcription cannot occur.

16. When the sequence of the 5′ or 3′ splice site or the branch point is altered by mutation, proper splicing cannot occur and no functional mRNA can be produced.

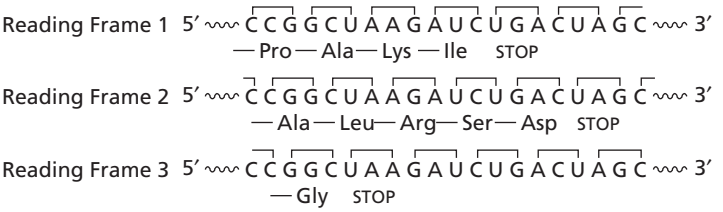
17. Yes. Once the U2 snRNP binds to the branch site it will occlude the U5 snRNP from binding to the 3′ splice acceptor and interfere with splicing. Furthermore, the deletion will have removed a large part of the pyrimidine stretch required for binding to the 3′ splice site. Both of these will prevent proper mRNA processing and the aberrant RNA will not be properly translated.

Chapter 22 Protein Synthesis

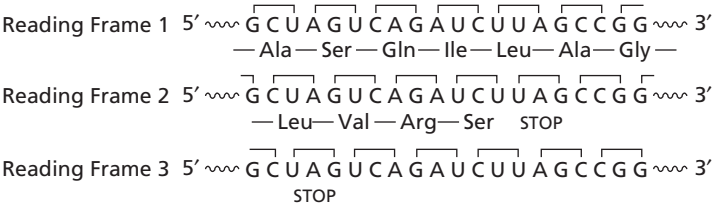
1. One strand of DNA has three different overlapping reading frames, therefore a double-stranded DNA has six reading frames. This can be seen by examining the DNA sequence beginning at the 5′ end of each strand and marking off the triplet codons. This identifies one reading frame on each strand. Now start at the second nucleotide in from the 5′ ends and mark off triplet codons; that is reading frame 2. The third reading frame on each strand begins at the third nucleotide in from the 5′ ends. The "fourth" reading frame is identical to the first—test this for yourself.

Using similar logic, it follows that if the genetic code were read in codons four nucleotides in length, then one strand of DNA could be read in four different reading frames and therefore a double-stranded piece of DNA would contain eight reading frames (four on each strand).

2. Each mRNA sequence could be translated in three different reading frames. For the first mRNA sequence, the possible codons and polypeptide sequences are



For the second mRNA sequence, the possible codons and polypeptide sequences are



Since only a reading frame without a stop codon can encode a polypeptide, the second mRNA sequence corresponds to the actual transcript. The sequence of the encoded polypeptide is –Ala–Ser–Gln–Ile–Leu–Ala–Gly–.

3. Two phosphoanhydride bonds are hydrolyzed for each amino acid activated by an aminoacyl-tRNA synthetase.

$$\text{Amino acid} + \text{tRNA} + \text{ATP} \longrightarrow \text{Aminoacyl} - \text{tRNA} + \text{AMP} + PP_i$$

$$PP_i + H_2O \longrightarrow 2\ P_i$$

The rest of the energy needed to synthesize the protein is provided by hydrolysis of GTP: one "high energy" bond is hydrolyzed in the formation of the 70S initiation complex, another during the insertion of each aminoacyl-tRNA into the A site of the ribosome, and another at each translocation step. Since the initial methionyl-tRNA is inserted into the P site, 599 new insertions and 599 translocations occur during the synthesis of a 600-residue protein. Finally, one phosphoanhydride bond is hydrolyzed during release of the completed polypeptide chain from the ribosome. The total number of phosphoanhydride bonds hydrolyzed during synthesis of the protein is

Activation (600×2)	1200
Initiation	1
Insertion	599
Translocation	599
Termination	1
Total	2400

4. The answer depends on your frame of reference. For example, relative to the ribosome, the mRNA and both tRNAs get translocated by one triplet codon. Relative to the mRNA, it is the ribosome that is shifted by three nucleotides.

5. The region of the mRNA molecule upstream of the true initiation codon contains the purine-rich Shine-Dalgarno sequence, which is complementary to a pyrimidine-rich sequence at the 3′ end of the 16S rRNA component of the 30S ribosomal subunit (Figure 22.17). By correctly positioning the 30S subunit on the mRNA transcript, the Shine-Dalgarno sequence allows fMct-$tRNA_f^{Mct}$ to bind to the initiation codon. Once protein synthesis begins, all subsequent methionine codons are recognized by Met-$tRNA^{Met}$.

6. No, because proper translation initiation in an *E. coli* cell requires a Shine-Dalgarno sequence located in the 5′ untranslated region of the mRNA. Since eukaryotic ribosomes do not have this requirement, it is extremely unlikely that an mRNA from a plant would fortuitously contain a Shine-Dalgarno sequence in the proper location.

If, however, the part of the gene encoding the plant mRNA were fused to a bacterial Shine-Dalgarno sequence, then the open reading frame for the plant protein would be properly translated in the bacterial cell.

7. The transcript of each rRNA gene is an rRNA molecule that is directly incorporated into a ribosome. Thus, multiple copies of rRNA genes are needed to assemble the large number of ribosomes that the cell requires. In contrast, the transcript of each ribosomal protein gene is an mRNA that can be translated many times. Because of this amplification of RNA to protein, fewer genes are needed for each ribosomal protein than for rRNA.

8. Possible suppressor tRNA species include all those that recognize codons differing from UAG by a single nucleotide, that is, tRNAs whose anticodons differ by a single nucleotide from the sequence CUA, which is complementary to the stop codon UAG. $tRNA^{Gln}$, $tRNA^{Lys}$, and $tRNA^{Gln}$ all recognize codons that differ only at the first position (codons CAG, AAG, and GAG, respectively). $tRNA^{Leu}$, $tRna^{Ser}$, and $tRNA^{Trp}$ recognize codons that differ only at the second position (UUG, UCG, and UGG, respectively). $tRNA^{Tyr}$ recognizes codons that differ only at the third position (UAU or UAC).

A cell that contains a suppressor tRNA can survive despite the loss of a normal tRNA because the cell also contains isoacceptor tRNA molecules that carry the same amino acid. Although the suppressor tRNA may occasionally insert an amino acid at a normally occurring stop codon, the resulting protein, which is larger than the normal gene product, is usually not lethal to the cell. In fact, strains of *E. coli* that contain suppressor tRNAs do survive but are often not as healthy as wild-type strains.

9. (a) Aminoacyl-tRNA synthetases—the enzymes that bind to tRNAs and catalyze aminoacylation.
(b) IF-2 in bacteria and eIF-2 in eukaryotes, a protein that binds to aminoacylated initiator-tRNA and loads it into the ribosome's P site during translation initiation.
(c) EF-Tu in bacteria and EF-1α in eukaryotes—a protein that binds to charged tRNAs and loads them into the ribosome's A site during polypeptide elongation.
(d) Ribosomes. These large complexes of RNA and protein contain two sites that can bind specifically to tRNAs, the A site and the P site.
(e) mRNA—tRNAs bind to mRNA through codon-anticodon hydrogen bonds.

The enzymes that modify specific residues on individual tRNAs during the maturation process must also be able to bind to tRNAs.

10. Under normal circumstances, when the translation machinery encounters UGA in RF-2 mRNA, RF-2 recognizes the stop codon and terminates protein synthesis. When the cellular concentration of RF-2 is low, however, the ribosome pauses at the termination codon, shifts frame, and continues translating the RF-2 mRNA to produce the full-length functional protein. Thus, the presence of the stop codon encourages translational frameshifting in the absence of RF-2 and allows RF-2 to regulate its own production.

11. (a) If the entire leader region were deleted, attenuation would not be possible, and transcription would be controlled exclusively by *trp* repressor. The overall rate of transcription of the *trp* operon would increase.
(b) If the region encoding the leader peptide were deleted, transcription would be controlled exclusively by *trp* repressor. Deletion of the sequence encoding the leader peptide would remove Sequence 1, thus allowing the stable 2–3 hairpin to form. Since neither the pause site (1–2 hairpin) nor the terminator (3–4 hairpin) could form, initiated transcripts would always continue into the *trp* operon.
(c) If the leader region did not contain an AUG codon, the operon would be rarely transcribed. Because of the absence of the initiation codon, the leader peptide would not be synthesized, and 1–2 hairpins and 3–4 hairpins would almost always form, leading to termination of transcription.

12. No, this is difficult to imagine. One of the important features of the attenuation model is that one or more codons in the leader peptide usually encode the amino acid that is synthesized by that operon. It is the relative shortage or abundance of particular aminoacylated tRNAs that modulates the attenuation. The products of the *lac* operon are not directly involved in amino acid biosynthesis, so we would not expect cellular levels of one class of aminoacylated tRNAs to vary with the activity of the operon.

13. The presence of codons specifying valine and leucine in the leader regions of isoleucine operons suggests that a scarcity of these amino acids would promote transcription of the genes for isoleucine biosynthesis. Many of the enzymes required to synthesize isoleucine are also required in the pathways to valine and leucine (Section 18.5A). Thus, even when the isoleucine concentration is high, a low concentration of valine or leucine ensures that transcription of the isoleucine operon does not terminate prematurely.

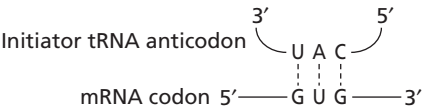
14. As the newly synthesized protein is extruded from the ribosome, the *N*-terminal signal peptide is recognized and bound by a signal-recognition particle (SRP). Further translation is inhibited until the SRP binds to its receptor on the cytosolic face of the endoplasmic reticulum. Ribophorins anchor the ribosome to the endoplasmic reticulum. When translation resumes, the polypeptide chain passes through a pore into the lumen. If the polypeptide does not pass completely through the membrane, the result is an integral membrane protein with its *N*-terminus in the lumen of the endoplasmic reticulum and its *C*-terminus in the cytosol.

Glycosylation of specific residues takes places in the lumen of the endoplasmic reticulum and in the Golgi apparatus. The protein, still embedded in the membrane, is transported between the endoplasmic reticulum and the Golgi apparatus in transfer vesicles that bud off the endoplasmic reticulum.

Secretory vesicles transport the fully glycosylated protein from the Golgi apparatus to the plasma membrane. When the vesicles fuse with the plasma membrane, the *N*-terminal portion of the protein, which was in the lumen, is now exposed to the extracellular space, and the *C*-terminal portion remains in the cytosol.

15. Yes. A hydrophobic secretion signal sequence located at the *N*-terminus of a protein is necessary and sufficient for entry into the cell's secretory pathway.

16. The initiator tRNA anticodon pairs with GUG by forming a G/U base pair between the 5′ nucleotide of the codon and the 3′ position of the anticodon.



This interaction is unrelated to wobble since the 5′ position of the anticodon is the wobble position.

Glossary of Biochemical Terms

A site. Aminoacyl site. The site on a ribosome that is occupied during protein synthesis by an aminoacyl-tRNA molecule.

acceptor stem. The sequence at the 5′ end and the sequence near the 3′ end of a tRNA molecule that are base paired, forming a stem. The acceptor stem is the site of amino acid attachment. Also known as the amino acid stem.

accessory pigments. Pigments other than chlorophyll that are present in photosynthetic membranes. The accessory pigments include carotenoids and phycobilins.

acid. A substance that can donate protons. An acid is converted to its conjugate base by loss of a proton. (The Lewis theory defines an acid as an electron-pair acceptor [Lewis acid].)

acid anhydride. The product formed by condensation of two molecules of acid.

acid dissociation constant (K_a). The equilibrium constant for the dissociation of a proton from an acid.

acid-base catalysis. Catalysis in which the transfer of a proton accelerates a reaction.

ACP. *See* acyl carrier protein.

activation energy. The free energy required to promote reactants from the ground state to the transition state in a chemical reaction.

activator. *See* transcriptional activator.

active site. The portion of an enzyme that contains the substrate-binding site and the amino-acid residues involved in catalyzing the conversion of substrate(s) to product(s). Active sites are usually located in clefts between domains or subunits of proteins or in indentations on the protein surface.

active transport. The process by which a solute specifically binds to a transport protein and is transported across a membrane against the solute concentration gradient. Energy is required to drive active transport. In primary active transport, the energy source may be light, ATP, or electron transport. Secondary active transport is driven by ion concentration gradients.

acyl carrier protein (ACP). A protein (in prokaryotes) or a domain of a protein (in eukaryotes) that binds activated intermediates of fatty acid synthesis via a thioester linkage.

adipocyte. A triacylglycerol-storage cell found in animals. An adipocyte consists of a fat droplet surrounded by a thin shell of cytosol in which the nucleus and other organelles are suspended.

adipose tissue. Animal tissue composed of specialized triacylglycerol-storage cells known as adipocytes.

A-DNA. The conformation of DNA commonly observed when purified DNA is dehydrated. A-DNA is a right-handed double helix containing approximately 11 base pairs per turn.

aerobic. Occurring in the presence of oxygen.

affinity chromatography. A chromatographic technique used to separate a mixture of proteins or other macromolecules in solution based on specific binding to a ligand that is covalently attached to the chromatographic matrix.

affinity labeling. A process by which an enzyme (or other macromolecule) is covalently inhibited by a reaction with a molecule that specifically interacts with the active site (or other binding site).

aldoses. A class of monosaccharides in which the most oxidized carbon atom, designated C-1, is aldehydic.

allosteric effector. *See* allosteric modulator.

allosteric interaction. The modulation of activity of a protein that occurs when a molecule binds to the regulatory site of the protein.

allosteric modulator. A biomolecule that binds to the regulatory site of an allosteric protein and thereby modulates its activity. An allosteric modulator may be an activator or an inhibitor. Also known as an allosteric effector.

allosteric protein. A protein whose activity is modulated by the binding of another molecule.

allosteric site. *See* regulatory site.

allosteric transitions. The changes in conformation of a protein between the active (R) state and the inactive (T) state.

α helix. A common secondary structure of proteins, in which the carbonyl oxygen of each amino acid residue (residue n) forms a hydrogen bond with the amide hydrogen of the fourth residue further toward the *C*-terminus of the polypeptide chain (residue $n + 4$). In an ideal right-handed α helix, equivalent positions recur every 0.54 nm, each amino acid residue advances the helix by 0.15 nm along the long axis of the helix, and there are 3.6 amino acid residues per turn.

amino acid. An organic acid consisting of an α-carbon atom to which an amino group, a carboxylate group, a hydrogen atom, and a specific side chain (R group) are attached. Amino acids are the building blocks of proteins.

amino acid analysis. A chromatographic procedure used for the separation and quantitation of amino acids in solutions such as protein hydrolysates.

amino terminus. *See N*-terminus.

aminoacyl site. *See* A site

aminoacyl-tRNA synthetase. An enzyme that catalyzes the activation and attachment of a specific amino acid to the 3′ end of a corresponding tRNA molecule.

amphibolic reaction. A metabolic reaction that can be both catabolic and anabolic.

amphipathic. Describes a molecule that has both hydrophobic and hydrophilic regions.

amyloplast. Modified chloroplasts that specialize in starch synthesis.

anabolic reaction. A metabolic reaction that synthesizes a molecule needed for cell maintenance and growth.

anaplerotic reaction. A reaction that replenishes metabolites removed from a central metabolic pathway (cf. cataplerotic).

angstrom (Å). A unit of length equal to 1×10^{-10} m, or 0.1 nm.

anion. An ion with an overall negative charge.

anode. A positively charged electrode. In electrophoresis, anions move toward the anode.

anomeric carbon. The most oxidized carbon atom of a cyclized monosaccharide. The anomeric carbon has the chemical reactivity of a carbonyl group.

anomers. Isomers of a sugar molecule that have different configurations only at the anomeric carbon atom.

antenna pigments. Light-absorbing pigments associated with the reaction center of a photosystem. These pigments may form a separate antenna complex or may be bound directly to the reaction-center proteins.

antibiotic. A compound, produced by one organism, that is toxic to other organisms. Clinically useful antibiotics must be specific for pathogens and not affect the human host.

antibody. A glycoprotein synthesized by certain white blood cells as part of the immunological defense system. Antibodies specifically bind to foreign compounds, called antigens, forming antibody-antigen complexes that mark the antigen for destruction. Also known as an immunoglobulin.

anticodon. A sequence of three nucleotides in the anticodon loop of a tRNA molecule. The anticodon binds to the complementary codon in mRNA during translation.

anticodon arm. The stem-and-loop structure in a tRNA molecule that contains the anticodon.

antigen. A molecule or part of a molecule that is specifically bound by an antibody.

antiport. The cotransport of two different species of ions or molecules in opposite directions across a membrane by a transport protein.

antisense strand. In double-stranded DNA the antisense strand is the strand that does not contain codons. Also called the template strand. The opposite strand is called the sense strand or the coding strand.

antisense RNA. An RNA molecule that binds to a complementary mRNA molecule, forming a double-stranded region that inhibits translation of the mRNA.

apoprotein. A protein whose cofactor(s) is absent. Without the cofactor(s), the apoprotein lacks the biological activity characteristic of the corresponding holoprotein.

apoptosis. The programed death of a cell.

atomic mass unit. The unit of atomic weight equal to 1/12th the mass of the ^{12}C isotope of carbon. The mass of the ^{12}C nuclide is exactly 12 by definition.

attenuation. A mechanism of regulation of gene expression that couples translation and transcription. Generally, the translation of a short reading frame at the beginning of a prokaryotic operon will determine whether transcription terminates before the rest of the operon is transcribed.

autophosphorylation. Phosphorylation of a protein kinase catalyzed by another molecule of the same kinase.

autosome. A chromosome other than a sex chromosome.

autotroph. An organism that can grow and reproduce using only inorganic substances (such as CO_2) as its only source of essential elements.

backbone. 1. The repeating $N—C_\alpha—C$ units connected by peptide bonds in a polypeptide chain. **2.** The repeating sugar-phosphate units connected by phosphodiester linkages in a nucleic acid.

bacteriophage. A virus that infects a bacterial cell.

base. 1. A substance that can accept protons. A base is converted to its conjugate acid by addition of a proton. (The Lewis theory defines a base as an electron-pair donor [Lewis base].) **2.** The substituted pyrimidine or purine of a nucleoside or nucleotide. The heterocyclic bases of nucleosides and nucleotides can participate in hydrogen bonding.

base pairing. The interaction between the bases of nucleotides in single-stranded nucleic acids to form double-stranded molecules, such as DNA, or regions of double-stranded secondary structure. The most common base pairs are formed by hydrogen bonding of adenine (A) with thymine (T) or uracil (U) and of guanine (G) with cytosine (C).

B-DNA. The most common conformation of DNA and the one proposed by Watson and Crick. B-DNA is a right-handed double helix with a diameter of 2.37 nm and approximately 10.4 base pairs per turn.

β-oxidation pathway. The metabolic pathway that degrades fatty acids to acetyl CoA, producing NADH and QH_2 and thereby generating large amounts of ATP. Each round of β-oxidation of fatty acids consists of four steps: oxidation, hydration, further oxidation, and thiolysis.

β pleated sheet. *See* β sheet.

β sheet. A common secondary structure of proteins that consists of extended polypeptide chains stabilized by hydrogen bonds between the carbonyl oxygen of one peptide bond and the amide hydrogen of another on the same or an adjacent polypeptide chain. The hydrogen bonds are nearly perpendicular to the extended polypeptide chains, which may be either parallel (running in the same *N*- to *C*-terminal direction) or antiparallel (running in opposite directions).

β strand. An extended polypeptide chain within a β sheet secondary structure or having the same conformation as a strand within a β sheet.

β turn. *See* turn.

bile. A suspension of bile salts, bile pigments, and cholesterol that originates in the liver and is stored in the gall bladder. Bile is secreted into the small intestine during digestion.

binding-change mechanism. A proposed mechanism for the phosphorylation of ADP and release of ATP from F_0F_1 ATP synthase. The mechanism proposes three different binding-site conformations for ATP synthase: an open site from which ATP has been released, an ATP-bearing tight-binding site that is catalytically active, and an ADP and P_i loose-binding site that is catalytically inactive. Inward passage of protons through the ATP synthase complex into the mitochondrial matrix causes the open site to become a loose site; the loose site, already filled with ADP and P_i, to become a tight site; and the ATP-bearing site to become an open site.

bioenergetics. The study of energy changes in biological systems.

biological membrane. *See* membrane.

biopolymer. A biological macromolecule in which many identical or similar small molecules are covalently linked to one another to form a long chain. Proteins, polysaccharides, and nucleic acids are biopolymers.

Bohr effect. The phenomenon observed when exposure to carbon dioxide, which lowers the pH inside the cells, causes the oxygen affinity of hemoglobin in red blood cells to decrease.

branch migration. The movement of a crossover, or branch point, resulting in further exchange of DNA strands during recombination.

branch site. The point within an intron that becomes attached to the 5′ end of the intron during splicing of mRNA precursors.

buffer. A solution of an acid and its conjugate base that resists changes in pH.

buffer capacity. The ability of a solution to resist changes in pH. For a given buffer, maximum buffer capacity is achieved at the pH at which the concentrations of the weak acid and its conjugate base are equal (i.e., when $pH = pK_a$).

C_4 pathway. A pathway for carbon fixation in several plant species that minimizes photorespiration by concentrating CO_2. In this pathway, CO_2 is incorporated into C_4 acids in the mesophyll cells, and the C_4 acids are decarboxylated in the bundle sheath cells, releasing CO_2 for use by the reductive pentose phosphate cycle.

calorie (cal). The amount of energy required to raise the temperature of 1 gram of water by 1°C (from 14.5°C to 15.5°C). One calorie is equal to 4.184 J.

Calvin cycle. A cycle of reactions that involve the fixation of carbon dioxide and the net production of glyceraldehyde-3-phosphate. Usually associated with photosynthesis. Also known as the Calvin-Benson cycle, the C3 pathway, and the reductive pentose phosphate (RPP) cycle.

Calvin-Benson cycle. *See* Calvin cycle.

CAM. *See* Crassulacean acid metabolism.

cap. A 7-methylguanosine residue attached by a pyrophosphate linkage to the 5′ end of a eukaryotic mRNA molecule. The cap is added posttranscriptionally and is required for efficient translation. Further covalent modifications yield alternative cap structures.

carbanion. A carbon anion that results from the cleavage of a covalent bond between carbon and another atom in which both electrons from the bond remain with the carbon atom.

carbocation. A carbon cation that results from the cleavage of a covalent bond between carbon and another atom in which the carbon atom loses both electrons from the bond.

carbohydrate. Loosely defined as a compound that is a hydrate of carbon in which the ratio of C:H:O is 1:2:1. Carbohydrates include monomeric sugars (i.e., monosaccharides) and their polymers. Also known as a saccharide.

carboxyl terminus. *See C*-terminus.

carnitine shuttle system. A cyclic pathway that shuttles acetyl CoA from the cytosol to the mitochondria by formation and transport of acyl carnitine.

cascade. Sequential activation of several components, resulting in signal amplification.

catabolic reaction. A metabolic reaction that degrades a molecule to provide smaller molecular building blocks and energy to an organism.

catabolite repression. A regulatory mechanism that results in increased rates of transcription of many bacterial genes and operons when glucose is present. A complex between cAMP and cAMP regulatory protein (CRP) activates transcription.

catalytic antibodies. Antibody molecules that have been genetically manipulated so that they catalyze reactions involving the antigen.

catalytic center. The polar amino acids in the active site of an enzyme that participate in chemical changes during catalysis.

catalytic constant (k_{cat}). A kinetic constant that is a measure of how rapidly an enzyme can catalyze a reaction when saturated with its

substrate(s). The catalytic constant is equal to the maximum velocity (V_{max}) divided by the total concentration of enzyme ($[E]_{total}$), or the number of moles of substrate converted to product per mole of enzyme active sites per second, under saturating conditions. Also known as the turnover number.

catalytic proficiency. The ratio of the rate constants for a reaction in the presence of enzyme (k_{cat}/K_m) to the rate constant for the chemical reaction in the absence of enzyme.

cataplerotic reaction. A reaction that removes intermediates in a pathway, especially the citric acid cycle (cf., anaplerotic).

cathode. A negatively charged electrode. In electrophoresis, cations move toward the cathode.

cation. An ion with an overall positive charge.

cDNA. *See* complementary DNA.

Central Dogma. The concept that the flow of information from nucleic acid to protein is irreversible. The term is often applied incorrectly to the actual pathway of information flow from DNA to RNA to protein.

ceramide. A molecule that consists of a fatty acid linked to the C-2 amino group of sphingosine by an amide bond. Ceramides are the metabolic precursors of all sphingolipids.

cerebroside. A glycosphingolipid that contains one monosaccharide residue attached via a β-glycosidic linkage to C-1 of a ceramide. Cerebrosides are abundant in nerve tissue and are found in myelin sheaths.

channel. An integral membrane protein with a central aqueous passage, which allows appropriately sized molecules and ions to traverse the membrane in either direction. Also known as a pore.

channeling. *See* metabolite channeling.

chaotropic agent. A substance that enhances the solubility of nonpolar compounds in water by disrupting regularities in hydrogen bonding among water molecules. Concentrated solutions of chaotropic agents, such as urea and guanidinium salts, decrease the hydrophobic effect and are thus effective protein denaturants.

chaperone. A protein that forms complexes with newly synthesized polypeptide chains and assists in their correct folding into biologically functional conformations. Chaperones may also prevent the formation of incorrectly folded intermediates, prevent incorrect aggregation of unassembled protein subunits, assist in translocation of polypeptide chains across membranes, and assist in the assembly and disassembly of large multiprotein structures.

charge-charge interaction. A noncovalent electrostatic interaction between two charged particles.

chelate effect. The phenomenon by which the constant for binding of a ligand having two or more binding sites to a molecule or atom is greater than the constant for binding of separate ligands to the same molecule or atom.

chemiosmotic theory. A theory proposing that a proton concentration gradient established during oxidation of substrates provides the energy to drive processes such as the formation of ATP from ADP and P_i.

chemoautotroph. An autotroph that derives chemical energy by oxidizing inorganic compounds (cf., photoautotroph).

chemoheterotroph. Non-photosynthetic organism that requires organic molecules as a carbon source and derives energy from oxidizing organic molecules.

chemotaxis. A mechanism that couples signal transduction to flagella movement in bacteria causing them to move toward a chemical (positive chemotaxis) or away from a chemical (negative chemotaxis).

chiral atom. An atom with asymmetric substitution that can exist in two different configurations.

chloroplast. A chlorophyll-containing organelle in algae and plant cells that is the site of photosynthesis.

chromatin. A DNA-protein complex in the nuclei of eukaryotic cells.

chromatography. A technique used to separate components of a mixture based on their partitioning between a mobile phase, which can be gas or liquid, and a stationary phase, which is a liquid or solid.

chromosome. A single DNA molecule containing many genes. An organism may have a genome consisting of a single chromosome or many.

chylomicron. A type of plasma lipoprotein that transports triacylglycerols, cholesterol, and cholesteryl esters from the small intestine to the tissues.

citric acid cycle. A metabolic cycle consisting of eight enzyme-catalyzed reactions that completely oxidizes acetyl units to CO_2. The energy released in the oxidation reactions is conserved as reducing power when the coenzymes NAD^1 and ubiquinone (Q) are reduced. Oxidation of one molecule of acetyl CoA by the citric acid cycle generates three molecules of NADH, one molecule of QH_2, and one molecule of GTP or ATP. Also known as the Krebs cycle and the tricarboxylic acid cycle.

clone. One of the identical copies derived from the replication or reproduction of a single molecule, cell, or organism.

cloning. The generation of many identical copies of a molecule, cell, or organism. Cloning sometimes refers to the entire process of constructing and propagating a recombinant DNA molecule.

cloning vector. A DNA molecule that carries a segment of foreign DNA. A cloning vector introduces the foreign DNA into a cell where it can be replicated and sometimes expressed.

coding strand. The strand of DNA within a gene whose nucleotide sequence is identical to that of the RNA produced by transcription (with the replacement of T by U in RNA).

codon. A sequence of three nucleotide residues in mRNA (or DNA) that specifies a particular amino acid according to the genetic code.

coenzyme. An organic molecule required by an enzyme for full activity. Coenzymes can be further classified as cosubstrates or prosthetic groups.

coenzyme A. A large coenyme used in transferring acyl groups.

cofactor. An inorganic ion or organic molecule required by an apoenzyme to convert it to a holoenzyme. There are two types of cofactors: essential ions and coenzymes.

column chromatography. A technique for purifying proteins. *See* affinity chromatography, gel-filtration chromatography, ion-exchange chromatography, HPLC, and affinity chromatography.

competitive inhibition. Reversible inhibition of an enzyme-catalyzed reaction by an inhibitor that prevents substrate binding.

complementary DNA (cDNA). DNA synthesized from an mRNA template by the action of reverse transcriptase.

concerted theory of cooperativity and allosteric regulation. A model of the cooperative binding of ligands to oligomeric proteins. According to the concerted theory, the change in conformation of a protein due to the binding of a substrate or an allosteric modulator shifts the equilibrium of the conformation of the protein between T (a low substrate-affinity conformation) and R (a high substrate-affinity conformation). This theory suggests that all subunits of the protein have the same conformation, either all T or all R. Also known as the symmetry-driven theory.

condensation. A reaction involving the joining of two or more molecules accompanied by the elimination of water, alcohol, or other simple substance.

configuration. A spatial arrangement of atoms that cannot be altered without breaking and re-forming covalent bonds.

conformation. Any three-dimensional structure, or spatial arrangement, of a molecule that results from rotation of functional groups around single bonds. Because there is free rotation around single bonds, a molecule can potentially assume many conformations.

conjugate acid. The product resulting from the gain of a proton by a base.

conjugate base. The product resulting from the loss of a proton by an acid.

consensus sequence. The sequence of nucleotides most commonly found at each position within a region of DNA or RNA.

cooperativity. 1. The phenomenon whereby the binding of one ligand or substrate molecule to a protein influences the affinity of the protein for additional molecules of the same substance. Cooperativity may be positive or negative. **2.** The phenomenon whereby formation

of structure in one part of a macromolecule promotes the formation of structure in the rest of the molecule.

core particle. *See* nucleosome core particle.

corepressor. A ligand that binds to a repressor of a gene causing it to bind DNA and prevent transcription.

Cori cycle. An interorgan metabolic loop that recycles carbon and transports energy from the liver to the peripheral tissues. Glucose is released from the liver and metabolized to produce ATP in other tissues. The resulting lactate is then returned to the liver for conversion back to glucose by gluconeogenesis.

cosubstrate. A coenzyme that is a substrate in an enzyme-catalyzed reaction. A cosubstrate is altered during the course of the reaction and dissociates from the active site of the enzyme. The original form of the cosubstrate can be regenerated in a subsequent enzyme-catalyzed reaction.

cotransport. The coupled transport of two different species of solutes across a membrane, in the same direction (symport) or the opposite direction (antiport), carried out by a transport protein.

coupled reactions. Two metabolic reactions that share a common intermediate.

covalent catalysis. Catalysis in which one substrate, or part of it, forms a covalent bond with the catalyst and then is transferred to a second substrate. Many enzymatic group-transfer reactions proceed by covalent catalysis.

Crassulacean acid metabolism (CAM). A modified sequence of carbon-assimilation reactions used primarily by plants in arid environments to reduce water loss during photosynthesis. In these reactions, CO_2 is taken up at night, resulting in the formation of malate. During the day, malate is decarboxylated, releasing CO_2 for use by the reductive pentose phosphate cycle.

***C*-terminus.** The amino acid residue bearing a free carboxyl group at one end of a peptide chain. Also known as the carboxyl terminus.

cyclic electron transport. A modified sequence of electron transport steps in chloroplasts that operates to provide ATP without the simultaneous formation of NADPH.

cytoplasm. The part of a cell enclosed by the plasma membrane, excluding the nucleus.

cytoskeleton. A network of proteins that contributes to the structure and organization of a eukaryotic cell.

cytosol. The aqueous portion of the cytoplasm minus the subcellular structures.

D arm. The stem-and-loop structure in a tRNA molecule that contains dihydrouridylate (D) residues.

dalton. A unit of mass equal to one atomic mass unit.

dark reactions. The photosynthetic reactions in which NADPH and ATP are used to fix CO_2 to carbohydrate. Also known as the light-independent reactions.

degeneracy. When referring to the genetic code, degeneracy refers to the fact that several different codons specify the same amino acid.

dehydrogenase. An enzyme that catalyzes the removal of hydrogen from a substrate or the oxidation of a substrate. Dehydrogenases are members of the IUBMB class of enzymes known as oxidoreductases.

denaturation. 1. A disruption in the native conformation of a biological macromolecule that results in loss of the biological activity of the macromolecule. **2.** The complete unwinding and separation of complementary strands of DNA.

detergent. An amphipathic molecule consisting of a hydrophobic portion and a hydrophilic end that may be ionic or polar. Detergent molecules can aggregate in aqueous media to form micelles. Also known as a surfactant.

dialysis. A procedure in which low-molecular-weight solutes in a sample are removed by diffusion through a semipermeable barrier and replaced by solutes from the surrounding medium.

diffusion controlled reaction. A reaction that occurs with every collision between reactant molecules. In enzyme-catalyzed reactions, the k_{cat}/K_m ratio approaches a value of $10^8 - 10^9\ M^{-1}\ s^{-1}$.

diploid. Having two sets of chromosomes or two copies of the genome.

dipole. Two equal but opposite charges, separated in space, resulting from the uneven distribution of charge within a molecule or a chemical bond.

direct repair. The removal of DNA damage by proteins that recognize damaged nucleotides and mismatched bases and repair them without cleaving the DNA or excising the base.

distributive enzyme. An enzyme that dissociates from its growing polymeric product after addition of each monomeric unit and must reassociate with the polymer for polymerization to proceed (cf., progressive enzyme).

disulfide bond. A covalent linkage formed by oxidation of the sulfhydryl groups of two cysteine residues. Disulfide bonds are important in stabilizing the three-dimensional structures of some proteins.

domain. A discrete, independent folding unit within the tertiary structure of a protein. Domains are usually combinations of several motifs forming a characteristic fold.

double helix. A nucleic acid conformation in which two antiparallel polynucleotide strands wrap around each other to form a two-stranded helical structure stabilized largely by stacking interactions between adjacent hydrogen-bonded base pairs.

double-reciprocal plot. A plot of the reciprocal of initial velocity versus the reciprocal of substrate concentration for an enzyme-catalyzed reaction. The x and y intercepts indicate the values of the reciprocals of the Michaelis constant and the maximum velocity, respectively. A double-reciprocal plot is a linear transformation of the Michaelis-Menten equation. Also known as a Lineweaver-Burk plot.

E. *See* reduction potential.

E°′. *See* standard reduction potential.

E site. Exit site. The site on a ribosome from which a deaminoacylated tRNA is released during protein synthesis.

Edman degradation. A procedure used to determine the sequence of amino acid residues from a free *N*-terminus of a polypeptide chain. The *N*-terminal residue is chemically modified, cleaved from the chain, and identified by chromatographic procedures, and the rest of the polypeptide is recovered. Multiple reaction cycles allow identification of the new *N*-terminal residue generated by each cleavage step.

effector enzyme. A membrane-associated protein that produces an intracellular second messenger in response to a signal from a transducer.

eicosanoid. An oxygenated derivative of a 20-carbon polyunsaturated fatty acid. Eicosanoids function as short-range messengers in the regulation of various physiological processes.

electromotive force (emf). A measure of the difference between the reduction potentials of the reactions on the two sides of an electrochemical cell (i.e., the voltage difference produced by the reactions).

electrolyte. A molecule such as NaCl that can dissociated to form ions.

electron transport. A set of reactions in which compounds such as NADH and reduced ubiquinone (QH_2) are aerobically oxidized and ATP is generated from ADP and P_i. Membrane-associated electron transport consists of two tightly coupled phenomena: oxidation of substrates by the respiratory electron transport chain, accompanied by the translocation of protons across the inner mitochondrial membrane to generate a proton concentration gradient; and formation of ATP, driven by the flux of protons into the matrix through a channel in ATP synthase.

electrophile. A positively charged or electron-deficient species that is attracted to chemical species that are negatively charged or contain unshared electron pairs (nucleophiles).

electrophoresis. A technique used to separate molecules by their migration in an electric field, primarily on the basis of their net charge.

electrospray mass spectrometry. A technique in mass spectrometry where the target molecule is sprayed into the detector in tiny droplets.

electrostatic interaction. A general term for the electronic interaction between particles. Electrostatic interactions include charge-charge interactions, hydrogen bonds, and van der Waals forces.

elongation factor. A protein that is involved in extending the peptide chain during protein synthesis.

enantiomers. Stereoisomers that are nonsuperimposable mirror images.

endocytosis. The process by which matter is engulfed by a plasma membrane and brought into the cell within a lipid vesicle derived from the membrane.

endonuclease. An enzyme that catalyzes the hydrolysis of phosphodiester linkages at various sites within polynucleotide chains.

endoplasmic reticulum. A membranous network of tubules and sheets continuous with the outer nuclear membrane of eukaryotic cells. Regions of the endoplasmic reticulum coated with ribosomes are called the rough endoplasmic reticulum; regions having no attached ribosomes are known as the smooth endoplasmic reticulum. The endoplasmic reticulum is involved in the sorting and transport of certain proteins and in the synthesis of lipids.

endosomes. Smooth vesicles inside the cell that are receptacles for endocytosed material.

energy-rich compound. A compound whose hydrolysis occurs with a large negative free-energy change (equal to or greater than that for ATP $\rightarrow$ ADP + P_i).

enthalpy (*H*). A thermodynamic state function that describes the heat content of a system.

entropy (*S*). A thermodynamic state function that describes the randomness or disorder of a system.

enzymatic reaction. A reaction catalyzed by a biological catalyst, an enzyme. Enzymatic reactions are 10^3 to 10^{17} times faster than the corresponding uncatalyzed reactions.

enzyme. A biological catalyst, almost always a protein. Some enzymes may require additional cofactors for activity. Virtually all biochemical reactions are catalyzed by specific enzymes.

enzyme assay. A method used to analyze the activity of a sample of an enzyme. Typically, enzymatic activity is measured under selected conditions such that the rate of conversion of substrate to product is proportional to enzyme concentration.

enzyme inhibitor. A compound that binds to an enzyme and interferes with its activity by preventing either the formation of the ES complex or its conversion to E + P.

enzyme-substrate complex (ES). A complex formed when substrate molecules bind noncovalently within the active site of an enzyme.

epimers. Isomers that differ in configuration at only one of several chiral centers.

equilibrium. The state of a system in which the rate of conversion of substrate to product is equal to the rate of conversion of product to substrate. The free-energy change for a reaction or system at equilibrium is zero.

equilibrium constant (K_{eq}). The ratio of the concentrations of products to the concentrations of reactants at equilibrium. The equilibrium constant is related to the standard Gibbs free energy change of reaction.

essential amino acid. An amino acid that cannot be synthesized by an animal and must be obtained in the diet.

essential fatty acid. A fatty acid that cannot be synthesized by an animal and must be obtained in the diet.

essential ion. An ion required as a cofactor for the catalytic activity of certain enzymes. Some essential ions, called activator ions, are reversibly bound to enzymes and often participate in the binding of substrates, whereas tightly bound metal ions frequently participate directly in catalytic reactions.

eukaryote. An organism whose cells generally possess a nucleus and internal membranes (cf., prokaryote).

excision repair. The reversal of DNA damage by excision-repair endonucleases. Gross lesions that alter the structure of the DNA helix are repaired by cleavage on each side of the lesion and removal of the damaged DNA. The resulting single-stranded gap is filled by DNA polymerase and sealed by DNA ligase.

exocytosis. The process by which material destined for secretion from a cell is enclosed in lipid vesicles that are transported to and fuse with the plasma membrane, releasing the material into the extracellular space.

exon. A nucleotide sequence that is present in the primary RNA transcript and in the mature RNA molecule. The term exon also refers to the region of the gene that corresponds to a sequence present in the mature RNA (cf., intron).

exonuclease. An enzyme that catalyzes the sequential hydrolysis of phosphodiester linkages from one end of a polynucleotide chain.

extrinsic membrane protein. *See* peripheral membrane protein.

facilitated diffusion. *See* passive transport.

facultative anaerobe. An organism that can survive in the presence or absence of oxygen.

fatty acid. A long chain aliphatic hydrocarbon with a single carboxyl group at one end. Fatty acids are the simplest type of lipid and are components of many more complex lipids, including triacylglycerols, glycerophospholipids, sphingolipids, and waxes.

feedback inhibition. Inhibition of an enzyme that catalyzes an early step in a metabolic pathway by an end product of the same pathway.

feed-forward activation. Activation of an enzyme in a metabolic pathway by a metabolite produced earlier in the pathway.

fermentation. The anaerobic catabolism of metabolites for energy production. In alcoholic fermentation, pyruvate is converted to ethanol and carbon dioxide.

fibrous proteins. A major class of water-insoluble proteins that associate to form long fibers. Many fibrous proteins are physically tough and provide mechanical support to individual cells or entire organisms.

first-order reaction. A reaction whose rate is directly proportional to the concentration of only one reactant.

Fischer projection. A two-dimensional representation of the three-dimensional structures of sugars and related compounds. In a Fischer projection, the carbon skeleton is drawn vertically, with C-1 at the top. At a chiral center, horizontal bonds extend toward the viewer and vertical bonds extend away from the viewer.

fluid mosaic model. A model proposed for the structure of biological membranes. In this model, the membrane is depicted as a dynamic structure in which lipids and membrane proteins (both integral and peripheral) rotate and undergo lateral diffusion.

fluorescence. A form of luminescence in which visible radiation is emitted from a molecule as it passes from a higher to a lower electronic state.

flux. The flow of material through a metabolic pathway. Flux depends on the supply of substrates, the removal of products, and the catalytic capabilities of the enzymes involved in the pathway.

fold. A combination of secondary structures that form the core of a protein domain. Many different folds have been characterized.

frameshift mutation. An alteration in DNA caused by the insertion or deletion of a number of nucleotides not divisible by three. A frameshift mutation changes the reading frame of the corresponding mRNA molecule and affects translation of all codons downstream of the mutation.

free energy change. *See* Gibbs free energy change.

free radical. A molecule or atom with an unpaired electron.

furanose. A monosaccharide structure that forms a five-membered ring as a result of intramolecular hemiacetal formation.

G protein. A protein that binds guanine nucleotides.

ΔG. *See* Gibbs free energy change.

$\Delta G^{\circ\prime}$. *See* standard Gibbs free energy change.

ganglioside. A glycosphingolipid in which oligosaccharide chains containing *N*-acetylneuraminic acid are attached to a ceramide. Gangliosides are present on cell surfaces and provide cells with distinguishing surface markers that may serve in cellular recognition and cell-to-cell communication.

gas chromatography. A chromatographic technique used to separate components of a mixture based on their partitioning between the gas phase and a stationary phase, which can be a liquid or solid.

gel-filtration chromatography. A chromatographic technique used to separate a mixture of proteins or other macromolecules in solution based on molecular size, using a matrix of porous beads. Also known as molecular-exclusion chromatography.

gene. Loosely defined as a segment of DNA that is transcribed. In some cases, the term gene may also be used to refer to a segment of DNA that encodes a functional protein or corresponds to a mature RNA molecule.

genetic code. The correspondence between a particular three nucleotide codon and the amino acid it specifies. The standard genetic code of 64 codons is used by almost all organisms. The genetic code is used to translate the sequence of nucleotides in mRNA into protein.

genetic recombination. The exchange or transfer of DNA from one molecule of DNA to another (cf., homologous recombination).

genome. One complete set of the genetic information in an organism. It may be a single chromosome or a set of chromosomes (haploid). Mitochondria and chloroplasts have genomes separate from that in the nucleus of eukaryotic cells.

Gibbs free energy change (ΔG). A thermodynamic quantity that defines the equilibrium condition in terms of the changes in enthalpy (H) and entropy (S) of a system at constant pressure. $\Delta G = \Delta H - T\Delta S$, where T is absolute temperature. Free energy is a measure of the energy available within a system to do work.

globular proteins. A major class of proteins, many of which are water soluble. Globular proteins are compact and roughly spherical, containing tightly folded polypeptide chains. Typically, globular proteins include indentations, or clefts that specifically recognize and transiently bind other compounds.

glucogenic compound. A compound, such as an amino acid, that can be used for gluconeogenesis in animals.

gluconeogenesis. A pathway for synthesis of glucose from a noncarbohydrate precursor. Gluconeogenesis from pyruvate involves the seven near-equilibrium reactions of glycolysis traversed in the reverse direction. The three metabolically irreversible reactions of glycolysis are bypassed by four enzymatic reactions that do not occur in glycolysis.

glucoside. A glycoside where the anomeric carbon atom is from glucose.

glycan. A general term for an oligosaccharide or a polysaccharide. A homoglycan is a polymer of identical monosaccharide residues; a heteroglycan is a polymer of different monosaccharide residues.

glycerophospholipid. A lipid consisting of two fatty acyl groups bound to C-1 and C-2 of glycerol 3-phosphate and, in most cases, a polar substituent attached to the phosphate moiety. Glycerophospholipids are major components of biological membranes.

glycoconjugate. A carbohydrate derivative in which one or more carbohydrate chains are covalently linked to a peptide chain, protein, or lipid.

glycoforms. Glycoproteins containing identical amino acid sequences but different oligosaccharide-chain compositions.

glycogen. A branched homopolymer of glucose residues joined by α-(1 → 4) linkages with α-(1 → 6) linkages at branch points. Glycogen is a storage polysaccharide in animals and bacteria.

glycolysis. A catabolic pathway consisting of 10 enzyme-catalyzed reactions by which one molecule of glucose is converted to two molecules of pyruvate. In the process, two molecules of ATP are formed from ADP + P_i, and two molecules of NAD^+ are reduced to NADH.

glycoprotein. A protein that contains covalently bound carbohydrate residues.

glycosaminoglycan. An unbranched polysaccharide of repeating disaccharide units. One component of the disaccharide is an amino sugar; the other component is usually a uronic acid.

glycoside. A molecule containing a carbohydrate in which the hydroxyl group of the anomeric carbon has been replaced through condensation with an alcohol, an amine, or a thiol.

glycosidic bond. Acetal linkage formed by condensation of the anomeric carbon atom of a saccharide with a hydroxyl, amino, or thiol group of another molecule. The most commonly encountered glycosidic bonds are formed between the anomeric carbon of one sugar and a hydroxyl group of another sugar. Nucleosidic bonds are *N*-linked glycosidic bonds.

glycosphingolipid. A lipid containing sphingosine and carbohydrate moieties.

glycosylation. *See* protein glycosylation.

glyoxylate cycle. A variation of the citric acid cycle in certain plants, bacteria, and yeast that allows net production of glucose from acetyl CoA via oxaloacetate. The glyoxylate cycle bypasses the two CO_2 producing steps of the citric acid cycle.

glyoxysome. An organelle that contains specialized enzymes for the glyoxylate cycle.

Golgi apparatus. A complex of flattened, fluid-filled membranous sacs in eukaryotic cells, often found in proximity to the endoplasmic reticulum. The Golgi apparatus is involved in the modification, sorting, and targeting of proteins.

granum. A stack of flattened vesicles formed from the thylakoid membrane in chloroplasts.

group transfer potential. *See* photsphoryl group transfer potential.

group transfer reaction. A reaction in which a substituent or functional group is transferred from one substrate to another.

H. *See* enthalpy.

hairpin. 1. A secondary structure adopted by single-stranded polynucleotides that arises when short regions fold back on themselves and hydrogen bonds form between complementary bases. Also known as a stem-loop. **2.** A tight turn connecting two consecutive β strands of a polypeptide.

haploid. Having one set of chromosomes or one copy of the genome (cf., diploid).

high energy molecule. *See* energy-rich compound.

Haworth projection. A representation in which a cyclic sugar molecule is depicted as a flat ring that is projected perpendicular to the plane of the page. Heavy lines represent the part of the molecule that extends toward the viewer.

HDL. *See* high density lipoprotein.

heat of vaporization. The amount of heat required to evaporate 1 gram of a liquid.

heat shock protein. A protein whose synthesis is increased in response to stresses such as high temperature. Many heat shock proteins are chaperones that are also expressed in the absence of stress.

helicase. An enzyme that is involved in unwinding DNA.

hemiacetal. The product formed when an alcohol reacts with an aldehyde.

hemiketal. The product formed when an alcohol reacts with a ketone.

Henderson-Hasselbalch equation. An equation that describes the pH of a solution of a weak acid or a weak base in terms of the pK_a and the concentrations of the proton donor and proton acceptor forms.

heterochromatin. Regions of chromatin that are highly condensed.

heterocyclic molecule. A molecule that contains a ring structure made up of more than one type of atom.

heteroglycan (heteropolysaccharide). A carbohydrate polymer whose residues consist of two or more different types of monosaccharide.

heterotroph. An organism that requires at least one organic nutrient, such as glucose, as a carbon source.

high density lipoprotein (HDL). A type of plasma lipoprotein that is enriched in protein and transports cholesterol and cholesteryl esters from tissues to the liver.

high-performance liquid chromatography (HPLC). A chromatographic technique used to separate components of a mixture by dissolving the mixture in a liquid solvent and forcing it to flow through a chromatographic column under high pressure.

histones. A class of proteins that bind to DNA to form chromatin. The nuclei of eukaryotic cells contain five histones, known as H1, H2A, H2B, H3, and H4.

Holliday junction. The region of strand crossover resulting from recombination between two molecules of homologous double-stranded DNA.

homoglycan (homopolysaccharide). A carbohydrate polymer whose residues consist of a single type of monosaccharide.

homologous. Referring to genes or proteins that descend from a common ancestor.

homologous recombination. Recombination between molecules of DNA that have closely related sequences (i.e., they are homologous). This is the standard form of recombination that occurs between chromosomes in eukaryotic cells.

homology. The similarity of genes or proteins as a result of evolution from a common ancestor.

hormone response element. A DNA sequence that binds a transcriptional activator consisting of a steroid hormone receptor complex.

housekeeping genes. Genes that encode proteins or RNA molecules that are essential for the normal activities of all living cells.

HPLC. *See* high-performance liquid chromatography.

hydration. A state in which a molecule or ion is surrounded by water.

hydrogen bond. A weak electrostatic interaction the formed when a hydrogen atom bonded covalently to a strongly electronegative atom is partially shared by interacting with electron pair of another electronegative atom.

hydrolase. An enzyme that catalyzes the hydrolytic cleavage of its substrate(s) (i.e., hydrolysis).

hydropathy. A measure of the hydrophobicity of amino acid side chains. The more positive the hydropathy value, the greater the hydrophobicity.

hydrophilic. "Water loving"—describing molecules that interact favorably with water.

hydrophilicity. The degree to which a compound or functional group interacts with water or is preferentially soluble in water.

hydrophobic. "Water fearing"—describing molecules that do not interact favorably with water and are much less soluble than hydrophilic molecules.

hydrophobic effect. The exclusion of hydrophobic groups or molecules by water. The hydrophobic effect appears to depend on the increase in entropy of solvent water molecules that are released from an ordered arrangement around the hydrophobic group.

hydrophobic interaction. A weak, noncovalent interaction between nonpolar molecules or substituents that results from the strong association of water molecules with one another. Such association leads to the shielding or exclusion of nonpolar molecules from an aqueous environment.

hydrophobicity. The degree to which a compound or functional group that is soluble in nonpolar solvents is insoluble or only sparingly soluble in water.

IDL. *See* intermediate density lipoprotein.

induced fit. Activation of an enzyme by a substrate-initiated conformational change.

inducer. A ligand that binds to and inactivates a repressor thereby increasing the transcription of the gene controlled by the repressor.

inhibition constant (K_i). The equilibrium constant for the dissociation of an inhibitor from an enzyme-inhibitor complex.

inhibitor. A compound that binds to an enzyme and inhibits its activity.

initial velocity (v_0). The rate of conversion of substrate to product in the early stages of an enzymatic reaction, before appreciable product has been formed.

initiation codon. A codon that specifies the initiation site for protein synthesis. The methionine codon (AUG) is the most common initiation codon.

initiation factor. *See* translation initiation factor.

initiator tRNA. The tRNA molecule that is used exclusively at initiation codons. The initiator tRNA is usually a specific methionyl-tRNA.

integral membrane protein. A membrane protein that penetrates the hydrophobic core of the lipid bilayer and usually spans the bilayer completely. Also known as an intrinsic membrane protein.

intercalating agent. A compound containing a planar ring structure that can fit between the stacked base pairs of DNA. Intercalating agents distort the DNA structure, partially unwinding the double helix.

intermediary metabolism. The metabolic reactions by which the small molecules of cells are interconverted.

intermediate density lipoprotein (IDL). A type of plasma lipoprotein that is formed during the breakdown of VLDLs.

intermediate filament. A structure composed of different protein subunits, found in the cytoplasm of most eukaryotic cells. Intermediate filaments are components of the cytoskeletal network.

intron. An internal nucleotide sequence that is removed from the primary RNA transcript during processing. The term intron also refers to the region of the gene that corresponds to the corresponding RNA intron (cf., exon).

inverted repeat. A sequence of nucleotides that is repeated in the opposite orientation within the same polynucleotide strand. An inverted repeat in double-stranded DNA can give rise to a cruciform structure.

ion pair. An electrostatic interaction between ionic groups of opposite charge within the interior of a macromolecule such as a globular protein.

ion product for water (K_w). The product of the concentrations of hydronium ions and hydroxide ions in an aqueous solution, equal to $1.0 \times 10^{-14}\ M^2$.

ion-exchange chromatography. A chromatographic technique used to separate a mixture of ionic species in solution, using a charged matrix. In anion-exchange chromatography, a positively charged matrix binds negatively charged solutes, and in cation-exchange chromatography, a negatively charged matrix binds positively charged solutes. The bound species can be serially eluted from the matrix by gradually changing the pH or increasing the salt concentration in the solvent.

ionophore. A compound that facilitates the diffusion of ions across bilayers and membranes by serving as a mobile ion carrier or by forming a channel for ion passage.

irreversible enzyme inhibition. A form of enzyme inhibition where the inhibitor binds covalently to the enzyme.

isoacceptor tRNA molecules. Different tRNA molecules that bind the same amino acid.

isoelectric focusing. A modified form of electrophoresis that uses buffers to create a pH gradient within a polyacrylamide gel. Each protein migrates to its isoelectric point (pI), that is, the pH in the gradient at which it no longer carries a net positive or negative charge.

isoelectric point (pI). The pH at which a zwitterionic molecule does not migrate in an electric field because its net charge is zero.

isoenzymes. *See* isozymes.

isomerase. An enzyme that catalyzes an isomerization reaction, a change in geometry or structure within one molecule.

isoprene. A branched, unsaturated five-carbon molecule that forms the basic structural unit of all isoprenoids, including the steroids and lipid vitamins.

isoprenoid. A lipid that is structurally related to isoprene.

isozymes. Different proteins from a single biological species that catalyze the same reaction. Also known as isoenzymes.

junk DNA. Regions of the genome with no known function.

K_a. *See* acid dissociation constant.

kb. *See* kilobase pair.

k_{cat}. *See* catalytic constant.

k_{cat}/K_m. The second-order rate constant for conversion of enzyme and substrate to enzyme and product at low substrate concentrations. The ratio of k_{cat} to K_m, when used to compare several substrates, is called the specificity constant.

K_{eq}. *See* equilibrium constant.

ketogenesis. The pathway that synthesizes ketone bodies from acetyl CoA in the mitochondrial matrix in mammals.

ketogenic compound. A compound, such as an amino acid, that can be degraded to form acetyl CoA and can thereby contribute to the synthesis of fatty acids or ketone bodies.

ketone bodies. Small molecules that are synthesized in the liver from acetyl CoA. During starvation, the ketone bodies β-hydroxybutyrate and acetoacetate become major metabolic fuels.

ketoses. A class of monosaccharides in which the most oxidized carbon atom, usually C-2, is ketonic.

K_j. *See* inhibition constant.

kilobase pair (kb). A unit of length of double-stranded DNA, equivalent to 1000 base pairs.

kinase. An enzyme that catalyzes transfer of a phosphoryl group to an acceptor molecule. A protein kinase catalyzes the phosphorylation of protein substrates. Kinases are also known as phosphotransferases.

kinetic mechanism. A scheme used to describe the sequence of steps in a multisubstrate enzyme-catalyzed reaction.

kinetic order. The sum of the exponents in a rate equation, which reflects how many molecules are reacting in the slowest step of the reaction. Also known as reaction order.

K_m. *See* Michaelis constant.

Krebs cycle. *See* citric acid cycle.

K_w. *See* ion product of water.

lagging strand. The newly synthesized DNA strand formed by discontinuous $5' \rightarrow 3'$ polymerization in the direction opposite replication fork movement.

lateral diffusion. The rapid motion of lipid or protein molecules within the plane of one leaflet of a lipid bilayer.

LDL. *See* low density lipoprotein.

leader peptide. The peptide encoded by a portion of the leader region of certain regulated operons. Synthesis of a leader peptide is the basis for regulating transcription of the entire operon by the mechanism of attenuation.

leader region. The sequence of nucleotides that lie between the transcription start site and the first coding region of an operon.

leading strand. The newly synthesized DNA strand formed by continuous $5' \rightarrow 3'$ polymerization in the same direction as replication fork movement.

leaflet. One layer of a lipid bilayer.

lectin. A plant protein that binds specific saccharides in glycoproteins.

leucine zipper. A structural motif found in DNA-binding proteins and other proteins. The zipper is formed when the hydrophobic faces (frequently containing leucine residues) of two amphipathic α-helices from the same or different polypeptide chains interact to form a coiled-coil structure.

LHC. *See* light-harvesting complex.

ligand. A molecule, group, or ion that binds noncovalently to another molecule or atom.

ligand-gated ion channel. A membrane ion channel that opens or closes in response to binding of a specific ligand.

ligase. An enzyme that catalyzes the joining, or ligation, of two substrates. Ligation reactions require the input of the chemical potential energy of a nucleoside triphosphate such as ATP. Ligases are commonly referred to as synthetases.

light reactions. The photosynthetic reactions in which protons derived from water are used in the chemiosmotic synthesis of ATP from ADP + P_i and a hydride ion from water reduces to NADPH. Also known as the light-dependent reactions.

light-harvesting complex (LHC). A large pigment complex in the thylakoid membrane that aids a photosystem in gathering light.

limit dextrin. A branched oligosaccharide derived from a glucose polysaccharide by the hydrolytic action of amylase or the phosphorolytic action of glycogen phosphorylase or starch phosphorylase. Limit dextrins are resistant to further degradation catalyzed by amylase or phosphorylase. Limit dextrins can be further degraded only after hydrolysis of the α-$(1 \rightarrow 6)$ linkages.

Lineweaver-Burk plot. *See* double-reciprocal plot.

linker DNA. The stretch of DNA (approximately 54 base pairs) between two adjacent nucleosome core particles.

lipase. An enzyme that catalyzes the hydrolysis of triacylglycerols.

lipid. A water-insoluble (or sparingly soluble) organic compound found in biological systems, which can be extracted by using relatively nonpolar organic solvents.

lipid bilayer. A double layer of lipids in which the hydrophobic tails associate with one another in the interior of the bilayer and the polar head groups face outward into the aqueous environment. Lipid bilayers are the structural basis of biological membranes.

lipid raft. A patch of membrane rich in cholesterol and sphingolipid.

lipid vitamin. A polyprenyl compound composed primarily of a long hydrocarbon chain or fused ring. Unlike water-soluble vitamins, lipid vitamins can be stored by animals. Lipid vitamins include vitamins A, D, E, and K.

lipid anchored membrane protein. A membrane protein that is tethered to a membrane through covalent linkage to a lipid molecule.

lipopolysaccharide. A macromolecule composed of lipid A (a disaccharide of phosphorylated glucosamine residues with attached fatty acids) and a polysaccharide. Lipopolysaccharides are found in the outer membrane of gram-negative bacteria. These compounds are released from bacteria undergoing lysis and are toxic to humans and other animals. Also known as an endotoxin.

lipoprotein. A macromolecular assembly of lipid and protein molecules with a hydrophobic core and a hydrophilic surface. Lipids are transported via lipoproteins.

liposome. A synthetic vesicle composed of a phospholipid bilayer that encloses an aqueous compartment.

loop. A nonrepetitive polypeptide region that connects secondary structures within a protein molecule and provides directional changes necessary for a globular protein to attain its compact shape. Loops contain from 2 to 16 residues. Short loops of up to 5 residues are often called turns.

low density lipoprotein (LDL). A type of plasma lipoprotein that is formed during the breakdown of IDLs and is enriched in cholesterol and cholesteryl esters.

lumen. The aqueous space enclosed by a biological membrane, such as the membrane of the endoplasmic reticulum or the thylakoid membrane.

lyase. An enzyme that catalyzes a nonhydrolytic or nonoxidative elimination reaction, or lysis, of a substrate, with the generation of a double bond. In the reverse direction, a lyase catalyzes addition of one substrate to a double bond of a second substrate.

lysophosphoglyceride. An amphipathic lipid that is produced when one of the two fatty acyl moieties of a glycerophospholipid is hydrolytically removed. Low concentrations of lysophosphoglycerides are metabolic intermediates, whereas high concentrations disrupt membranes, causing cells to lyse.

lysosome. A specialized digestive organelle in eukaryotic cells. Lysosomes contain a variety of enzymes that catalyze the breakdown of cellular biopolymers, such as proteins, nucleic acids, and polysaccharides, and the digestion of large particles, such as some bacteria ingested by the cell.

major groove. The wide groove on the surface of a DNA double helix created by the stacking of base pairs and the resulting twist in the sugar-phosphate backbones.

MALDI. *See* matrix-assisted laser desorption ionization.

mass action ratio (Q). The ratio of the concentrations of products to the concentrations of reactants of a reaction.

mass spectrometry. A technique that determines the mass of a molecule.

matrix. *See* mitochondrial matrix.

matrix-assisted laser desorption ionization (MALDI). A technique in mass spectrometry where the target molecule is released from a solid matrix by a laser beam.

maximum velocity (V_{max}). The initial velocity of a reaction when the enzyme is saturated with substrate, that is, when all the enzyme is in the form of an enzyme-substrate complex.

melting curve. A plot of the change in absorbance versus temperature for a DNA molecule. The change in absorbance indicates unfolding of the double helix.

melting point (T_m). The midpoint of the temperature range in which double-stranded DNA is converted to single-stranded DNA or a protein is converted from its native form to the denatured state.

membrane. A lipid bilayer containing associated proteins that serves to delineate and compartmentalize cells or organelles. Biological membranes are also the site of many important

biochemical processes related to energy transduction and intracellular signaling.

membrane-associated electron transport. *See* electron transport.

membrane potential ($\Delta\psi$). The charge separation across a membrane that results from differences in ionic concentrations on the two sides of the membrane.

messenger ribonucleic acid. *See* mRNA.

metabolic fuel. A small compound that can be catabolized to release energy. In multicellular organisms, metabolic fuels may be transported between tissues.

metabolically irreversible reaction. A reaction in which the value of the mass action ratio is two or more orders of magnitude smaller than the value of the equilibrium constant. The Gibbs free energy change for such a reaction is a large negative number; thus, the reaction is essentially irreversible.

metabolism. The sum total of biochemical reactions carried out by an organism.

metabolite. An intermediate in the synthesis or degradation of biopolymers and their component units.

metabolite channeling. Transfer of the product of one reaction of a multifunctional enzyme or a multienzyme complex directly to the next active site or enzyme without entering the bulk solvent. Channeling increases the rate of a reaction pathway by decreasing the transit time for an intermediate to reach the next enzyme and by producing high local concentrations of the intermediate.

metalloenzyme. An enzyme that contains one or more firmly bound metal ions. In some cases, such metal ions constitute part of the active site of the enzyme and are active participants in catalysis.

micelle. An aggregation of amphipathic molecules in which the hydrophilic portions of the molecules project into the aqueous environment and the hydrophobic portions associated with one another in the interior of the structure to minimize contact with water molecules.

Michaelis constant (K_m). The concentration of substrate that results in an initial velocity (v_0) equal to one-half the maximum velocity (V_{max}) for a given reaction.

Michaelis-Menten equation. A rate equation relating the initial velocity (v_0) of an enzymatic reaction to the substrate concentration ([S]), the maximum velocity (V_{max}), and the Michaelis constant (K_m).

microfilament. *See* actin filament.

microtubule. A protein filament composed of a and b tubulin heterodimers. Microtubules are components of the cytoskeletal network and can form structures capable of directed movement.

minor groove. The narrow groove on the surface of a DNA double helix created by the stacking of base pairs and the resulting twist in the sugar-phosphate backbones.

mismatch repair. Restoration of the normal nucleotide sequence in a DNA molecule containing mismatched bases. In mismatch repair, the correct strand is recognized, a portion of the incorrect strand is excised, and correctly base-paired, double-stranded DNA is synthesized by the actions of DNA polymerase and DNA ligase.

missense mutation. An alteration in DNA that involves the substitution of one nucleotide for another, resulting in a change in the amino acid specified by that codon.

mitochondrial matrix. The gel-like phase enclosed by the inner membrane of the mitochondrion. The mitochondrial matrix contains many enzymes involved in aerobic energy metabolism.

mitochondrion. An organelle that is the main site of oxidative energy metabolism in most eukaryotic cells. Mitochondria contain an outer and an inner membrane, the latter characteristically folded into cristae.

mixed inhibition. A form of enzyme inhibition where both K_m and V_{max} are affected.

molar mass. The weight in grams of one mole of a compound.

molecular chaperone. *See* chaperone.

molecular crowding. The decrease in diffusion rate that occurs when molecules collide with each other.

molecular weight. *See* relative molecular mass.

monocistronic mRNA. An mRNA molecule that encodes only a single polypeptide. Most eukaryotic mRNA molecules are monocistronic.

monomer. 1. A small compound that becomes a residue when polymerized with other monomers. **2.** A single subunit of a multisubunit protein.

monosaccharide. A simple sugar of three or more carbon atoms with the empirical formula $(CH_2O)_n$.

monounsaturated fatty acid. An unsaturated fatty acid with a single carbon-carbon double bond.

motif. A combination of secondary structure that appears in a number of different proteins. Also known as supersecondary structure.

M_r. *See* relative molecular mass.

mRNA. A class of RNA molecules that serve as templates for protein synthesis.

mRNA precursor. A class of RNA molecules synthesized by eukaryotic RNA polymerase II. mRNA precursors are processed posttranscriptionally to produce mature messenger RNA.

mucin. A high-molecular-weight *O*-linked glycoprotein containing as much as 80% carbohydrate by mass. Mucins are extended, negatively charged molecules that contribute to the viscosity of mucus, the fluid found on the surfaces of the gastrointestinal, genitourinary, and respiratory tracts.

multienzyme complex. An oligomeric protein that catalyzes several metabolic reactions.

mutagen. An agent that can cause DNA damage.

mutation. A heritable change in the sequence of nucleotides in DNA that causes a permanent alteration of genetic information.

near-equilibrium reaction. A reaction in which the value of the mass action ratio is close to the value of the equilibrium constant. The Gibbs free energy change for such a reaction is small; thus, the reaction is reversible.

Nernst equation. An equation that relates the observed change in reduction potential (ΔE) to the change in standard reduction potential ($\Delta E^{\circ\prime}$) of a reaction.

neutral phospholipids. Glycerophospholipids, such as phosphatidyl choline, having no net charge.

neutral solution. An aqueous solution that has a pH value of 7.0.

nick translation. The process in which DNA polymerase binds to a gap between the 3′ end of a nascent DNA chain and the 5′ end of the next RNA primer, catalyzes hydrolytic removal of ribonucleotides using $5' \rightarrow 3'$ exonuclease activity, and replaces them with deoxyribonucleotides using $5' \rightarrow 3'$ polymerase activity.

nitrogen cycle. The flow of nitrogen from N_2 to nitrogen oxides ($NO_2^{\ominus}$ and $NO_3^{\ominus}$) ammonia, nitrogenous biomolecules, and back to N_2.

nitrogen fixation. The reduction of atmospheric nitrogen to ammonia. Biological nitrogen fixation occurs in only a few species of bacteria and algae.

***N*-linked oligosaccharide.** An oligosaccharide chain attached to a protein through covalent bonds to the amide nitrogen atom of side chain of asparagine residues. The oligosaccharide chains of *N*-linked glycoproteins contain a core pentasaccharide of two *N*-acetylglucosamine residues and three mannose residues.

NMR spectroscopy. *See* nuclear magnetic resonance spectroscopy.

noncompetitive inhibition. Inhibition of an enzyme-catalyzed reaction by a reversible inhibitor that binds to either the enzyme or the enzyme- substrate complex.

nonessential amino acid. An amino acid that an animal can produce in sufficient quantity to meet metabolic needs.

nonhomologous recombination. Recombination between unrelated sequences that do not share significant sequence similarity.

nonrepetitive structure. An element of protein structure in which consecutive residues do not have a single repeating conformation.

nonsense mutation. An alteration in DNA that involves the substitution of one nucleotide for another, changing a codon that specifies an amino acid to a termination

codon. A nonsense mutation results in premature termination of a protein's synthesis.

***N*-terminus.** The amino acid residue bearing a free α-amino group at one end of a peptide chain. In some proteins, the *N*-terminus is blocked by acylation. The *N*-terminal residue is usually assigned the residue number 1. Also known as the amino terminus.

nuclear envelope. The double membrane that surrounds the nucleus and contains protein-lined nuclear pore complexes that regulate the import and export of material to and from the nucleus. The outer membrane of the nuclear envelope is continuous with the endoplasmic reticulum; the inner membrane is lined with filamentous proteins, constituting the nuclear lamina.

nuclear magnetic resonance spectroscopy (NMR spectroscopy). A technique used to study the structures of molecules in solution. In nuclear magnetic resonance spectroscopy, the absorption of electromagnetic radiation by molecules in magnetic fields of varying frequencies is used to determine the spin states of certain atomic nuclei.

nuclease. An enzyme that catalyzes hydrolysis of the phosphodiester linkages of a polynucleotide chain. Nucleases can be classified as endonucleases and exonucleases.

nucleic acid. A polymer composed of nucleotide residues linked in a linear sequence by 3′–5′ phosphodiester linkages. DNA and RNA are nucleic acids composed of deoxyribonucleotide residues and ribonucleotide residues, respectively.

nucleoid region. The region within a prokaryotic cell that contains the chromosome.

nucleolus. The region of the eukaryotic nucleus where rRNA transcripts are processed and ribosomes are assembled.

nucleophile. An electron-rich species that is negatively charged or contains unshared electron pairs and is attracted to chemical species that are positively charged or electron-deficient (electrophiles).

nucleophilic substitution. A reaction in which one nucleophile (e.g., $Y^{\ominus}$) displaces another (e.g., $X^{\ominus}$).

nucleoside. A purine or pyrimidine *N*-glycoside of ribose or deoxyribose.

nucleosome. A DNA-protein complex that forms the fundamental unit of chromatin. A nucleosome consists of a nucleosome core particle (approximately 146 base pairs of DNA plus a histone octamer), linker DNA (approximately 54 base pairs), and histone H1 (which binds the core particle and linker DNA).

nucleosome core particle. A DNA-protein complex composed of approximately 146 base pairs of DNA wrapped around an octamer of histones (two each of H2A, H2B, H3, and H4).

nucleotide. The phosphate ester of a nucleoside, consisting of a nitrogenous base linked to a pentose phosphate. Nucleotides are the monomeric units of nucleic acids.

nucleus. An organelle that contains the principal genetic material of eukaryotic cells and functions as the major site of RNA synthesis and processing.

obligate aerobe. An organism that requires the presence of oxygen for survival.

obligate anaerobe. An organism that requires an oxygen-free environment for survival.

Okazaki fragments. Relatively short strands of DNA that are produced during discontinuous synthesis of the lagging strand of DNA.

oligomer. A multisubunit molecule whose arrangement of subunits always has a defined stoichiometry and almost always displays symmetry.

oligonucleotide. A polymer of several (up to about 20) nucleotide residues linked by phosphodiester bonds.

oligopeptide. A polymer of several (up to about 20) amino acid residues linked by peptide bonds.

oligosaccharide. A polymer of 2 to about 20 monosaccharide residues linked by glycosidic bonds.

oligosaccharide processing. The enzyme-catalyzed addition and removal of saccharide residues during the maturation of a glycoprotein.

***O*-linked oligosaccharide.** An oligosaccharide attached to a protein through a covalent bond to the hydroxyl oxygen atom of a serine or threonine residue.

open reading frame. A stretch of nucleotide triplets that contains no termination codons. Protein-encoding regions are examples of open reading frames.

operator. A DNA sequence to which a specific repressor protein binds, thereby blocking transcription of a gene or operon.

operon. A bacterial transcriptional unit consisting of several different coding regions cotranscribed from one promoter.

ordered sequential reaction. A reaction in which both the binding of substrates to an enzyme and the release of products from the enzyme follow an obligatory order.

organelle. Any specialized membrane-bounded structure within a eukaryotic cell. Organelles are uniquely organized to perform specific functions.

origin of replication. A DNA sequence at which replication is initiated.

osmosis. The movement of solvent molecules from a less concentrated solution to an adjacent, more concentrated solution.

osmotic pressure. The pressure required to prevent the flow of solvent from a less concentrated solution to a more concentrated solution.

oxidase. An enzyme that catalyzes an oxidation-reduction reaction in which O_2 is the electron acceptor. Oxidases are members of the IUBMB class of enzymes known as oxidoreductases.

oxidation. The loss of electrons from a substance through transfer to another substance (the oxidizing agent). Oxidations can take several forms, including the addition of oxygen to a compound, the removal of hydrogen from a compound to create a double bond, or an increase in the valence of a metal ion.

oxidative phosphorylation. *See* electron transport.

oxidizing agent. A substance that accepts electrons in an oxidation-reduction reaction and thereby becomes reduced.

oxidoreductase. An enzyme that catalyzes an oxidation-reduction reaction. Some oxidoreductases are known as dehydrogenases, oxidases, peroxidases, oxygenases, or reductases.

oxygenation. The reversible binding of oxygen to a macromolecule.

Δp. *See* protonmotive force.

PAGE. *See* polyacrylamide gel electrophoresis.

passive transport. The process by which a solute specifically binds to a transport protein and is transported across a membrane, moving with the solute concentration gradient. Passive transport occurs without the expenditure of energy. Also known as facilitated diffusion.

Pasteur effect. The slowing of glycolysis in the presence of oxygen.

pathway. A sequence of metabolic reactions.

pause site. A region of a gene where transcription slows. Pausing is exaggerated at palindromic sequences, where newly synthesized RNA can form a hairpin structure.

PCR. *See* polymerase chain reaction.

pentose phosphate pathway. A pathway by which glucose 6-phosphate is metabolized to generate NADPH and ribose 5-phosphate. In the oxidative stage of the pathway, glucose 6-phosphate is converted to ribulose 5-phosphate and CO_2 rating two molecules of NADPH. In the nonoxidative stage, ribulose 5-phosphate can be isomerized to ribose 5-phosphate or converted to intermediates of glycolysis. Also known as the hexose monophosphate shunt.

peptide. Two or more amino acids covalently joined in a linear sequence by peptide bonds.

peptide bond. The covalent secondary amide linkage that joins the carbonyl group of one amino acid residue to the amino nitrogen of another in peptides and proteins.

peptide group. The nitrogen and carbon atoms involved in a peptide bond and their four substituents: the carbonyl oxygen atom, the amide hydrogen atom, and the two adjacent α-carbon atoms.

peptidoglycan. A macromolecule containing a heteroglycan chain of alternating *N*-acetylglucosamine and *N*-acetylmuramic acid cross-linked to peptides of varied composition. Peptidoglycans are the major components of the cell walls of many bacteria.

peptidyl site. *See* P site.

peptidyl transferase. The enzymatic activity responsible for the formation of a peptide bond during protein synthesis.

peptidyl-tRNA. The tRNA molecule to which the growing peptide chain is attached during protein synthesis.

peripheral membrane protein. A membrane protein that is weakly bound to the interior or exterior surface of a membrane through ionic interactions and hydrogen bonding with the polar heads of the membrane lipids or with an integral membrane protein. Also known as an extrinsic membrane protein.

periplasmic space. The region between the plasma membrane and the cell wall in bacteria.

permeability coefficient. A measure of the ability of an ion or small molecule to diffuse across a lipid bilayer.

peroxisome. An organelle in all animal and many plant cells that carries out oxidation reactions, some of which produce the toxic compound hydrogen peroxide (H_2O_2). Peroxisomes contain the enzyme catalase, which catalyzes the breakdown of toxic H_2O_2 to water and O_2.

pH. A logarithmic quantity that indicates the acidity of a solution, that is, the concentration of hydronium ions in solution. pH is defined as the negative logarithm of the hydronium ion concentration.

pH optimum. In an enzyme-catalyzed reaction, the pH at the point of maximum catalytic activity.

phage. *See* bacteriophage.

phase-transition temperature (T_m). The midpoint of the temperature range in which lipids or other macromolecular aggregates are converted from a highly ordered phase or state (such as a gel) to a less-ordered state (such as a liquid crystal).

ϕ (phi). The angle of rotation around the bond between the α-carbon and the nitrogen of a peptide group.

phosphagen. A "high energy" phosphate storage molecule found in animal muscle cells. Phosphagens are phosphoamides and have a higher phosphoryl-group-transfer potential than ATP.

phosphatase. An enzyme that catalyzes the hydrolytic removal of a phosphoryl group.

phosphatidate. A glycerophospholipid that consists of two fatty acyl groups esterified to C-1 and C-2 of glycerol 3-phosphate. Phosphatidates are metabolic intermediates in the biosynthesis or breakdown of more complex glycerophospholipids.

phosphoanhydride. A compound formed by condensation of two phosphate groups.

phosphodiester linkage. A linkage in nucleic acids and other molecules in which two alcoholic hydroxyl groups are joined through a phosphate group.

phosphoester linkage. The bond by which a phosphoryl group is attached to an alcoholic or phenolic oxygen.

phospholipid. A lipid containing a phosphate moiety.

phosphorolysis. Cleavage of a bond within a molecule by group transfer to an oxygen atom of phosphate.

phosphorylase. An enzyme that catalyzes the cleavage of its substrate(s) via nucleophilic attack by inorganic phosphate (P_i) (i.e., via phosphorolysis).

phosphorylation. A reaction involving the addition of a phosphoryl group to a molecule.

phosphoryl group transfer potential. A measure of the ability of a compound to transfer a phosphoryl group to another compound. Under standard conditions, group transfer potentials have the same values as the standard free energies of hydrolysis but are opposite in sign.

photoautotroph. A photosynthetic organism that can utilize CO_2 as its main carbon source.

photon. A quantum of light energy.

photophosphorylation. The light-dependent formation of ATP from ADP and P_i catalyzed by chloroplast ATP synthase.

photoheterotroph. Photosynthetic organism that requires organic molecules as a carbon source.

photoreactivation. The direct repair of damaged DNA by an enzyme that is activated by visible light.

photorespiration. The light-dependent uptake of O_2 and the subsequent metabolism of phosphoglycolate that occurs primarily in C_3 photosynthetic plants. Photorespiration can occur because O_2 competes with CO_2 for the active site of ribulose 1,5-*bis*phosphate carboxylase-oxygenase, the enzyme that catalyzes the first step of the reductive pentose phosphate cycle.

photosynthesis. The conversion of light energy (photons) to chemical energy in the form of ATP and/or NADPH.

photosystem. A functional unit of the light-dependent electron-transfer reactions of photosynthesis. Each membrane-embedded photosystem contains a reaction center, which forms the core of the photosystem, and a pool of light-absorbing antenna pigments.

phototroph. An organism that can convert light energy into chemical potential energy (i.e., an organism capable of photosynthesis).

physiological pH. The normal pH of human blood, which is 7.4.

pI. *See* isoelectric point.

ping-pong reaction. A reaction in which an enzyme binds one substrate and releases a product, leaving a substituted enzyme that then binds a second substrate and releases a second product, thereby restoring the enzyme to its original form.

pitch. The axial distance for one complete turn of a helical structure.

pK_a. A logarithmic value that indicates the strength of an acid. pK_a is defined as the negative logarithm of the acid dissociation constant, K_a.

plasma membrane. The membrane that surrounds the cytoplasm of a cell and thus defines the perimeter of the cell.

plasmalogen. A glycerophospholipid that has a hydrocarbon chain linked to C-1 of glycerol 3-phosphate through a vinyl ether linkage. Plasmalogens are found in the central nervous system and in peripheral nerve and muscle tissue.

plasmid. A relatively small, extrachromosomal DNA molecule that is capable of autonomous replication. Plasmids are usually closed, circular, double-stranded DNA molecules.

P:O ratio. The ratio of molecules of ADP phosphorylated to atoms of oxygen reduced during oxidative phosphorylation.

polar. Having uneven distribution of charge. A molecule or functional group is polar if its center of negative charge does not coincide with its center of positive charge.

poly A tail. A stretch of polyadenylate, up to 250 nucleotide residues long, that is added to the 3′ end of a eukaryotic mRNA molecule following transcription.

polyacrylamide gel electrophoresis (PAGE). A technique used to separate molecules of different net charge and/or size based on their migration through a highly cross-linked gel matrix in an electric field.

polycistronic mRNA. An mRNA molecule that contains multiple coding regions. Many prokaryotic mRNA molecules are polycistronic.

polymerase chain reaction (PCR). A method for amplifying the amount of DNA in a sample and for enriching a particular DNA sequence in a population of DNA molecules. In the polymerase chain reaction, oligonucleotides complementary to the ends of the desired DNA sequence are used as primers for multiple rounds of DNA synthesis.

polynucleotide. A polymer of many (usually more than 20) nucleotide residues linked by phosphodiester bonds.

polypeptide. A polymer of many (usually more than 20) amino acid residues linked by peptide bonds.

polyribosome. *See* polysome.

polysaccharide. A polymer of many (usually more than 20) monosaccharide residues linked by glycosidic bonds. Polysaccharide chains can be linear or branched.

polysome. The structure formed by the binding of many translation complexes to a large mRNA molecule. Also known as a polyribosome.

polyunsaturated fatty acid. An unsaturated fatty acid with two or more carbon-carbon double bonds.

pore. *See* channel.

posttranscriptional processing. RNA processing that occurs after transcription is complete.

posttranslational modification. Covalent modification of a protein that occurs after synthesis of the polypeptide is complete.

prenylated protein. A lipid-anchored protein that is covalently linked to an isoprenoid moiety via the sulfur atom of a cysteine residue at the *C*-terminus of the protein.

primary structure. The sequence in which residues are covalently linked to form a polymeric chain.

primary transcript. A newly synthesized RNA molecule before processing.

primase. An enzyme in the primosome that catalyzes the synthesis of short pieces of RNA about 10 residues long. These oligonucleotides are the primers for synthesis of Okazaki fragments.

primosome. A multiprotein complex, including primase and helicase in *E. coli*, that catalyzes the synthesis of the short RNA primers needed for discontinuous DNA synthesis of the lagging strand.

processive enzyme. An enzyme that remains bound to its growing polymeric product through many polymerization steps (cf., distributive enzyme).

prochiral atom. An atom with multiple substituents, two of which are identical. A prochiral atom can become chiral when one of the identical substituents is replaced.

prokaryote. An organism, usually a single cell, which contains no nucleus or internal membranes (cf., eukaryote).

promoter. The region of DNA where RNA polymerase binds during transcription initiation.

prostaglandin. An eicosanoid that has a cyclopentane ring. Prostaglandins are metabolic regulators that act in the immediate neighborhood of the cells in which they are produced.

prosthetic group. A coenzyme that is tightly bound to an enzyme. A prosthetic group, unlike a cosubstrate, remains bound to a specific site of the enzyme throughout the catalytic cycle of the enzyme.

protease. An enzyme that catalyzes hydrolysis of peptide bonds. The physiological substrates of proteases are proteins.

protein. A biopolymer consisting of one or more polypeptide chains. The biological function of each protein molecule depends not only on the sequence of covalently linked amino acid residues, but also on its three-dimensional structure (conformation).

protein coenzyme. A protein that does not itself catalyze reactions but is required for the action of certain enzymes.

protein glycosylation. The covalent addition of carbohydrate to proteins. In *N*-glycosylation, the carbohydrate is attached to the amide group of the side chain of an asparagine residue. In *O*-glycosylation, the carbohydrate is attached to the hydroxyl group of the side chain of a serine or threonine residue.

protein kinase. *See* kinase.

protein phosphatase. *See* phosphatase.

proteoglycan. A complex of protein with glycosaminoglycan chains covalently bound through their anomeric carbon atoms. Up to 95% of the mass of a proteoglycan may be glycosaminoglycan.

proteomics. The study of all proteins produced in a certain cell type, tissue, organ, or organism.

protonmotive force (Δp). The energy stored in a proton concentration gradient across a membrane.

proximity effect. The increase in the rate of a nonenzymatic or enzymatic reaction attributable to high effective concentrations of reactants, which result in more frequent formation of transition states.

pseudo first-order reaction. A multi-reactant reaction carried out under conditions where the rate depends on the concentration of only one reactant.

pseudogene. A nonexpressed sequence of DNA that evolved from a protein-encoding gene. Pseudogenes often contain mutations in their coding regions and cannot produce functional proteins.

ψ (psi). The angle of rotation around the bond between the α-carbon and the carbonyl carbon of a peptide group.

$\Delta\psi$. *See* membrane potential.

P site. Peptidyl site. The site on a ribosome that is occupied during protein synthesis by a tRNA molecule attached to the growing polypeptide chain (peptidyl tRNA).

purine. A nitrogenous base having a two-ring structure in which a pyrimidine is fused to imidazole. Adenine and guanine are substituted purines found in both DNA and RNA.

pyranose. A monosaccharide structure that forms a six-membered ring as a result of intramolecular hemiacetal formation.

pyrimidine. A nitrogenous base having a heterocyclic ring that consists of four carbon atoms and two nitrogen atoms. Cytosine, thymine, and uracil are substituted pyrimidines found in nucleic acids (cytosine in DNA and RNA, uracil in RNA, and thymine principally in DNA).

***Q*.** *See* mass action ratio.

Q cycle. A cyclic pathway proposed to explain the sequence of electron transfers and proton movements within Complex III of mitochondria or the cytochrome *bf* complex in chloroplasts. The net result of the two steps of the Q cycle is oxidation of two molecules of QH_2 or plastoquinol (PQH_2); formation of one molecule of QH_2 or PQH_2; transfer of two electrons; and net translocation of four protons across the inner mitochondrial membrane to the intermembrane space or across the thylakoid membrane to the lumen.

quaternary structure. The organization of two or more polypeptide chains within a multisubunit protein.

R state. The more active conformation of an allosteric protein; opposite of T state.

Ramachandran plot. A plot of c versus f values for amino acid residues in a polypeptide chain. Certain f and c values are characteristic of different conformations.

random sequential reaction. A reaction in which neither the binding of substrates to an enzyme nor the release of products from the enzyme follows an obligatory order.

rate acceleration. The ratio of the rate constant for a reaction in the presence of enzyme (k_{cat}) divided by the rate constant for that reaction in the absence of enzyme (k_n). The rate acceleration value is a measure of the efficiency of an enzyme.

rate equation. An expression of the observed relationship between the velocity of a reaction and the concentration of each reactant.

rate determining step. The slowest step in a chemical reaction. The rate determining step has the highest activation energy among the steps leading to formation of a product from the substrate.

reaction center. A complex of proteins, electron transport cofactors, and a special pair of chlorophyll molecules that forms the core of a photosystem. The reaction center is the site of conversion of photochemical energy to electrochemical energy during photosynthesis.

reaction mechanism. The step-by-step atomic or molecular events that occur during chemical reactions.

reaction order. *See* kinetic order.

reaction specificity. The lack of formation of wasteful by-products by an enzyme. Reaction specificity results in essentially 100% product yields.

reactive center. The part of a coenzyme to which mobile metabolic groups are attached.

reading frame. The sequence of nonoverlapping codons of an mRNA molecule that specifies the amino acid sequence. The reading frame of an mRNA molecule is determined by the position where translation begins; usually an AUG codon.

receptor. A protein that binds a specific ligand, such as a hormone, leading to some cellular response.

recombinant DNA. A DNA molecule that includes DNA from different sources.

recombination. *See* genetic recombination.

reducing agent. A substance that loses electrons in an oxidation-reduction reaction and thereby becomes oxidized.

reducing end. The residue containing a free anomeric carbon in a polysaccharide. A polysaccharide usually contains no more than one reducing end.

reduction. The gain of electrons by a substance through transfer from another substance (the reducing agent). Reductions can take several forms, including the loss of oxygen from a compound, the addition of

hydrogen to a double bond of a compound, or a decrease in the valence of a metal ion.

reduction potential (*E*). A measure of the tendency of a substance to reduce other substances. The more negative the reduction potential, the greater the tendency to donate electrons.

regulated enzyme. An enzyme located at a critical point within one or more metabolic pathways, whose activity may be increased or decreased based on metabolic demand. Most regulated enzymes are oligomeric.

regulatory protein. A protein that is involved in the regulation of gene expression, usually at the point of transcription initiation. Repressors and activators are examples of regulatory proteins.

regulatory site. A ligand-binding site in a regulatory enzyme distinct from the active site. Allosteric modulators alter enzyme activity by binding to the regulatory site. Also known as an allosteric site.

relative molecular mass (M_r). The mass of a molecule relative to 1/12th the mass of ^{12}C. There are no units associated with the values for relative molecular mass.

release factor. A protein involved in terminating protein synthesis.

renaturation. The restoration of the native conformation of a biological macromolecule, usually resulting in restoration of biological activity.

replication. The duplication of double-stranded DNA, during which parental strands separate and serve as templates for synthesis of new strands. Replication is carried out by DNA polymerase and associated factors.

replication fork. The Y-shaped junction where double-stranded, template DNA is unwound and new DNA strands are synthesized during replication.

replisome. A multiprotein complex that includes DNA polymerase, primase, helicase, single-strand binding protein, and additional components. The replisomes, located at each of the replication forks, carry out the polymerization reactions of bacterial chromosomal DNA replication.

repressor. A regulatory DNA-binding protein that prevents transcription by RNA polymerase.

residue. A single component within a polymer. The chemical formula of a residue is that of the corresponding monomer minus the elements of water.

resonance energy transfer. A form of excitation energy transfer between molecules that does not involve transfer of an electron.

respiratory electron transport chain. A series of enzyme complexes and associated cofactors that are electron carriers, passing electrons from reduced coenzymes or substrates to molecular oxygen (O_2), the terminal electron acceptor of aerobic metabolism.

restriction endonuclease. An endonuclease that catalyzes the hydrolysis of double-stranded DNA at a specific nucleotide sequence. Type I restriction endonucleases catalyze both the methylation of host DNA and the cleavage of nonmethylated DNA, whereas type II restriction endonucleases catalyze only the cleavage of nonmethylated DNA.

restriction map. A diagram showing the size and arrangement of fragments produced from a DNA molecule by the action of various restriction endonucleases.

reverse transcriptase. A type of DNA polymerase that catalyzes the synthesis of a strand of DNA from an RNA template.

reverse turn. *See* turn.

ribonucleic acid (RNA). A polymer consisting of ribonucleotide residues joined by 3′–5′ phosphodiester bonds. The sugar moiety in RNA is ribose. Genetic information contained in DNA is transcribed in the synthesis of RNA, some of which (mRNA) is translated in the synthesis of protein.

ribonucleoprotein. A complex containing both ribonucleic acid and protein.

ribosome. A large ribonucleoprotein complex composed of multiple ribosomal RNA molecules and proteins. Ribosomes are the site of protein synthesis.

ribozyme. An RNA molecule with enzymatic activity.

rise. The distance between one residue and the next along the axis of a helical macromolecule.

RNA processing. The reactions that transform a primary RNA transcript into a mature RNA molecule. The three general types of RNA processing include the removal of RNA nucleotides from primary transcripts, the addition of RNA nucleotides not encoded by the gene, and the covalent modification of bases.

rRNA. *See* ribosomal ribonucleic acid.

S. *See* Svedberg unit.

S. *See* entropy.

salt bridge. *See* charge-charge interactions.

salvage pathway. A pathway in which a major metabolite, such as a purine or pyrimidine nucleotide, can be synthesized from a preformed molecular entity, such as a purine or pyrimidine.

saturated fatty acid. A fatty acid that does not contain a carbon-carbon double bond.

Schiff base. A complex formed by the reversible condensation of a primary amine with an aldehyde (to form an aldimine) or a ketone (to form a ketimine).

SDS-PAGE. *See* sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

second messenger. A compound that acts intracellularly in response to an extracellular signal.

secondary structure. The regularities in local conformations within macromolecules. In proteins, secondary structure is maintained by hydrogen bonds between carbonyl and amide groups of the backbone. In nucleic acids, secondary structure is maintained by hydrogen bonds and stacking interactions between the bases.

second-order reaction. A reaction whose rate depends on the concentrations of two reactants.

self-splicing intron. An intron that is excised in a reaction mediated by the RNA precursor itself.

sense strand. In double-stranded DNA the sense strand is the strand that contains codons. Also called the coding strand. The opposite strand is called the antisense strand or the template strand.

sequential reaction. An enzymatic reaction in which all the substrates must be bound to the enzyme before any product is released.

sequential theory of cooperativity and allosteric regulation. A model of the cooperative binding of identical ligands to oligomeric proteins. According to the simplest form of the sequential theory, the binding of a ligand may induce a change in the tertiary structure of the subunit to which it binds and may alter the conformations of neighboring subunits to varying extents. Only one subunit conformation has a high affinity for the ligand. Also known as the ligand-induced theory.

Shine-Dalgarno sequence. A purine-rich region just upstream of the initiation codon in prokaryotic mRNA molecules. The Shine-Dalgarno sequence binds to a pyrimidine-rich sequence in the ribosomal RNA, thereby positioning the ribosome at the initiation codon.

σ factor. *See* σ subunit.

σ subunit (sigma subunit). A subunit of prokaryotic RNA polymerase, which acts as a transcription initiation factor by binding to the promoter. Different σ subunits are specific for different promoters. Also known as a σ factor.

signal peptidase. An integral membrane protein of the endoplasmic reticulum that catalyzes cleavage of the signal peptide of proteins translocated to the lumen.

signal peptide. The *N*-terminal sequence of residues in a newly synthesized polypeptide that targets the protein for translocation across a membrane.

signal transduction. The process whereby an extracellular signal is converted to an intracellular signal by the action of a membrane-associated receptor, a transducer, and an effector enzyme.

signal recognition particle (SRP). A eukaryotic protein-RNA complex that binds a newly synthesized peptide as it is extruded from the ribosome. The signal-recognition particle is involved in anchoring the ribosome to the cytosolic face of the endoplasmic

reticulum so that protein translocation to the lumen can occur.

single-strand binding protein (SSB). A protein that binds tightly to single-stranded DNA, preventing the DNA from folding back on itself to form double-stranded regions.

site-directed mutagenesis. An in vitro procedure by which one particular nucleotide residue in a gene is replaced by another, resulting in production of an altered protein sequence.

site-specific recombination. An example of recombination that occurs at specific sites in the genome.

small nuclear ribonucleoprotein (snRNP). An RNA-protein complex composed of one or two specific snRNA molecules plus a number of proteins. snRNPs are involved in splicing mRNA precursors and in other cellular events.

small RNA. A class of RNA molecules. Some small RNA molecules have catalytic activity. Some small nuclear RNA molecules (snRNA) are components of small nuclear ribonucleoproteins (snRNPs).

snRNA. *See* small nuclear RNA.

snRNP. *See* small nuclear ribonucleoprotein.

sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). Polyacrylamide gel electrophoresis performed in the presence of the detergent sodium dodecyl sulfate. SDS-PAGE allows separation of proteins on the basis of size only rather than charge and size.

solvation. A state in which a molecule or ion is surrounded by solvent molecules.

solvation sphere. The shell of solvent molecules that surrounds an ion or solute.

special pair. A specialized pair of chlorophyll molecules in reaction centers that is the primary electron donor during the light-dependent reactions of photosynthesis.

specific heat. The amount of heat required to raise the temperature of 1 gram of a substance by 1°C.

specificity constant. *See* k_{cat}/K_m.

sphingolipid. An amphipathic lipid with a sphingosine (trans-4-sphingenine) backbone. Sphingolipids, which include sphingomyelins, cerebrosides, and gangliosides, are present in plant and animal membranes and are particularly abundant in the tissues of the central nervous system.

sphingomyelin. A sphingolipid that consists of phosphocholine attached to the C-1 hydroxyl group of a ceramide. Sphingomyelins are present in the plasma membranes of most mammalian cells and are a major component of myelin sheaths.

splice site. The conserved nucleotide sequence surrounding an exon-intron junction. It includes the site where the RNA molecule is cleaved during intron excision.

spliceosome. The large protein-RNA complex that catalyzes the removal of introns from mRNA precursors. The spliceosome is composed of small nuclear ribonucleoproteins.

splicing. The process of removing introns and joining exons to form a continuous RNA molecule.

SRP. *See* signal recognition particle.

SSB. *See* single-strand binding protein.

stacking interactions. The weak noncovalent forces between adjacent bases or base pairs in single-stranded or double-stranded nucleic acids, respectively. Stacking interactions contribute to the helical shape of nucleic acids.

standard Gibbs free energy change ($\Delta G°'$). The free energy change for a reaction under biochemical standard state conditions.

standard reduction potential ($E°'$). A measure of the tendency of a substance to reduce other substances under biochemical standard state conditions.

standard state. A set of reference conditions for a chemical reaction. In biochemistry, the standard state is defined as a temperature of 298 K (25°C), a pressure of 1 atmosphere, a solute concentration of 1.0 M, and a pH of 7.0.

starch. A homopolymer of glucose residues that is a storage polysaccharide in plants. There are two forms of starch: amylose, an unbranched polymer of glucose residues joined by α-(1 → 4) linkages; and amylopectin, a branched polymer of glucose residues joined by α-(1 → 4) linkages with α-(1 → 6) linkages at branch points.

steady state. A state in which the rate of synthesis of a compound is equal to its rate of utilization or degradation.

stem-loop. *See* hairpin.

stereoisomers. Compounds with the same molecular formula but different spatial arrangements of their atoms.

stereospecificity. The ability of an enzyme to recognize and act upon only a single stereoisomer of a substrate.

steroid. A lipid containing a fused, four-ring isoprenoid structure.

sterol. A steroid containing a hydroxyl group.

stomata. Structures on the surface of a leaf through which carbon dioxide diffuses directly into photosynthetic cells.

stop codon. *See* termination codon.

strand invasion. The exchange of single strands of DNA from two nicked molecules having homologous nucleotide sequences.

stroma. The interior of a chloroplast corresponding to the cytoplasm of the ancestral cyanobacterium.

stromal lamellae. Regions of the thylakoid membrane that are in contact with the stroma.

substrate. A reactant in a chemical reaction. In enzymatic reactions, substrates are specifically acted upon by enzymes, which catalyze the conversion of substrates to products.

substrate cycle. A pair of opposing reactions that catalyzes a cycle between two pathway intermediates.

substrate level phosphorylation. Phosphorylation of a nucleoside diphosphate by transfer of a phosphoryl group from a nonnucleotide substrate.

supercoil. A topological arrangement assumed by over- or underwound double-stranded DNA. Underwinding gives rise to negative supercoils; overwinding produces positive supercoils.

supersecondary structure. *See* motif.

Svedberg unit (S). A unit of 10^{-13} second used for expressing the sedimentation coefficient, a measure of the rate at which a large molecule or particle sediments in an ultracentrifuge. Large S values usually indicate large masses.

symport. The cotransport of two different species of ions or molecules in the same direction across a membrane by a transport protein.

synonymous codons. Different codons that specify the same amino acid.

synthase. A common name for an enzyme, often a transferase, that catalyzes a synthetic reaction.

synthetase. An enzyme that catalyzes the joining of two substrates and requires the input of the chemical potential energy of a nucleoside triphosphate. Synthetases are members of the IUBMB class of enzymes known as ligases.

T state. The less active conformation of an allosteric protein; opposite of R state.

TATA box. An A/T-rich DNA sequence found within the promoter of both prokaryotic and eukaryotic genes.

template strand. The strand of DNA within a gene whose nucleotide sequence is complementary to that of the transcribed RNA. During transcription, RNA polymerase binds to and moves along the template strand in the 3′ → 5′ direction, catalyzing the synthesis of RNA in the 5′ → 3′ direction.

termination codon. A codon that is recognized by specific proteins that cause newly synthesized peptides to be released from the translation machinery thus terminating translation. The three termination codons (UAG, UAA, and UGA) are also known as stop codons.

termination sequence. A sequence at the 3′ end of a gene that mediates transcription termination.

tertiary structure. The compacting of polymeric chains into one or more domains within a macromolecule. In proteins, tertiary structure is stabilized mainly by hydrophobic interactions between side chains.

thermodynamics. The branch of physical science that studies transformations of heat and energy.

30 nm fiber. A chromatin structure in which nucleosomes are coiled into a solenoid 30 nm in diameter.

−35 region. A sequence found within the promoter of some prokaryotic genes about 30 to 35 base pairs upstream of the transcription initiation site.

3_{10} helix. A secondary structure of proteins, consisting of a helix in which the carbonyl oxygen of each amino acid residue (residue *n*) forms a hydrogen bond with the amide hydrogen of the third residue further toward the *C*-terminus of the polypeptide chain (residue $n + 3$).

thylakoid lamella. *See* thylakoid membrane.

thylakoid membrane. A highly folded, continuous membrane network suspended in the aqueous matrix of the chloroplast. The thylakoid membrane is the site of the light-dependent reactions of photosynthesis, which lead to the formation of NADPH and ATP. Also known as the thylakoid lamella.

T_m. *See* melting point and phase-transition temperature.

topoisomerase. An enzyme that alters the supercoiling of a DNA molecule by cleaving a phosphodiester linkage in either one or both strands, rewinding the DNA, and resealing the break. Some topoisomerases are also known as DNA gyrases.

topology. 1. The arrangement of membrane-spanning segments and connecting loops in an integral membrane protein. **2.** The overall morphology of a nucleic acid molecule.

TψC arm. The stem-and-loop structure in a tRNA molecule that contains the sequence ribothymidylate–pseudouridylate–cytidylate (TψC).

trace element. An element required in very small quantities by living organisms. Examples include copper, iron, and zinc.

transaminase. An enzyme that catalyzes the transfer of an amino group from an α-amino acid to an α-keto acid. Transaminases require the coenzyme pyridoxal phosphate. They are also called aminotransferases.

transcription. The copying of biological information from a double-stranded DNA molecule to a single-stranded RNA molecule, catalyzed by a transcription complex consisting of RNA polymerase and associated factors.

transcription bubble. A short region of double-stranded DNA that is unwound by RNA polymerase during transcription.

transcription factor. A protein that binds to the promoter region, to RNA polymerase, or to both during assembly of the transcription initiation complex. Some transcription factors remain bound during RNA chain elongation.

transcription initiation complex. The complex of RNA polymerase and other factors that assembles at the promoter at the start of transcription.

transcriptional activator. A regulatory DNA-binding protein that enhances the rate of transcription by increasing the activity of RNA polymerase at specific promoters.

transducer. The component of a signal-transduction pathway that couples receptor-ligand binding with generation of a second messenger catalyzed by an effector enzyme.

transfer ribonucleic acid. *See* tRNA.

transferase. An enzyme that catalyzes a group-transfer reaction. Transferases often require a coenzyme.

transition state. An unstable, high-energy arrangement of atoms in which chemical bonds are being formed or broken. Transition states have structures between those of the substrates and the products of a reaction.

transition-state analog. A compound that resembles a transition state. Transition-state analogs characteristically bind extremely tightly to the active sites of appropriate enzymes and thus act as potent inhibitors.

transition-state stabilization. The increased binding of transition states to enzymes relative to the binding of substrates or products. Transition-state stabilization lowers the activation energy and thus contributes to catalysis.

translation. The synthesis of a polypeptide whose sequence reflects the nucleotide sequence of an mRNA molecule. Amino acids are donated by activated tRNA molecules, and peptide bond synthesis is catalyzed by the translation complex, which includes the ribosome and other factors.

translation complex. The complex of a ribosome and protein factors that carries out the translation of mRNA in vivo.

translation initiation complex. The complex of ribosomal subunits, an mRNA template, an initiator tRNA molecule, and initiation factors that assembles at the start of protein synthesis.

translation initiation factor. A protein involved in the formation of the initiation complex at the start of protein synthesis.

translocation. 1. The movement of the ribosome by one codon along an mRNA molecule. **2.** The movement of a polypeptide through a membrane.

transposon. A mobile genetic element that jumps between chromosomes or parts of a chromosome by taking advantage of recombination mechanisms. Also known as a transposable element.

transverse diffusion. The passage of lipid or protein molecules from one leaflet of a lipid bilayer to the other leaflet. Unlike lateral diffusion within one leaflet of a bilayer, transverse diffusion is extremely slow.

triacylglycerol. A lipid containing three fatty acyl residues esterified to glycerol. Fats and oils are mixtures of triacylglycerols. Formerly known as a triglyceride.

tricarboxylic acid cycle. *See* citric acid cycle.

triglyceride. *See* triacylglycerol.

triose. A three-carbon sugar.

tRNA. A class of RNA molecules that carry activated amino acids to the site of protein synthesis for incorporation into growing peptide chains. tRNA molecules contain an anticodon that recognizes a complementary codon in mRNA.

turn (in proteins). A protein loop of 4-5 residues that causes a change in the direction of a polypeptide chain in a folded protein.

turnover. The dynamic metabolic steady state in which molecules are degraded and replaced by newly synthesized molecules.

turnover number. *See* catalytic constant.

twist. The angle of rotation between adjacent residues within a helical macromolecule.

type I reaction center. The special pair of chlorophyll molecules and associated electron transfer chain found in photosystem I.

type II reaction center. The reaction center found in photosystem II.

uncompetitive inhibition. Inhibition of an enzyme-catalyzed reaction by a reversible inhibitor that binds only to the enzyme-substrate complex, not to the free enzyme.

uncouplers. *See* uncoupling agent.

uncoupling agent. A compound that disrupts the usual tight coupling between electron transport and phosphorylation of ADP.

uniport. The transport of a single type of solute across a membrane by a transport protein.

unsaturated fatty acid. A fatty acid with at least one carbon-carbon double bond. An unsaturated fatty acid with only one carbon-carbon double bond is called a monounsaturated fatty acid. A fatty acid with two or more carbon-carbon double bonds is called a polyunsaturated fatty acid. In general, the double bonds of unsaturated fatty acids are of the *cis* configuration and are separated from each other by methylene ($-CH_2-$) groups.

urea cycle. A metabolic cycle consisting of four enzyme-catalyzed reactions that converts nitrogen from ammonia and aspartate to urea. Four ATP equivalents are consumed during formation of one molecule of urea.

***v*.** *See* velocity.

v_0. *See* initial velocity.

vacuole. A fluid-filled organelle in plant cells that is a storage site for water, ions, or nutrients.

van der Waals force. A weak intermolecular force produced between neutral atoms by transient electrostatic interactions. Van der Waals attraction is strongest when atoms are separated by the sum of their van der Waals radii; strong van der Waals repulsion precludes closer approach.

van der Waals radius. The effective size of an atom. The distance between the nuclei of two nonbonded atoms at the point of maximal attraction is the sum of their van der Waals radii.

variable arm. The arm of a tRNA molecule that is located between the anticodon arm and the TψC arm. The variable arm can range in length from about 3 to 21 nucleotides.

velocity (V). The rate of a chemical reaction, expressed as amount of product formed per unit time.

very low density lipoprotein (VLDL). A type of plasma lipoprotein that transports endogenous triacylglycerols, cholesterol, and cholesteryl esters from the liver to the tissues.

vitamin. An organic micronutrient that cannot be synthesized by an animal and must be obtained in the diet. Many coenzymes are derived from vitamins.

VLDL. *See* very low density lipoprotein.

V_{max}. *See* maximum velocity.

wax. A nonpolar ester that consists of a long chain monohydroxylic alcohol and a long chain fatty acid.

wobble position. The 5′ position of an anticodon, where non-Watson-Crick base pairing with a nucleotide in mRNA is permitted. The wobble position makes it possible for a tRNA molecule to recognize more than one codon.

X-ray crystallography. A technique used to determine secondary, tertiary, and quaternary structures of biological macromolecules. In X-ray crystallography, a crystal of the macromolecule is bombarded with X rays, which are diffracted and then detected electronically or on a film. The atomic structure is deduced by mathematical analysis of the diffraction pattern.

Z-DNA. A conformation of oligonucleotide sequences containing alternating deoxycytidylate and deoxyguanylate residues. Z-DNA is a left-handed double helix containing approximately 12 base pairs per turn.

zero-order reaction. A reaction whose rate is independent of reactant concentration.

Z-scheme. A zigzag scheme that illustrates the reduction potentials associated with electron flow through photosynthetic electron carriers.

zwitterion. A molecule containing negatively and positively charged groups.

Photo and Illustration Credits

Chapter 1 Page 2 top, Science Photo Library/Photo Researchers, Inc.; 2 middle, Photos 12/Alamy; 2 bottom, Science Photo Library/Photo Researchers, Inc.; 3 top, Corbis; 3 bottom, Shutterstock; 11, Shutterstock; 12, Manuscripts & Archives—Yale University Library; 15 top, SSPL/The Image Works; 15 bottom, Richard Bizley/Photo Researchers, Inc.; 18 top, Lee D. Simon/Photo Researchers, Inc.; 18 bottom, National Library of Medicine Profiles in Science; 20, Matthew Daniels, Wellcome Images; 22, Dr. Torsten Wittmann/Photo Researchers, Inc.; and 23, David S. Goodsell, the RCSB Protein Data Bank. Coordinates from PDB entry 1atn.

Chapter 2 Page 28 top, NASA; 28 bottom, Michael Charters; 31, iStockphoto; 32, NOAA; 33, Valley Vet Supply; 37, Travel Ink/Getty Images; 41, Elemental-Imaging/iStockphoto; 44 top, Edgar Fahs Smith Memorial Collection; 44 bottom, Fotolia; and 48, Library of Congress.

Chapter 3 Page 56, Thomas Deerinck, NCMIR/Photo Researchers, Inc.; 57, Argonne National Laboratory; 58, Pascal Goetgheluck/Photo Researchers, Inc.; 60, iStockphoto; 69, iStockphoto; 70, MARKA/Alamy; 71, Bio-Rad Laboratories, Inc.; 73 top, REUTERS/William Philpott WP/HB; 73 bottom, AFP Photo/Newscom; and 78, Bettmann/CORBIS.

Chapter 4 Page 85, Shutterstock; 86, Swiss Institute of Bioinformatics; 88, Lisa A. Shoemaker; 89 top, Bror Strandberg; 89 bottom, Hulton Archive/Getty Images; 93, Custom Life Science Images/Alamy; 94, Bettmann/Corbis; 95, Julian Voss-Andreae; 108, From Kühner et al., "Proteome Organization in a Genome-Reduced Bacterium" Science 27 Nov 2009 Vol. 326 no. 5957 pp. 1235–1240. American Association for the Advancement of Science.; 109, Howard Ochman; 111, From Butland et al., "Interaction network containing conserved and essential protein complexes in Escherichia coli," Nature 433 (2005), 531–537; 113, National Library of Medicine; 117, Laurence A. Moran; 119, Easawara Subramanian, http://www.nature.com/nsmb/journal/v8/n6/full/nsb0601_489.html; 121, Danielle Anthony; 122, SSPL/The Image Works; 123, Janice Carr/Centers for Disease Control; 126, Ed Uthman, licensed via Creative Commons http:// creativecommons.org/licenses/by/2.0/; and 127, Julian Voss-Andreae.

Chapter 5 Page 135, Dorling Kindersley; 136, Jonathan Elegheert; 137, Michael P. Walsh/IUBMB; 138, Leonardo DaVinci; 142 top, Rockefeller Archives Center; 142 bottom left, University of Pittsburgh, Archives Service Center; 142 bottom right, Laurence A. Moran; and 149, AP Photo/Paul Sakuma.

Chapter 6 Page 167, Ronsdale Press, photo copyright Dina Goldstein; 174, Bettmann/CORBIS; 183, Paramount/Photofest; and 186, Shutterstock.

Chapter 7 Page 198, Shutterstock; 200, Library of Congress; 204, Heath Folp/Industry & Investment NSW; 209, History Press; 212, Christian Heintzen, University of Manchester; 214, iStockphoto; 215, John Olive; 216, Stephanie Schuller/Photo Researchers, Inc.; 219 left, Meg and Raul via Flickr/CC-BY-2.0 http://creativecommons.org/licenses/by/2.0/deed.en 219 right, and 220, Shutterstock; and 223, both, ©® The Nobel Foundation.

Chapter 8 Pages 227, 239, 240, Shutterstock; 244 top, Image Source/Alamy; 244 bottom, Jack Griffith; 245, Jakob Jeske/Fotolia; 246, Jens Stougaard; 247 top, Eric Erbe, Christopher Pooley, Beltsville Agricultural Research/USDA; 247 bottom, Robert Hubert, Microbiology Program, Iowa State University; and 252, Christine Ortlepp.

Chapter 9 Page 258, imagebroker/Alamy; 262 top, Steve Gschmeissner/Photo Researchers, Inc.; 262 bottom, Shutterstock; 268 bottom, Shutterstock; 270, John Ross; 273 top, Professors Pietro M. Motta & Tomonori Naguro/Photo Researchers, Inc.; 273 bottom, Biophoto Associates/Photo Researchers, Inc.; 277, Lisa A. Shoemaker; 278 bottom, Julie Marie/Fotolia; 284 top, M.M. Perry; and 284 bottom, Shutterstock.

Chapter 10 Page 294, Quade Paul, Echo Medical Media; 296, Charles Boone, From Costanzo et al. "The Genetic Landscape of a Cell" *Science 327*; (2010):425–432; 297, Roche Applied Science; 303, Shutterstock; 305 top, University of Edinburgh/Wellcome Images; 305 bottom, Biophoto Associates/Photo Researchers, Inc.; and 312, National Library of Medicine.

Chapter 11 Page 325, Barton W. Spear—Pearson Education; 331 left, Super-Stock, Inc;. 331 right, Bettmann/CORBIS; 336, Warner Bros./Photofest; 341, ChinaFotoPress/Zuma/ICON/Newscom; and 349, dreambigphotos/Fotolia.

Chapter 12 Page 359, CBS/Landov; 369, United States Postal Service; 370 top, A. Jones/Photo Researchers, Inc.; 370 bottom, Laura Van Niftrik; and 375, Tim Crosby/Getty Images.

Chapter 13 Page 386, Science Photo Library/Photo Researchers, Inc.; 387, From Zhou, Z.H. et al. (2001) Proc. Natl. Acad. Sci. USA 98, pp. 14802–14807; 390 top, From Zhou, Z.H. et al. (2001) Proc. Natl. Acad. Sci. USA 98, pp. 14802–14807; 390 bottom, NASA; and 396, 401, Shutterstock.

Chapter 14 Page 417 top and left, Shutterstock; 417 bottom, Dirk Freder/iStockphoto; 419 top, Lisa A. Shoemaker; 419 middle and bottom, Shutterstock; 420 top Roberto Danovaro; 420 left, Milton Saier; 426, Michael Radermacher; 433, Alexander Tzagoloff; and 438, NASA/Sandra Joseph and Kevin O'Connell.

Chapter 15 Page 443, Mary Ginsburg; 444, Arizona State University—Plant Bio Department; 447 top, Makoto Kusaba; 447 bottom, Shutterstock; 448 top, CHINE NOUVELLE/SIPA/Newscom; 448 bottom, Robert Lucking; 452, Niels Ulrik Frigaard; 457, Michelle Liberton, Howard Berg, and Himadri Pakrasi, of the Donald Danforth Plant Science Center and of Washington University, St. Louis; 458 top, Andrew Syred/Photo Researchers, Inc.; 458 bottom, NSF Polar Programs/NOAA; 459, Lisa A. Shoemaker; 462, Lawrence Berkeley National Laboratory; 468, Shutterstock; 469 top, From Bhattacharyya et al, "The wrinkled-seed ..." *Cell*, Vol 60, No 1, 1990, pp 115–122; 469 middle, Peter Arnold/Photolibrary; 469 bottom, Fotolia; 470, From David F. Savage et al., "Spatially Ordered Dynamics of the Bacterial Carbon Fixation Machinery," 2011. American Association for the Advancement of Science; 471 top, AP Photo/Charlie Neibergall; and 471 bottom, Shutterstock.

Chapter 16 Page 475, Kennan Ward/Corbis; 486, Shutterstock; 490 top, Bettmann/CORBIS; 490 bottom, Hulton Archive/Getty Images; 493, Environmental Justice Foundation, Ltd.; 495 top, David Leys, Toodgood et al., 2004; 495 bottom, Eric Clark/Molecular Expressions; 501 top, Shutterstock; 501 bottom, Steve Gschmeissner/SPL/Alamy; 504, Donald Nicholson/IUBMB; 506, Shutterstock; and 507, Robin Fraser.

Chapter 17 Page 515 top, NASA Visible Earth; 515 bottom, NOAA; 516, Inga Spence/Photo Researchers, Inc.; 531, Shutterstock; 532, iStockphoto.com; 534, National Library of Medicine; and 540, U.S Air Force photo/Staff Sgt Eric T. Sheler.

Chapter 18 Page 552, G. Robert Greenberg; 554, National Library of Medicine; 561, Peter Reichard; 564, Shutterstock; and 568, Fotolia.

Chapter 19 Page 574, National Cancer Institute; 581, SSPL/The Image Works; 587, Andrew Paterson/Alamy; 589, Lisa A. Shoemaker; 591 both, Ulrich K. Laemmli; 597 top left, 597 top right, Lisa A. Shoemaker; 597 middle, Stanford University School of Medicine; and 597 bottom, Steve Northup/Time&Life Images/Getty Images.

Chapter 20 Page 603 top, John Cairns; 603 bottom left, David S. Hogness; 603 bottom right, Regional Oral History Office, The Bancroft Library, University of California, Berkeley; 613 both, Timothy Lohman; 615, From *Structure,* 6, Dec. 2008 Copyright Elsevier. Original artwork by Glass Egg Design, Jessica Eichman, www.glasseggdesign.com; 618, Lisa A. Shoemaker; 619, David Bentley; 627 top, Laguna Design/Photo Researchers, Inc.; 627 bottom, Paul Sabatier/Art Life Images/Superstock; 628 top, James Kezer/Stanley Sessions; 628 bottom, Dr. L. Caro/Photo Researchers, Inc; 630, Institute of Molecular-biology and Biophysics, From Yamada et al., *Molecular Cell* Vol 10 p 671 (2002). Figure 4b (right), with permission from Elsevier.; and 630, Vanderbilt University, Genes and Development. From Wang et al. *BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures.* Vol. 14, No. 8, pp. 927–939, April 15, 2000 Fig 3M.

Chapter 21 Page 634, Marc Gantier/Getty Images; 636, From Murakami.et al., Science 296: 1285–1290 (2002) Fig5A (left) American Association for the Advancement of Science; 638, Oscar L. Miller, Jr.; and 651, Lisa A. Shoemaker.

Chapter 22 Page 666, National Security Agency; 667, US Navy Office of Information; 675, David Goodsell; 677, Stanford University School of Medicine; 681, Oscar L. Miller, Jr.; and 692, H. H. Mollenhauer/USDA.

This page intentionally left blank

Index

In this index, the page numbers listed indicate tables (with a T added to that page number) and figures (with an F added to that page number).

A
A-DNA, 585–586F
ABO blood group, 250–251F
absorption spectrum of DNA, 584–585F
acceptor stem, 668F
accessory pigments, 447–448F
acetaldehyde, lyases catalyzation, 137
acetaminophen, structure of, 486F
acetate, gluconeogenesis precursor, 362–363
acetic acid (CH_3COOH), 45
  buffer range of, 50F
  dissociation of, 45
  pH and, 45, 47, 50F
  titration of, 47F
acetyl CoA, 315–316, 387–394
  cholesterol and, 488
  citric acid cycle reactions, 385, 387–394
  isopentenyl diphosphate conversion from, 488
  nucleotidyl group transfer, 315
  oxidation of, 385, 391–394F
  pyruvate, conversion from, 385, 387–391
  thioester hydrolysis, 316
acetylcholinesterase, 134F
acid–base catalysis, 168–169
acid solutions, 42–49F
  base solutions combined with, 47–48
  base solutions dissociated from, 44–45
  dissociation constant, K_a, 44–48T
  Henderson–Hasselbach equation for, 46–47
  ionization and, 42
  pH scale for, 43F, 49
  parameter value, pK_a, 45–48T
  titration, curves for, 47–48F
  weak, 44–49
aconitase, citrus cycle reactions, 396–397F
actin filaments, 23F
activation energy, $G^{\ddagger}$, 14F, 165F
activator ions, 196
active membrane transport, 280–283F
acute lymphoblastic leukemia treatment, 521
acyl, general formula of, 5F
acyl carrier protein (ACP), 111F, 204–206F
acyl CoA transport into mitochondria, 497–498
adenine (A), 8–9F, 310–311T, 551F
adenosine deaminase, 181–182F
adenosine 5-monophosphate (AMP), 550–551F
adenosine triphosphate (ATP), 8–9F, 198–199F, 308–315, 417–442
  active membrane transport, 282–283F, 435–436
  β-oxidation, generation from, 498–499
  citric acid cycle reactions, 405–406F
  coenzyme metabolic property, 198–199F
  cyclic adenosine monophosphate (cAMP), 287–288F
  electron transport and, 417–442
  eukaryotic mitochondria and, 21
  Gibbs free energy change, ΔG, 308–312
  hexokinase reactions, 326–327, 328F, 330F
  high energy bond, ~, 311
  hydrolysis, 308–312
    electrostatic repulsion, 309
    metabolically irreversible changes, 308–312
    resonance stabilization, 310
    solvation effects, 309–310
  metabolic changes, 198–199F, 304, 308–315
  nucleotide metabolic reactions, 551F
  nucleotidyl group transfer, 315F
  phosphofruktokinase-1 (PFK-1) regulation by, 345–346F
  phosphoryl group transfer, 312–315
  photosynthesis photosystems and, 459–460F
  production of, 314–315F
  reduced coenzyme production of, 405–406F
  structure of, 8–9F
  synthase, 433–435F, 456, 459–460F
  synthesis of, 417–442
    ATP synthase catalysis, 433–435F
    chemiosmotic theory, 420–423
    mitochondria, 418–420F
    mitochondrial membrane transport, 435–436
    NADH shuttle mechanisms in eukaryotes, 436–439F
    P/O (phosphorylated/oxygen) ratio, 436
    proton leaks and heat production, 435
    protonmotive force, 421–420F
    superoxide anions, 440–441
adenylyl cyclase signaling pathway, 287–288F
adenylyl kinase (pig), 105F
affinity chromatography, 70
aggrecan, 245–246
aggregation from protein folding, 119
Agre, Peter, 280
Agrobacterium sp., 528
alanine (A, Ala), 56, 59F, 64T
  catabolism of, 535
  gluconeogenic precursor, 361
  glucose-alanine cycle, 361F
  ionization of, 64–65F
  isomerases catalyzation, 137–138
  nomenclature, 56, 64T
  pyruvate, conversion from, 361F
  structure and properties of, 56, 59F
  synthesis of, 521–523F
  titration of, 64–65F
  transferases catalyzation, 136–137
alcohol groups with side chains, 60–61
alcohols, 5F
  cyclization of monosaccharides and reactions of, 230–231F
  general formula of, 5F
  solubility in water, 35T
aldehyde, general formula of, 5F
aldohexoses, 229F
aldolase cleavage, 330–332F
aldopentoses, 229F
aldoses, 228–234F
  cyclization of, 230–234F
  epimers, 230
  Fischer projections of, 228–230F
  structure of, 228–230F
aldotetroses, 229F
aliphatic R groups, 59
alkaline hydrolysis, 591–592F
alkaptonuria, 544
allose, 229F
allosteric enzymes, 153–158F
  concerted (symmetry) model for, 156–157F
  phosphofructokinase, 154–155F
  properties of, 155–156F
  regulation of enzyme activity using, 153–158
  sequential model for, 157–158F
allosteric protein interactions, 127–129F
allosteric regulation of eukaryotic ribonucleotide reductase, 561T
allysine residues, 121F
α-carbon atom, 56
α-globin subunits, 122–123F
α helix proteins, 94–97F, 98–99
  amphipathic, 95–97A
  β strand and sheet connections, 98–99F
  collagen type III triple helix, 119F
  left-handed, 119–120F
  leucine zipper, 96–97A
  membranes, 270–271F
  protein conformation of, 94–97F
  right-handed, 94–95F
  rotation of, 95
  side chains in, 95
  3_{10} helix compared to, 96–97F
α-ketoglutarate, transferases catalyzation, 56
α-ketoglutarate dehydrogenase complex, citrus cycle reactions, 398–399F
α subunits, RNA transcription, 641–642T
α-tocopherol (vitamin E), 218F
α/β barrel, domain fold, 106F
$\alpha_2\beta_2$ tetramer (insulin), 290–291F
altrose, 229F
amide linkages, 4–5F
amino acid metabolism, 514–549
  ammonia assimilation, 518–519
    glutamate and glutamine incorporation, 518F
    transanimation reactions, 518–519F
  catabolism, 534–542
    alanine, asparagine, aspartate, glutamate, and glutamine, 535
    argenine, histidine, and proline, 535–536F
    branched chain amino acids and, 537–539F
    cysteine, 540–541F
    glycine and serine, 536–537F
    lysine, 542F
    methionine conversion and, 539–540F
    threonine, 537–538F
    tyrosine, 541–542F
  diseases of, 544
  essential amino acids, 529T
  functions of, 514–515
  nitrogen cycle, 515–517F
  nitrogen fixation, 515
  nitrogenases, 516–517
  nonessential amino acids, 514, 529T
  precursors, 529–532
    glutamate, glutamine, and aspartate, 529
    lignin from phenylalanine, 531–532F
    melanin from tyrosine, 531, 533F
    nitric oxide from arginine, 530–531F
    serine and glycine, 529–530F
  protein turnover, 531–533
  renal glutamine metabolism, 547–548
  synthesis of amino acids, 520–529
    alanine, valine, leucine, and isoleucine, 521–523F
    aspartate and asparagine, 520–521F
    citric acid cycle, 520F

amino acid metabolism (*Continued*)
glutamate, glutamine, arginine, and proline, 523F
histidine, 527F
lysine, methionmine, and threonine, 520–522F
phenylalanine, tyrosine, and tryptophan, 524–527F
serine, glycine, and cysteine, 523–525F
urea cycle, conversion of ammonia to urea, 542–547
amino acids, 6F, 55–84
α-carbon atom, 56
active sites of enzymes, 168T
catabolism of, 519, 534–542
catalytic functions of residues, 166–168T
chromatographic procedure for, 73–74F
common types of, 58–62
alcohol groups with side chains, 60–61
aliphatic R groups, 59
aromatic R groups, 59–60
derivatives, 62–63
hydrophobicity of side chains, 62
negatively charged R groups, 62
positively charged R groups, 61–62
sulfur-containing R groups, 60
defined, 56
evolution and ancestors from, 57–58, 79–81F
free-energy change of transfer for, 63T
glucose precursors, 360–361
hydrolysis for analysis of, 73–74F
hydropathy scale, 62T
ionization of, 63–67
molecular weight of, 74–75T
nomenclature, 56–58, 61F, 64T
peptide bonds, 67–68
pK_a values, 168T
protein composition with, 67–68, 73–74T
protein purification and analysis, 68–73
racemization, 58
residues, 67–68F, 74–75F
RS system configuration, 61F
sequencing, 68, 74–81F
side chains, 56, 59–62
site-directed mutagenesis, 167
structure of, 6F, 56–62F
abbreviations for, 58–59F
ball-and-stick model of, 56–57F
mirror-image pairs, 57F
numbering conventions, 56F
titration of, 64–65F
amino sugars, 235–236, 237F
aminoacyl-tRNA, 670–673
binding sites, 671–672F, 675F, 677F
docked at A site, 675, 677F, 680–682F
elongation factors and docking of, 680–682F
ribosome binding sites, 675, 677F
synthetases, 670–673F
proofreading for errors in, 673
protein synthesis and, 670–673F
reaction of, 670–672F
specificity of, 671–673F
substrate-binding sites, 677F
aminoimidazole carboxamide ribonucleotide (AICAR), 553F
aminoimidazole ribonucleotide (AIR), 553F
aminoimidazole succinylocarboxamide ribonucleotide (SAICAR), 553F
ammonia (NH_3), 45, 518–519
assimilation, 518–519
conversion to urea, 542–547
dissociation for formation of, 45
enzyme transfer from glutamate, 558
glutamate and glutamine incorporation, 518F
transanimation reactions, 518–519F
urea cycle, 542–547F
ammonium ions, general formula of, 5F
amphibolic pathways, 407–409
amphibolic reactions, 295
amphipathic helix, 95–97A
amphipathic molecules, 36
amplification, 285
DNA, 615–616
signal pathways, 285
amylase, 242F
amylopectin, 241–242F
amyloplasts, 469
amylose, 241F
Anabaena spherica, 305F
anabolic (biosynthetic) reactions, 294–295F, 302–303F
anaerobic conversion, 339–340F
Anfinsen, Christian B., 112–113
angstrom (Å), units of, 26
anionic forms of fatty acids, 258T
anomeric carbon, 231
anomers, 231
antenna chlorophylls, 446–447F
anti conformation of nucleotides, 577–578F
antibiotic inhibition of protein synthesis, 686
antibody binding to specific antigens, 129–130F
anticodon arm, 668–669F
anticodons, 668–671T
base pairing, 669–670T
defined, 668
wobble position of, 670–671F
antigens, antibody binding to, 129–130F
antiparallel β sheets, 97–98F
antiparallel DNA strands, 581–583
antiport, membrane transport, 281F
apoptosis, 534
aquaporin, 280F
Arabidopsis thalianna, 93
arabinose, 229F
L-arabinose-binding protein, 105F
arginine (R, Arg), 61–62F
catabolism of, 535–536F
nitric oxide synthesis from, 530–531F
nomenclature of, 64T
structure of, 61–62F
synthesis of, 523F
urea cycle and, 543F, 545–546F
arginine kinase, 190–192F
aromatic R groups, 59–60
arsenate (arsenic) poisoning, 336
arsenite (arsenic) poisoning, 336
ascorbic acid (vitamin C), 209–211
asparagine (N, Asn), 62F
acute lymphoblastic leukemia treatment, 521
catabolism of, 535
nomenclature, 64T
structure of, 62F
synthesis of, 520–521F
aspartame, 68F, 240
aspartate (D, Asp), 62F
catabolism of, 535
gluconeogenic precursor, 361
malate–aspartate shuttle, 348F
metabolic precursor use, 529
nomenclature, 64T
structure of, 62F
synthesis of, 520–521F
urea cycle and, 543F, 545–546F
aspirin, structure of, 486F
association constant, K_a, 109–110F
atmospheric pollution, photosynthesis and, 457
ATP, *see* adenosine triphosphate (ATP)
ATP synthase, 433–435F
binding change mechanism, 434–435F
chloroplasts, 459–460F
cytochrome complexes, 456
electron transfer from, 456
electron transport, complex V, 433–435F
photosynthesis and, 456, 459–460F
rotation of molecules, 434–435
structure of, 433F
attenuation, 688–689F
audioradiograph of replicating chromosome, 603F
autophosphorylation, 290
autotrophs, 302–303
Avery, Oswald, 3, 573
Azotobacter vinelandii nitrogenase, 516–517F

B
B-DNA, 582–584F
bacteria, 246–248. *See also Escherichia coli* (*E. coli*)
citric acid cycle and, 411–414
Entner-Doudoroff (ED) pathway, 351–352F
forked pathway, 412–413F
gloxylate pathway, 411–412F
Gram stain for, 247F
intestinal, 216F
metabolism and adaptation of, 295–296
penicillin, 247–248F
peptidoglycans, 246–248F
polysaccharide capsules, 247
Staphylococcus aureus (*S. Aureus*), 76, 247–248F
bacterial DNA, 3, 590
bacterial enzymes, 364F
bacterial flagellum, 109F
bacterial photosystems, 448–458
coupled, 453–455T
cytochrome *bf* complex, 453–455F
electron transfer in, 449–453
Gibbs free energy change, ΔG, 455–457
green filamentous bacteria, 448, 452F
internal membranes, 457
photosystem I (PSI), 448, 450–453F
photosystem II (PSII), 448–450F
purple bacteria, 448–450F
reaction equations, 450T, 452T, 455T
reduction potentials, 455–457F
bacterial reaction center (BRC), *see* photosystems
bacterial transducers, 285–286
bacteriophage MS2 capsid protein, 107F
bacteriorhodopsin, 270–271F, 461
ball-and-stick models, 56–57F
amino acids, 56–57F
DNA, 582–584F
monosaccharide (chiral) compounds, 228F, 235F
Barnum, P. T., 200
Bascillus stearothermophylus, 402
Bascillus subtilis, 186
base composition of DNA, 579T
base pairing, 604–606, 669–671
DNA, 604–606
protein synthesis, 668–671F
Watson–Crick, 668–670F
wobble positions of anticodon and codon, 670T-671F
base solutions, 42–43F, 47–48F
acid titration using, 47–48F
dissociated from acid solutions, 44–45
Henderson–Hasselbach equation for, 47–48
ionization and, 42
pH scale for, 43F
Beadle, George, 212, 634
β barrel, domain fold, 106F
β barrel protein membranes, 271–272F
β-carotene, 217F, 447F
β-globin subunits, 122–123F
β helix, domain fold, 106F
β–meander motif (structure), 100–101F
β-oxidation, 494–501
acyl CoA transport into mitochondria, 497–498
ATP generation from, 498–499
fatty acids, 494–501
lipid metabolism and, 494–501
odd-chain fatty acids, 499–500
trifunctional enzymes and, 498
unsaturated fatty acids, 500–501

β-sandwich motif (structure), 100–101F
β strands and sheets, 97–99F
  α helix connections, 98–99F
  antiparallel sheets, 97–98F
  β turns, 99F
  hydrophobic interactions, 98
  loops, 98
  parallel sheets, 97–98F
  pleated sheet, 97–98
  protein conformation of, 97–99F
  residues and, 99F
  reverse turns, 99
  turns, 99F
βαβ unit motif (structure), 100F
bicarbonate production by renal glutamine metabolism, 547–548
bidirectional DNA replication, 602–603F
bile salts, 505F
binding. *See also* oxygen binding; substrates
  aminoacyl-tRNA sites, 671–672F, 675F, 677F
  cap binding protein (CBP), 679
  change mechanism, ATP synthase, 434–435F
  DNA fragments, 609–611F
  hormones, 286–288
  protein synthesis, 671–672F, 675F, 677–679F
biochemistry, 1–27
  biopolymers, 4–10
  cells, 17–26
    E. coli, 17F, 23–24, 26F
    eukaryotic, 18–23F
    living, 23–26
    prokaryotic, 17–18F
  chemical elements of life, 3–4
  defined,
  energy, life and, 10–15
  evolution and, 15–17
  macromolecules, 4–10
    lipids, 9
    membranes, 9–10
    nucleic acids, 7–9F
    polysaccharides, 6–7F
    proteins, 6
  multidisciplinary nature of, 26
  special terminology of, 26–27
  20th century science and, 2–3
  units for, 26–27T
bioenergetics, 11. *See also* ATP; metabolism; thermodynamics
biological functions, 55–56, 119–129
  amino acid metabolism diseases, 544
  antibody binding to specific antigens, 129–130
  blood plasma, 33F, 35F, 51–52F
  cancer DCA inhibitors, 408F
  cartilage structure, 245–246F
  coronary heart disease and lipoprotein lipase, 507
  diabetes mellitus (DM), 381, 511
  dietary requirements and fatty acids, 261
  genetic defects, 265–266
  gout, 569
  hyperactivity, 359
  intestinal bacteria, 216F
  lactate buildup, 341
  lactose intolerance, 350
  Lesch–Nyhan syndrome, 569
  gout, 569
  lysosomal storage diseases, 492F
  liver metabolic functions, 344–345F, 379–380F
  mucin secretions, 252F
  oxygen binding to myoglobin and hemoglobin, 123–129
  proteins and, 55–56, 119–129
  scurvy, ascorbic acid and, 209–210
  sweetness receptors, 240
  vitamin deficiency, 198 T, 209–210, 214, 215
biological membranes, 9, 269–275. *See also* membranes
biopolymers, *see* polymers
  biosynthetic (anabolic) pathways, 302–303
biotin, 211–212F
2,3-bisphospho-D-glycerate (2,3BPG), 127–128F
1,3 *bis*phosphoglycerate, 334F
2,3 *bis*phosphoglycerate, 335–337F
bisubstrate enzyme reactions, 147–148F
blood, 33F, 35F, 250–251F
  ABO group, 250–251F
  2,3 *bis*phosphoglycerate in, 335F
  buffer capacity, 51–52F
  glycolysis reactions, 335
  plasma, 33F, 35F, 51–52F
  properties of, 33F, 35F
boat conformations, 235F
Bohr effect, 128F
Boyer, Herbert, 597
Boyer, Paul D., 223, 434
branched chain amino acids, 537–539F
breast cancer and DNA repair, 630
Briggs, George E., 141
Buchanan, John (Jack) M., 551, 554
Buchner, Eduard, 2, 331
buffered solutions, 50–52F
  acetic acid, 50F
  blood plasma, 51–52F
  capacity and pK_a, 50–52FT
  carbonic acid, 51F
  pH and, 50–52F
  preparation of, 50

C

C-terminus (carboxyl terminus), 68, 76F
C_3 pathway, *see* Calvin cycle
C_4 pathway, 469–471F
Caenorhabditis elegans, 296
Cahill, George, 380
calcium (Ca), 3
calories (cal), units of, 26
calorimeter, 13F
Calvin, Melvin, 462
Calvin cycle, 443, 461–467F
  carbon dioxide (CO_2) fixation, 461–467, 469–472
  NADPH reduction, 466–467
  ribulose 1,5-*bis*phosphate, 465–466F
  rubisco (rubilose 1,5-*bis*phosphate carboxylase-oxygenase), 462, 464–466F
  stages of, 462F
    oxygenation, 465–466F
    reduction, 466–467
    regeneration, 466–467F
cancer drug inhibition, synthesis for, 564
cap binding protein (CBP), 679
cap formation, mNRA, 658–659F
capsaicin, 284F
capsule, polysaccharide, 247
carbamate adducts, 129F
carbamoyl phosphate, urea cycle and, 543F, 545–546F
carbamoyl phosphate synthetase, 558F
carbocation, 164
carbohydrates, 227–255
  defined, 227
  disaccharides, 236–239
    glycosidic bonds in, 236–238F
    structures of, 237–239F
    sugars, 238–239
    glucosides and, 236–239, 241F
    nucleosides and, 239, 241F
  glucosides, 236–239, 241F
  glycoconjugates, 244–252
    glycoproteins, 248–252F
    peptidoglycans, 246–248F
    proteoglycans, 244–246F
  monosaccharides, 227–236
    aldoses, 228–234F
    ball-and-stick models of, 228F
    chiral compounds, 228–230F
    conformations of, 234–235F
    cyclization of, 230–234
    derivatives of, 235–236F
    epimers, 230
    Fischer projections of, 228–232F
    Haworth projections of, 232–235F
    ketoses, 228–234F
    trioses, 226
  oligosaccharides, 227, 248–252F
  polysaccharides, 227, 240–244
    cellulose, 243F
    chitin, 244F
    glycogen, 240–243F
    heteroglycans, 240
    homoglycans, 240
    starch, 240–242F
    structure of, 240–241T
carbolic acid, general formula of, 5F
carbon (C), 3
  glycolysis reactions, 333–334F
carbon dioxide (CO_2), lyases catalyzation, 137
carbon dioxide (CO_2) fixation, 461–467
  bacteria compartmentalization, 469
  C_4 pathway, 469–471F
  Calvin cycle, 443, 461–467F
  carboxysomes, 469–470F
  crassulacean acid metabolism (CAM), 471–472F
  NADPH reduction, 466–467
  ribulose 1,5-*bis*phosphate, 465–466F
  rubisco (rubilose 1,5-*bis*phosphate carboxylase-oxygenase), 462, 464–466F
carbonic acid, buffer capacity of, 51F
carbonic anhydrase, 197F
carbonyl, general formula of, 5F
carboxyaminoimidazole ribonucleotide (CAIR), 553F
carboxylate, general formula of, 5F
carboxysomes, 469–470F
carotenoids, 447–448F
cartilage structure, 245–246F
cascade amplification of signal pathways, 285
catabolic reactions, 295F, 303–304F. *See also* glycolysis
  glucose, 325–354
  metabolic pathways, 303–304F
  NADH, 304
catabolism, 534–542
  alanine, asparagine, aspartate, glutamate, and glutamine, 535
  amino acid metabolism and, 534–542
  argenine, histidine, and proline, 535–536F
  branched chain amino acids and, 537–539F
  cysteine, 540–541F
  glycine and serine, 536–537F
  lysine, 542F
  methionine conversion and, 539–540F
  purine, 565–568
  pyrimidine, 568–570
  threonine, 537–538F
  tyrosine, 541–542F
catalysis, 166–171, 175–182
  acid–base, 168–169
  amino acid residues and, 166–168T
  catalytic residue frequency distribution, 168T
  chemical modes of, 166–171
  covalent, 169–170F
  diffusion-controlled reactions, 171–175
  enzymatic modes, 175–182
    induced fit, 179–180
    proximity effect, 176–178F

catalysis (*Continued*)
transition–state stabilization, 176, 180–182F
weak binding and, 176, 179–179F
enzyme mechanism of, 166–171, 175–182
ionizable amino acid residue functions, 166–168T
pH effects on enzymatic rates, 170–172F
pK_a values of ionizable amino acids, 168T
RNA polymerase, 637–638F
serine proteases and modes of, 185–188
substrate binding and, 171–172T, 175–182F
catalysts, 2, 113–114, 134, 136–138
defined, 134
denaturation reduction from, 113–114
hydrolase enzymes, 137
isomerases enzymes, 137–138
ligases enzymes, 138
lyases enzymes, 137
oxidoreductase enzymes, 136
protein structures, 113–114
regulation of enzyme activity, 153–158
transferases enzymes, 136–137
catalytic activity, 89
catalytic constant, k_{cat}, 143–145
catalytic proficiency, 144–147T
catalytic triad, 185F
cellobiose, 237–238, 239F
cells, 17–26
cytosols, 23, 26F
E. coli, 17F, 23–24, 26F
diffusion in, 34F
eukaryotic, 18–23F
living, 23–26
prokaryotic, 17–18F
structure of, 17–23
solubility and concentrations of, 34F
cellular pathways, 302–304
cellulose, 243F
cellulose, 7–8F
Celsius scale (°C), units of, 26–27
Central Dogma, 3
cerebrosides, 265, 266F
ceremide, 264, 265F
chain elongation, 603, 679–684
DNA polymerase replication, 604–606F
protein synthesis translation, 673–674, 679–684
aminoacyl-tRNA docking sites for, 680–681F
elongation factors, 680–681F
microcycle steps for, 679–684F
peptidyl transferase catalysis, 681–682F
ribosomes and, 673–674
translocation of ribosome, 682–684F
RNA polymerase catalyzation, 636–637F
chair conformations, 189–190F, 235F
Chance, Britton, 420
Changeaux, Jean-Pierre, 157
channels for (animal) membrane transport, 279–280F
chaotropes, 36
chaotropic agents for denaturation, 111
chaperones, *see* molecular chaperones
Chargaff, Erwin, 579
charge–charge interactions, 37, 117, 584
chemiosmotic theory, 420–423
chemoautotrophs, 303, 439–440
chemoheterotrophs, 303
chemotaxis, 284
chiral atoms, 56–57
chiral compounds, 228–230F
chitin, 244F
Chlamydomonas sp., 458
chloride (Cl), 3
chlorophylls, 444–447F
antenna, 446–447F
photon (energy) absorption, 445–446
resonance energy transfer, 446
special pair, 446–447F
structure of, 444–445
chloroplasts, 21–22F, 458–460F
ATP synthase, 459–460F
cyanobacteria evolution of, 459
eukaryotic cell structure and, 20F, 21–22F
organization of, 459–460F
photosynthesis and, 22
structure of, 458–459F
cholecalciferol (vitamin D), 218–219F
cholesterol, 266–268
isoprenoid metabolism and, 490, 493–494F
level regulation, 493
lipid bilayers, 277–278F
lipid metabolism and, 488, 490–494
membrane fluidity and, 277–278F
steroids and, 266–268
synthesis of, 488, 490–494
chromasomal DNA replication, 602–603
chromatin, 588–591
bacterial DNA packaging, 590
higher levels of, 590
histones, 588–590F
nucleosomes and, 588–591
packing ratio, 588
RNA eukaryotic transcriptions and, 649
chromatography, 69–70F, 73–74F
amino acid analysis, 73–47F
techniques, 69–70F
chymotrypsin, 76–77F, 183–188F
Ciechanover, Aaron, 533
cis conformation, 91F, 93, 258, 259F
cis/trans isomerization, 93, 104F
cistine, formulation of, 60F
citrate synthase, citrus cycle reactions, 385F, 394–396F
citric acid cycle, 303–304, 326F, 385–416
amphibolic pathways, 407–409
ATP production, 405–406F
bacteria and, 411–414
coenzyme reduction, 405–406F
energy production in, 405T
enzymatic reactions of, 392
enzyme reactions, 386, 394–402
aconitase, 396–397F
α-ketoglutarate dehydrogenase complex, 398–399F
citrate synthase, 394–396F
conversion of from another, 402F
fumarase, 401
isocitrate dehydrogenase, 397–398F
malate dedrogenase, 401–402
succinate dehydrogenase complex, 399–401F
succinyl synthetase, 398–400F
eukaryotic cells and, 385
evolution of, 412–414
forked pathways, 413F
gloxylate pathway, 409–412
glucose synthesis from, 326F
glycolytic pathway, 408
history of, 385–386
metabolic pathway, 303–304
oxidation of acetyl CoA, 385, 391–394
prochiral substrate binding, 397
pyruvate conversion to acetyl CoA, 385, 387–391
pyruvate entry into mitochondria, 402–405F
regulation of, 406–407
cleavage, 76–77F, 112F, 163–164
bonds, 112F, 163–164
carbocation, 164
enzyme reactions and, 163–164
free radicals, 164
hydrolysis, 592F, 594F
nuclease sites, 592F
proteins by cyanogen bromide (CNBr), 76–77F
RNA, 594F
Cleland, W. W., 147
cobalamin (vitamin B_{12}), 215–216F
codons, 665–670T
anticodons, 668–671F
base pairing, 669–670T
defined, 665
genetic code, 665–668F
initiation, 667, 675–679F
mRNA reading frames, 666–667F
protein synthesis and, 665–684
RNA translation and, 675–679F
synonymous, 667
termination (stop), 667, 682, 684
translation of in chain elongation, 679–684F
wobble positions, 670–671F
coenzymes, 196–226, 316–321
acyl carrier protein (ACP), 204–206F
adenosine triphosphate (ATP), 198–199F, 405–406F
ascorbic acid (vitamin C), 209–211
biotin (vitamin B_7), 211–212F
citric acid cycle, 405–406F
cobalamin (vitamin B_{12}), 215–216F
coenzyme A, 204–206F
cofactors, 196F
cosubstrates, 197–199
cytochromes, 221–222F
electron transfer for free energy, 319–320
energy conservation from, 316–320
flavin adenine dinucleotide (FAD), 204–205F
flavin mononucleotide (FMN), 204–205F
Gibbs free energy change, ΔG, 317–319
half-reactions, 317–319T
inorganic cations, 197
lipid vitamins, 217–219F
lipoamide, 216–217F
mechanistic roles, 199T
metabolic roles of, 198–200T
metal-activated enzymes, 197
metalloenzymes, 197
NADH reactions, 319–320
nicotinamide adenine dinucleotide (NAD), 196F, 200–203F
nicotinamide adenine dinucleotide phosphate (NADP), 200–202F
nobel prizes for, 223
nucleotides, 198–199
oxidation–reduction, 221F, 316–320
prosthetic groups, 197, 205–206F
proteins as, 221
pyridoxal phosphate (PDP), 207–209F
reactive center, 196
reduced, 316–320, 405–406F
reduction potential, 317–319T
riboflavin, 204–205F
tetrahydrofolate, 213–214F
thiamine diphosphate (TDP), 206–207F
ubiquinone (coenzyme Q), 219–221F
vitamins, 196, 198–199T
cofactors, 196F, 425
Cohen, Stanley N., 597
coiled–coil motif (structure), 100F
collagen, 119–121F
covalent (bond) cross links in, 120–121F
interchain hydrogen bonding in, 120F
protein structure, study of, 119–121F
residue formation and, 120–121F
Schiff bases, 121F
type III triple helix, 119F
column chromatography, 69–70F
compartmentation, 304–305
complementary base pairing, double-helix DNA, 582–583F
concanavalin A (Jack bean), 104F
concerted (symmetry) model for enzyme regulation, 156–157F

configurations versus conformations, 234
conformational changes from oxygen binding, 124–126F
conformations versus configurations, 234
CorA, magnesium pump, 280–281F
Corey, Robert, 94
Cori, Gerty and Carl, 369–370, 375
Cori cycle, 360F
Cori ester, 369–370F
coronary heart disease and lipoprotein lipase, 507
cosubstrates, 197–199
cotranslational modifications, 690–691
coupled photosystems, 453–455T
covalent bonds, 37–38F, 120–121F, 392
  citric acid cycle, 392
  collagen protein structure, 120–121F
  hydrogen bonds and, 37–38F
covalent catalysis, 169–170F
covalent modification, 158F
crassulacean acid metabolism (CAM), 471–472F
Crick, Francis H. C., 3, 573–574, 601, 635, 665, 669
Critical Assessment of Methods to Protein Structure Prediction (CASP), 116
cyanobacteria evolution of chloroplast photosystems, 459
cyanogen bromide (CNBr), 76–77F
cyclic adenosine monophosphate (cAMP), 287–288F
  regulatory protein activation of RNA transcription, 653–655
cyclic electronic transfer, 452–453
cyclic guanosine monophosphate (cGMP), 287
cyclization of monosaccharide, 230–234
  anomeric carbon, 231
  anomers, 231
  furanos, 231F
  Haworth projections for, 232–234F
  pyranos, 231F
cysteine (C, Cys), 60F
  catabolism of, 540–541F
  nomenclature, 64T
  structure of, 60F
  synthesis of, 523–525F
cysteine desulfurate (IscS) interactions, 111F
cystinuria, 544
cytidine triphosphate (CTP) synthesis, 559–560F
cytochrome *bf* complex, 453–455F
cytochrome b_{562}, 104F
cytochrome *c*, 79–81F, 101F
  protein structure conservation, 101F
  sequencing, 79–81F
cytochrome *c* oxidase (electron transfer complex IV), 431–432F
cytochromes, 221–222F
cytoplasm, 34F
cytosine (C), 8
  hydrogen bonding, 38F
cytoskeleton, 20F, 23
cytosols, 20F, 23, 26F, 691F

D

D-amino acids, 57–58F
D arm, 668–669F
Dam, Henrik Carl Peter, 223
dark reactions, 443
Darwin, Charles, 15
degenerate genetic code, 667
degradation, *see* catabolism
dehydrogenases enzymes, 136, 203F
Delbruck, Max, 18
denaturation, 110–114F
  chemical, 111–114
    chaotropic agents, 111
    cleavage of bonds, 112F
    detergents, 111–112
    disulfide bonds and bridges, 112F
  double-stranded DNA, 584–585F
  enzyme catalyzation, 113–114
  heating, 111F
  melting curve, 584–585F
  proteins, 110–114F
  renaturation and, 112–113F
deoxy sugars, 235–236F
deoxyhemoglobin, 123
deoxymyoglobin, 123
deoxyribonucleic acid, *see* DNA
deoxyribose, 8F, 574F
deoxythymidylate (dTMP) production, 560–564F
deoxyuridine monophosphate (dUMP) methylation, 560–564F,
detergents, 36F
  denaturation by, 112
  solubility of, 36F
diabetes mellitus (DM), 381, 511
  lipid metabolism and, 511
dialysis, 69
dichloroacetate (DCA), 408F
Dickerson, Dick, 89
dideoxynucleotides for DNA sequencing, 616, 618
dietary lipids, absorption of, 505
diffusion, 34F, 275–276
  facilitated, 281
  lateral, 275F
  lipids in membranes, 275–276F
  membrane transport and, 281
  solubility and, 34F
  transverse, 275–276F
diffusion-controlled reactions, 171–175
  energy diagrams for, 174F
  substrate binding speed and, 171–172T
  superoxide dismutase, 175F
  triose phosphate isomerase (TPI), 172–174F
dihydrofolate, 213F
dihydroxyacetone, 228F, 231F, 236F
dihydroxyacetone phosphate, 332–333F
1,25 dihydroxycholecalciferol, 218F
dipeptide, 6F, 68
diploid cells, 20
disaccharides, 236–239
  cellobiose, 237–238, 239F
  glucosides and, 236–239, 241F
  glycosidic bonds in, 236–238F
  lactose, 238, 239F
  maltose, 237, 239F
  nucleosides and, 239, 241F
  reducing and non reducing sugars, 238–239
  structures of, 237–239F
  sucrose, 238, 239F
discontinuous DNA lagging strand synthesis, 608F
dissociation constant, K_d, 109
  acid solutions, K_a, 44–48T
disulfide bonds and bridges, 112F
DNA (deoxyribonucleic acid), 3, 8–9F, 601–633
  A-DNA, 585–586F
  absorption spectrum of, 584–585F
  amplification of, 615–616
  bacterial, 3, 590
  ball-and-stick model, 582–584F
  base composition of, 579T
  B-DNA, 582–584F, 586F
  chromatin, 588–591
  cloning vectors, 597–598F
  degradation, 373
  discovery of, 3
  double helix, 581–585
  double-stranded, 579–586
    anti-parallel strands, 581–583
    charge–charge interactions, 584
    chemical structure of, 581F
    complementary base pairing, 582–583F
    conformations of, 585–586F
    denaturation of, 584–585F
    hydrogen bonds in, 584
    hydrophobic effects, 584
    major and minor grooves in, 582–583F
    phosphodiester linkages (3–5′) in, 580–581F
    stability from weak forces, 583–585F
    stacking interactions, 582–583F, 585T
    sugar-phosphate backbones of,
    van der Waal forces on, 39
    ultraviolet light absorption, 584–585F
  eukaryotic cells and, 20
  fingerprints, 596–597F
  phosphodiester linkages in, 8–9F
  gene mutation, 322, 447, 469
  histones, 588–590F
  homologous recombination, 626–631
  hydrogen bonds in, 37–38F
  hydrolysis of, 593–596F
    *Eco*RI and, 595–596F
    nucleases and, 593–596F
    restriction endonucleosis and, 593, 595T
  history of, 601–602
  loops for attachment of, 590, 652F
  melting point, T_m, 584
  modified nucleotides, 564–565F
  nucleic acid and, 573–574
  pulling to fully extended form, 588F
  recombinant, 597–598F
  repair of damaged, 622–652
  restriction maps, 596
  sequencing of, 616–619F
  single-strand, 588
  space-filling model, 573F, 582–584F
  sticky ends on, 598
  structure of, 8–9F
  supercoiled, 586–587F
  synthesis, 373
  Watson–Crick model, 579
  Z-DNA, 586F
DNA repair, 622–625
  breast cancer and, 630
  excision, 624–625F
  photodimerization (direct repair), 622–623
DNA replication, 602–622
  base pairing in, 604–606
  bidirectional, 602–603F
  chromasomal, 602–603
  eukaryotes, 619–622
  forks, 602–603, 606, 608F, 613F
  initiation (origin) of, 615F
  polymerase chain reaction (PCR), 615–617F
  polymerases, 603–615
    chain elongation, 604–606F
    interactions, 111F
    nucleotide-group-transfer reaction, 604–605
    proofreading for error correction, 607
    protein types, 603–604T
  replisome model, 610, 612–615
  semiconservative, 602F
  sequencing, 616–619F
    dideoxynucleotides used for, 616, 618
    parallel DNA by synthesis, 618–619
    Sanger method, 616, 618
  synthesis of polymerases, 607–615
    binding fragments, 609–611F
    discontinuous, 608F
    Klenow fragment, 609–610F
    lagging strands, 608–609F, 613–614F
    Okazki fragments, 608–611F
    phosphodiester linkage, 610, 612F
    RNA primer for, 608–609
    single-strand binding (SSB) protein, 613F
    two strands simultaneously, 607–615
  termination (terminus) of, 615F

dnaA gene encoding, 615
Dobzhansky, Theodosius, 15
Doisy, Edward Adelbert, 223
domains, protein structure and, 101–102, 106F
Donahue, Jerry, 575
donepezil hydrochloride, 134F
double bonds, Δ*n*, in fatty acids, 258–259
double helix, 581–585
  anti-parallel strand formation of, 581–583
  B-DNA, 582–584F
  major and minor grooves in, 582–583F
  stability from weak forces, 583–585F
double membranes, 273F
double–reciprocal (Lineweaver–Burk) plot, 146–147F
double-stranded DNA, 579–586
  anti-parallel strands, 581–583
  charge–charge interactions, 584
  chemical structure of, 581F
  complementary base pairing, 582–583F
  conformations of, 585–586F
  denaturation of, 584–585F
  hydrogen bonds in, 584
  hydrophobic effects, 584
  major and minor grooves in, 582–583F
  phosphodiester linkages (3–5′) in, 580–581F
  stability from weak forces, 583–585F
  stacking interactions, 582–583F, 585T
  van der Waal forces on, 39
  ultraviolet light absorption, 584–585F
Drosophila melanogaster, 86, 296, 603F

E
E site (exit site), 682–684F
*Eco*RI, hydrolysis and, 595–596F
Edidin, Michael A., 276
Edman, Pehr, 74
Edman degradation procedure, 74–75F
effector enzymes, 285
eicosanoids, 268–269F
  structures of, 268–269F
  synthesis of, 483–486F
Eijkman, Christiaan, 198, 223
elastase, 183–185F
electrochemical cell, 317F
electrolytes, 32–34
electromotive force, 317
electron micrographs, 284, 603F
electron transfer, 319–320, 455–457
  bacterial photosystems, 449–453
  cyclic, 452–453
  free energy, 319–320
  noncyclic, 452
  photosynthesis, 449–453, 455–457
  Z-scheme, 455–456F
electron transport, 417–442
  adenosine triphosphate (ATP) synthesis and, 417–442
  chemoautotroph energy from, 439–440
  cofactors, 425
  enzyme complexes, 423–435
    complex I (NADH to ubiquinone catalysis), 426–427F
    complex II (succinate:ubiquinone oxidoreductase), 427–428F
    complex III (ubiquino1:cytochrome *c* oxidoreductase), 428–430F
    complex IV (cytochrome *c* oxidase), 431–432F
    complex V (ATP synthase), 433–435F
  Gibbs free energy change, Δ*G*, 423–425T
  NADH shuttle mechanisms in eukaryotes, 436–439F
  oxidation–reduction reactions, 423–425T
  oxygen uptake in mitochondria, 421F
  P/O (phosphorylated/oxygen) ratio, 436
  photosynthesis compared to, 439
  protonmotive force, 421–420F
  Q-cycle electron pathway, 430
  reduction potentials of oxidation–reduction components, 425T
  superoxide atoms, 440–441
  terminal electron acceptors and donors, 439–440
electrophiles, 39–40, 163
electrospray mass spectrometry, 72
electrostatic repulsion, 309
elongation, *see* chain elongation
Embden, Gustav, 331
Embden–Meyerhof–Parnas pathway, 331
enantiomers, 56
endo-envelope conformations, 234F
endocytosis, membrane transport and, 283–284F
endonucleases, defined, 591
endoplasmic reticulum (ER), 20–21F, 691F
endosymbiotic origins, 22
energy, 10–15
  activation, $G^{\ddagger}$, 14F
  bioenergetics, 11
  citric acid cycle, conserved in, 405T
  equilibrium and, 12–15
  flow of, 11F
  Gibbs free energy changes, 12–15
  living organisms and, 10–11
  metabolism, 11
  NADH oxidation–reduction, conservation from, 316–320
  photosynthesis and, 11F
  protein synthesis expense of, 684–685
  reaction rates, 11–12, 14–15
  thermodynamics, 12–13
energy equation, photon of light, 445, 445
energy-rich compounds, 310
enolase reactions, 338
enolpyruvate, 315F
enthalpy, *H*, 12
enthalpy changes, Δ*H*, 12–13, 306
Entner-Doudoroff (ED) pathway, 351–352F
entropy, *S*, 12
entropy change, Δ*S*, 12–13, 306
enzyme reactions, 386, 392, 394–402
  aconitase, 396–397F
  α-ketoglutarate dehydrogenase complex, 398–399F
  citrate synthase, 394–396F
  citric acid cycle, 386, 392, 394–402
  conversion of from another, 402F
  fumarase, 401
  isocitrate dehydrogenase, 397–398F
  malate dedrogenase, 401–402
  succinate dehydrogenase complex, 399–401F
  succinyl synthetase, 398–400F
enzyme–substrate complex (ES), 139–140, 142–143
enzymes, 2, 6–7F, 134–161, 162–195. *See also* coenzymes; substrates
  activation energy lowered by, 165–166F
  allosteric, 153–158F
    concerted (symmetry) model for, 156–157F
    phosphofructokinase, 154–155F
    properties of, 155–156F
    regulation of enzyme activity using, 153–158
    sequential model for, 157–158F
  ammonia transfer from glutamate, 558
  catalytic proficiency of, 144–147T
  catalytic constant, k_{cat}, 143–145
  catalysts, 2, 113–114, 134
  chemical reaction rates and, 15
  cell cytosol behavior of, 23, 26F
  citric acid cycle reactions, 386, 394–402
  classes of, 136–138
    oxidoreductases, 136
    transferases, 136–137, 395
    number system for, 137F
    hydrolases, 137
    lyases, 137
    isomerases, 137–138
    ligases, 138
  cofactors, 196F
  conversion of from another, 402F
  covalent modification of, 158F
  defined, 135
  electron transport, 423–435
    complex I (NADH to ubiquinone catalysis), 426–427F
    complex II (succinate:ubiquinone oxidoreductase), 427–428F
    complex III (ubiquino1:cytochrome *c* oxidoreductase), 428–430F
    complex IV (cytochrome *c* oxidase), 431–432F
    complex V (ATP synthase), 433–435F
  glycolysis, reactions of, 326–327T
  gluconeogenesis regulation, 363–364F
  inhibition, 148–153
    competitive, 149–150F
    constant, K_i,148
    irreversible, 152–153F
    noncompetitive, 149–151F
    pharmaceutical uses of, 151–152
    reversible, 148–152F
    uncompetitive, 149–150F
  inorganic cations and, 197
  kinetic constant, k_m, 144–147, 149T
  kinetics and, 23, 138–149
  lock-and-key theory of specificity, 180
  mechanisms of, 147, 162–195
    arginine kinase, 190–192F
    catalysis, 166–182
    cleavage reactions, 163–164
    diffusion-controlled reactions, 171–175
    lysozyme, 189–191F
    nucleophilic substitution, 163
    oxidation–reduction reactions, 164
    serine proteases, 183–189F
    transition states, 163, 164–166
  metal-activated, 197
  metabolite channeling, 158–159
  Michaelis–Menton equation for, 140–144
  multienzyme complexes, 158–159
  multifunctional, 158–159
  multisubstrate reactions, 147–148F
  pH and rates of, 170–172F
  properties of, 134–161
  protein structures and, 6–7F, 113–114
  reactions, 134–136F, 138–140F, 147–148
  regulation of, 153–158
  substrate binding and, 171–172T, 175–182F
epimers, 230
epinephrine, structure of, 63F, 199F
equilibrium, 11–15
  acid dissociation constant, K_a, 44–48
  association constant, K_a, 109–110F
  buffered solutions, 51–52
  constant, K_{eq}, 12, 14
  dissociation constant, K_d, 109
  energy and, 12–15
  Gibbs free energy change, Δ*G*, 12–15, 307–308
  metabolic changes and, 307–308
  near-equilibrium reaction, K_{eq}, 307–308
  protein–protein interactions, 109–110
  rate changes and, 11–12
erythrose, 229
erythrulose, 231F
Escherichia coli (*E. coli*), 17F, 23–24, 26F, 86F, 106, 108T
  allosteric enzyme regulation and, 154–155F

audioradiograph of replicating chromosome, 603F
carbamoyl phosphate synthetase, 558F
cells, 17F, 23–24, 26F
chaperonin (GroE), 118–119F
covalent catalysis, 169–170F
cytochrome b_{562}, 104F
flavodoxin, 105F
gloxylate pathway, 411–412
homologous recombination, 627–630
L-arabinose-binding protein, 105F
metabolic network of, 295–296
oligomeric proteins, 106, 108T
phosphofructokinase, 154–155F
ribosome, 665F, 647–675F
RNA content in, 636T
structure of, 17F, 104F
thiol-disulfide oxidoreductase, 105F
transketolase, 368F
trp operon, 688–690F
tryptophan biosynthesis enzyme, 105F
UDP N-acetylglucosamine acyl transference, 104F
essential amino acids, 529T
essential ions, 196
ester linkages, 4–5F
ethanol, pyruvate metabolism to, 339–340F
ether, synthesis of, 487F
eukaryotes, 15–16F
chromatin and, 649
DNA replication in, 619–622
evolution and, 15–16F
glucose synthesis in, 369–370F
initiation factors, 677, 679F
mRNA processing, 656, 658–663
NADH shuttle mechanisms in, 436–439
protein synthesis and, 674–677, 679F, 691–692F
polymerases, 646–648T
ribosomes, prokaryotic cells compared to, 674–675F
RNA transcription, 646–649
secretory pathways in, 691–692F
transcription factors, 648–649T
eukaryotic cells, 18–23F
citric acid cycle and, 385
chloroplasts, 21–22F
compartmentalization, 501–502
cytoskeleton, 23
DNA and, 20
endoplasmic reticulum (ER), 20–21F
Golgi apparatus, 21F
lipid metabolism and, 501–502
metabolic pathways in, 305F
mitochondria, 21–22F
mitosis, 20F
nucleus of, 20
organelles, 19–20F
structure of, 19–20F
vesicle specialization, 22
eukaryotic DNA polymerase, 620T
eukaryotic enzymes, 364F
eukaryotic (plant) photosystems, 458–461
ATP synthase, 459–460F
chloroplasts, 458–460F
cyanobacteria evolution of, 459
organization of components, 459–460F
eukaryotic ribonucleotide reductase, allosteric regulation of, 561T
eukaryotic transducers, 285
evolution, 15–17, 57–58
amino acids and, 57–58
bacterial enzymes, 364F
biochemistry and, 15–17
common ancestors, 57–58
cyanobacteria effects on chloroplast photosystems, 459
cytochrome *c* sequences, 79–81F
endosymbiotic origins, 22
eukaryotes, 15–16F
last common ancestor (LCA), 57–58
metabolic pathways, 301–302
mitochondria and chloroplasts, 459
phylogenetic tree representation, 79–80F
prokaryotes, 15–16F
protein primary structure, 79–81
exit site (E site), 682–684F
exocytosis, membrane transport and, 283–284F
exons, 660
exonucleases, 591
extreme thermophiles, 30F

F

facilitated diffusion, membrane transport and, 281
fat-soluble vitamins, 198
fatty acids, 9, 257–261
anionic forms of, 258T
cis configuration, 258, 259F
coenzymes and, 215, 221
dietary requirements and, 261
double bonds, Δn, in, 258–259
lipid structure of, 258–261
micromolecular structure of, 9
nomenclature, 257–258T
oxidation of, 494–501
acyl CoA synthase activation, 494
ATP generation from, 498–499
β-oxidation, 494–501F
mitochondria transport, 479–498
odd-chains, 499–500
unsaturated, 500–501
polyunsaturated, 258, 260F
saturated, 258, 260F
synthesis of, 475–481, 497F
activation reactions, 479F
β-oxidation and, 497F
desaturation, 479–481
elongation reactions, 477–479F
extension reactions, 479–481
initiation reaction, 477F
trans configuration, 258, 259F
unsaturated, 258, 260F
feed-forward activation, 300
feedback inhibition, 300
Fenn, John B., 73
fermentation process, 340F
fibrous proteins, 86, 119–121. *See also* collagens
Filmer, David, 157
fingerprints, 77–79F, 596–597F
DNA restriction endonucleases, 596–597F
tryptic, sequencing use of, 77–79F
Fischer, Edmund (Eddy) H., 375–376
Fischer, Emil, 2, 3, 180
Fischer projections, 7F, 228–232F
aldoses, 228–230F
ketoses, 230–231F
monosaccharide carbohydrates, 228–232F
trioses, 228F
flavin adenine dinucleotide (FAD), 204–205F
flavin mononucleotide (FMN), 204–205F
flavodoxin, 105F
Flemming, Walter, 585
fluid mosaic model, 274–275
fluorescent protein (jellyfish), 104F
flux in metabolic pathways, 300F
FMN oxidoreductase (yeast), 105F
folate (vitamin B_9), 213–214F
folding, 99–103F, 114–119F
aggregation from, 119
CASP, 116
characteristics of, 114–115F
charge–charge interactions and, 117
hydrogen bonding and, 115–116F
hydrophobic effect and, 114–115
molecular chaperones and, 117–119F
pathways, 114–115F
protein stability and, 99–103F, 114–119F
tertiary protein structure and, 99–103
van der Waals interactions and, 117
forked pathways, 413F
formamidoimidazole carboxamide ribonucleotide (FAICAR), 553F
formylglycinamide ribonucleotide (FGAR), 553F
formylglycinamidine ribonucleotide (FGAM), 553F
N-formylmethionine, structure of, 62–63F
fractional saturation, 124–125F
Franklin, Rosalind, 579
free-energy change, *see* Gibbs free energy change, ΔG
free radicals, 164
ribonucleotide reduction, 562
freeze-fracture electron microscopy, 276–277F
fructose, 231F
conversion to glyceraldehyde 3-phosphate, 348–349
gluconeogenesis regulation, 363–364F
invertase conversion to, 349
fructose 1,6 *bis*phosphate, 332F, 358–359F
fructose 6-phosphate, 330–331F, 358–359F
gluconeogenesis conversion, 358–359F
gluconeogenesis regulation, 363–364F
glycolysis conversion, 330–331F
Frye, L. D., 276
fuel metabolism, 295
fumarase, citrus cycle reactions, 401
fumarate, urea cycle and, 543F, 545–546F
Funk, Casimir, 198
furanos, 231F, 234
Furchgott, Robert F., 530

G

G proteins, 285–286F, 290
galactose, 229F
conversion to glucose 1-phosphate, 349–350
galactose mutarotase, 234F
galactosides, 239, 241F
γ-aminobutyrate, structure of, 63F
gamma crystallin (cow), 104F
Gamow, George, 666
gangliosides, 265, 266F
gel-filtration chromatography, 69–70
gene, defined, 634
gene mutation, 322, 447, 469
gene orientation, 639–640F
gene regulation, 649–651, 685–690
protein synthesis, 685–690
attenuation, 688–689F
globin regulation by heme availability, 687–688F
ribosomal assembly in *E. coli*, 685–687F
trp operon in *E. coli*, 688–690F
RNA transcription and, 649–651
gene sequences, metabolism and, 295–296
genetic code, 665–668T
codons, 665–668T
degenerate, 667
history of, 665–667F
mRNA and, 666–667F
reading frames, 666–667F
tRNA and, 666, 668–670F
genetic defects, sphingolipids and, 265–266
genetically modified food, 528
genome, defined, 573
gibberellins, 270
Gibbs, Josiah Willard, 12
Gibbs free energy change, ΔG, 12–15, 341–342F
actual, 306, 341–342F
adenosine triphosphate (ATP), 308–312
electron transport, 423–425T

enthalpy changes, Δ*H*, and, 306
entropy changes, Δ*S*, and, 306
formation of reactants, 308T
glycolysis reactions, 332, 341–342F
hydrolysis, 308–312
mass action ratio, *Q*, and, 306
membrane transport and, 278–279
metabolic reaction direction from, 306–312
metabolically irreversible reactions, 307, 308–312
near-equilibrium reaction, K_{eq}, 307–308
oxidation–reduction reactions, 316–320
photosynthesis photosystems, 455–457
reduction potential and, 317–319T
standard, 306, 341–342T
thermodynamic reactions and, 12–15, 278–279
globin protein synthesis regulation, 687–688
globular proteins, 86, 122–129. *See also* hemoglobin; myoglobin
gloxylate pathway, 409–412
glucokinase, 344–345F
glucolfuranose, 233F
gluconeogenesis, 303, 326F, 355–384
Cori cycle, 360F
fructose 1,6 *bis*phosphate, 358–359F
glucose level maintenance (mammals), 379–381
glucose 6-phosphatase, 359–360
glucose synthesis by, 326F
glycogen metabolism, 369–372
glycogen regulation (mammals), 372–379
glycogen storage diseases, 381–382
glycolysis compared to, 356–357F
hormone regulation of, 376, 378–379F
metabolic pathway, 303
pentose phosphate pathway, 364–369
phosphoenylpyruvate carboxykinase (PEPCK) reactions, 358F
precursors for, 360–363
acetate, 362–363
amino acids, 360–361
glycerol, 360–361F
lactate, 360, 361–362
propionate, 361–362
sorbitol, 362
pyruvate to glucose conversion, 356–360
pyruvate carcoxylase reaction, 357–358F
regulation of, 363–364, 376–379F
L-glucono-gamma-lactone oxidase (GULO), 210–211F
glucopyranose, 232F, 239F
glucose, 7–8F, 229–230F, 236F
cyclization of, 231–234F
diabetes mellitus (DM) and, 381
glycolysis, 325–354
hemeostasis phases, 380F
liver metabolic functions and, 379–380F
maintenance of levels in mammals, 379–381
monosaccharide structures of, 229–230F, 236F
pyruvate conversion via gluconeogenesis, 356–360F
pyruvate conversion via glycolysis, 328–329F, 338–340F
solubility of, 34F
sorbitol conversion, 362G
starch and, 240–242F
storage as starch and glycogen, 240–243F
structure of, 7–8F, 34F
sugar acids derived from, 238F
sugar phosphate structures, 236F
glucose-alanine cycle, 361F
glucose 1-phosphate, galactose conversion to, 349–350
glucose 6-phosphatase, 359–360
glucose 6-phosphate dehydrogenase deficiency, 367F
glucose 6-phosphate isomerase catalysis, 327, 330–331F, 345F
glucose 6-phosphate, liver metabolic functions and, 345F
glucosides, 236–239, 241F
glucuronate, 238F
glutamate (E, Glu), structure of, 62F
ammonia incorporated in, 518F
catabolism of, 535
enzyme transfer of ammonia from, 558
ionization of, 65–66F
malate–aspartate shuttle, 348F
metabolic precursor use, 529
nomenclature, 64T
phosphorol group transfer, 312–313
structure of, 62F
synthesis of, 312–313, 523F
transferases catalyzation, 136–137
urea cycle and, 545–546F
phosphorol group transfer, 312–313F
glutamine (Q, Gln), structure of, 62F
ammonia incorporated in, 518F
catabolism of, 535
ligases catalyzation, 138
metabolic precursor use, 529
nomenclature, 64T
structure of, 62F
synthesis of, 312–313, 523F
glycan, 227
glyceraldehyde, 228–229F, 236F
glyceraldehyde 3-phosphate, 332–334F
fructose conversion to, 348–349
shuttle mechanisms in eukaryote, 437F
glyceraldehyde 3-phosphate dehyrogenase, 333–334, 346–347F
glycerol, 360–361F
glyoxylate cycle, 361
gluconeogenesis precursor, 360–361F
oxidation of, 361F
glycerol 3-phosphate, 9–10F
micromolecular structure of, 9–10F
oxidation of, 361F
glycerol 3-phosphate dehyrogenase, 361F
glycerophospholipids, 6–10F, 262–265
micromolecular structure of, 9–10F
phosphatidates, 262–264F
plasmalogens, 263, 265F
synthesis of, 481–483F
types of, 263T
glycinamide ribonucleotide (GAR), 553F
glycine (G, Gly), 59F, 65–4T
catabolism of, 536–537F
metabolic precursor use, 529–530F
nomenclature, 64T
structure of, 59F
synthesis of, 523–524F
glycine encephalopathy, 544
glycoconjugates, 244–252
cartilage structure, 245–246F
glycoproteins, 248–252F
glycosaminoglycans, 244–245F
oligosaccharides, 248–252F
peptidoglycans, 246–248F
proteoglycans, 244–246F
glycogen, 240–243F, 369–382
cleavage of residues, 371–372F
degradation of, 371–372F, 373–374F
glucose level maintenance (mammals), 379–381
glucose storage (animals), 240–243
hormone regulation of, 376–379
linkages, 242–243F
Mendelian Inheritance in Man (MIM) numbers, 381–382
metabolism, 369–372
molecule, 371F
phosphorolysis reaction, 371–372F
regulation of (mammals), 372–379, 374F
storage diseases, 381–382
synthase reaction, 370–371F
synthesis of, 369–371F
glycogen phosphorylase, 373–374F
degradation of, 373–375F
phosphorylated state (GPa), 375F
unphosphorylated state (GPb), 347–375F
glycolysis, 303, 325–354
aldolase cleavage, 330–332F
enolase reactions, 338
Entner-Doudoroff (ED) pathway, 351–352F
enzymatic relations of, 326–327T
fructose conversion to glyceraldehyde 3-phosphate, 348–349
galactose conversion to glucose 1-phosphate, 349–350
Gibbs free energy change, Δ*G*, 341–342T
gluconeogenesis compared to, 356–357F
glucose catabolism, 325–354
glucose 6-phosphate isomerase catalysis, 327, 330–331F, 345F
glucose synthesis by, 326F
glucose to pyruvate conversion by, 328–329F
glyceraldehyde 3-phosphate dehyrogenase catalysis, 333–334
hexokinase reactions, 326–327, 328F, 330F
history of, 331
hormone regulation of, 376, 378–379F
mannose conversion to fructose 6-phosphate, 351
metabolic pathway, 303
phosphofruktokinase-1 (PFK-1) catalysis, 330
phosphoglycerate kinase catalysis, 335–336
phosphoglycerate mutase catalysis, 336–337F
pyruvate kinase catalysis, 338
pyruvate metabolic functions, 338–340F
metabolism to ethanol, 339–340F
reduction to lactate, 340
regulation of, 343–347
hexokinase, 344–345
hexose transports, 343–344
metabolic pathway in mammals, 343F
Pasteur effect for, 347
phosphofruktokinase-1 (PFK-1), 345–346F
pyruvate kinases, 346–347F
sucrose cleaved to monosaccharines, 348
triose phosphate isomerase catalysis, 332–334F
glycolytic pathway, 408
glycoproteins, 248–252F. *See also* oligosaccharides
glycosaminoglycans, 244–245F
glycosides, 241F
glycosidic bonds, 236–238F
glycosphingolipids, 256
glycosylation of proteins, 694F
glyoxylate cycle, 361
Golgi, Camillo, 21
Golgi apparatus, 20–21F, 691F
Goodsell, David S., 23, 34
gout, 569
Gram, Christian, 247
Gram stain, 247F
grana, 458
Greek key motif (structure), 100–101F
green filamentous bacteria, photosynthesis in, 448, 452F
Greenberg, G. Robert, 551, 552
group transfer reactions, 163
growth factors, signal transduction and, 284
guanine (G), 8, 551F
hydrogen bonding, 38F
structure of, 551F
guanosine 5′-monophosphate (GMP), 550–551F
gulose, 229F
gyrate atrophy, 544

H
hairpin formation, RNA transcription, 644F
hairpin motif (structure), 100F
Haldane, J. B. S., 141
half-chair conformation, 189–190F
half-reactions, 317–319T
Haloarcula marismortui, 675, 676F
Halobacterium halobium, 270
Halobacterium salinarium, 461
Hanson, Richard, 359
haploid cells, 20
Harden, Arthur, 331
Haworth, Sir Walter Norman, 223, 232–234
Haworth projections, 7–8F, 232–235F
head growth, 373
heat shock proteins, 117–118F
helical wheel, 95
Helicobacter pylori, 216F
3_{10} helix, 95
helix bundle motif (structure), 100F
helix–loop–helix (helix–turn–helix) structure, 100F
heme,122–126F, 221–222F
  globin protein synthesis regulation, 687–688
  prosthetic groups,122–126F, 221–222F
    absorption spectra, 221–222F
    cytochromes, 221–222F
    hemoglobin (Hg), 122–126F
    myoglobin (Mg), 122–126F
    oxygen binding in, 123–126F
    oxygenation and, 122
hemeostasis phases in glucose, 380F
hemiacetal, 232F
hemiketal, 232F
hemoglobin (Hb), 122–129F
  allosteric protein interactions, 127–129F
  α– and β–globin subunits of, 122–123F
  embryonic and fetal, 126F
  heme prosthetic group, 122–124F
  oxygen binding, 123–129
  protein structure, study of, 122–129F
  protein synthesis regulation by heme availability, 687–688
  tertiary structure of, 122–123F
Henderson–Hasselbach equation, 46–47, 66
Hereditary Persistence of Fetal Hemoglobin (HPFH), 126
Hershko, Avram, 533
heteroglycans, 240
heterotrophs, 302–303
hexokinase, glycolysis regulation of, 344–345
hexokinase reactions, 326–327, 328F, 330F
hexose transports, glycolysis regulation of, 343–344
high-density lipoproteins (HDL), 507–508
high energy bond, ~, 311
high-performance liquid chromatography (HPLC), 69–70F
histamine, structure of, 63F
histidine (H, His), 61F
  catabolism of, 535–536F
  ionization of, 65–66F
  nomenclature, 64T
  structure of, 61F
histones, 588–590F
HIV-1 aspartic protease, 107F
Hodgkin, Dorothy Crowfoot, 88, 215, 223
Holliday, Robin, 626
Holliday junction (model) for DNA recombination, 601, 626–627F
homocysteine, 216F
homoglycans, 240
homologous proteins, 79
homologous recombination, 626–631
  E. coli, 627–630
  Holliday junction (model), 626–627F
  repair as, 631
Hopkins, Sir Frederick Gowland, 223
hopotonic cells, 35F
Hoppe-Seyler, Felix, 573
hormones, 284–287
  adenylyl cyclase binding, 287–288F
  G protein binding, 286
  gluconeogenesis regulation by, 376, 378–379F
  glycogen metabolism regulation, 376–377F
  glycolysis regulation by, 376, 378–379F
  lipid metabolism regulation by, 502–504
  multicellular organism receptor functions, 284–285
  receptor binding, 287–288F
  signal transduction and, 284–287
hydrated molecules, 34
hydrochloric acid (HCL), dissociation of, 44–45
hydrogen (H), 3, 29F
  polarity of water and, 29F
hydrogen bonds, 30–32F, 37–38F
  α helix, 94–97F, 98–99F
  β sheets and strands, 97–99F
  collagen, 120F
  covalent bonds and, 37–38F
  DNA (deoxyribonucleic acid), 37–38F, 584
  double helix, 584
  ice, formation of, 30–31F
  interchain, 120F
  loops and turns stabilized by, 98–99F
  nucleic acid sites, 575–576F
  orientation of, 30–31F
  protein folding and, 115–116F
  protein structures and, 94–99F
  types of, 116T
  water, 30–32F, 37–38F
hydrolases enzymes, 137
hydrolysis, 2, 40F, 73–74F
  adenosine triphosphate (ATP), 308–312
    electrostatic repulsion, 309
    metabolically irreversible changes, 308–312
    resonance stabilization, 310
    solvation effects, 309–310
  amino acid analysis and, 73–74F
    chromotagraphic procedure for, 73–74F
    phenylisothiocyanate (PITC) treatment, 73F
    protein compositions, 74T
  arsenate (arsenic) poisoning and, 336
  Gibbs free energy change, ΔG, 308–312
  nucleic acids, 591–598
    alkaline, 591–592F
    DNA, 593–596F
    *Eco*RI and, 595–596F
    restriction endonucleosis and, 593, 595T
    ribonuclease A, 592–594
    RNA, 591–594F
  macromolecules, 40F
  proteins, 40
  signal transduction and, 285–289F
  thioesters, 316
hydronium ions, 41–43
hydropathy scale, amino acids, 62T
hydrophilic substances, 32
hydrophobic effects, double-stranded DNA, 584
hydrophobic interactions, 39, 98, 114–115
hydrophobic substances, 35, 123–124F
hydrophobicity of side chains, 62
hydroxide ions, 41–43
hydroxyethylthaimine diphoshate (HETDP), 207F
hydroxyl, general formula of, 5F
hydroxylysine residue, 120F
hydroxyproline residue, 120F
hyperactivity, 359
hyperbolic binding curve, 124–126F, 146
hypertonic cells, 35F
hypoxanthine-guanine phosphoribosyl transferase (HGBRT), 107–108F

I
ibuprofen, structure of, 486F
ice, formation of, 30–31F
idose, 229F
Ignarro, Louis J., 530
imazodole ($C_3H_4N_2$), titration of, 47F
immunoglobin, 129–130F
induced-fit enzymes, 179–180
inhibition, 148–153. *See also* regulation
  antibiotics for protein synthesis, 686F
  cancer drugs for, 564
  competitive, 149–150F
  constant, K_i,148
  dichloroacetate (DCA), 408F
  enzyme behavior and, 148–153
  kinetic constant, k_m, effects on, 144–147, 149T
  irreversible, 152–153F
  noncompetitive, 149–151F
  pharmaceutical uses of, 151–152, 408
  phosphorylation, 687–688F
  protein synthesis and, 686–688F
  reversible, 148–152F
  uncompetitive, 149–150F
inhibitors, defined, 148
initiation codons, 667, 675–679F
initiation factors, 675, 677–679F
  eukaryotic cells, 677, 679F
  prokaryotic cells, 677–678F
inorganic cations, 197
inosinate base pairs, 670F
inosine 5′-monophosphate (IMP) synthesis, 551–554F
inositol 1,4,5-*tris*phosphate (IP_3), 287–289F
inositol-phospholipid signaling pathway, 287–289F
insolubility of nonpolar substances, 35–36. *See also* solubility
insulin, 290–291F, 344F
  diabetes mellitus (DM) regulation by, 381
  glycogen metabolism regulation by, 376–377F
  glycolysis regulation by, 344F
  receptors, 290–291F
integral (transmember) proteins, 270–272F
interconversions, pentose phosphate pathway, 368–369F
intermediary metabolism, 294
intermediate-density lipoproteins (IDL), 507
intermediate filaments, 23
intermediates, enzyme transition states and, 165–166F
International Union of Biochemistry and Molecular Biology (IUBMB), 136, 401
International Union of Pure and Applied Chemistry (IUPAC), 257
interorgan metabolism, 304–305
intrinsically disordered (unstable) proteins, 102–103
intron/extron gene organization, 660–662F
introns, 658
invertase, 349
ion-exchange chromatography, 69
ion pairing, 37
ion product, *K*, 42–43
ionic state of side chains, 64–65F
ionic substances, solubility of, 32–35
ionization, 41–43, 63–67
  acids, 42
  amino acids, 63–67
  bases, 42
  Henderson–Hasselbach equation for, 66

ionization (*Continued*)
ion product, *K*, 42–43
pK_a values and, 63–67
titration and, 64–65F
water, 41–43
iron–sulfur clusters, 197–198F
irreversible changes, metabolic, 308–312
irreversible inhibition, 152–153F
isoacceptor tRNA molecules, 670–671
isocitrate dehydrogenase, citrus cycle reactions, 397–398F
isoleucine (I, Ile), 59F, 64T
nomenclature, 64T
stereosomers of, 59F
structure of, 59F
synthesis of, 521–523F
isomerases enzymes, 137–138
isopentenyl diphosphate, cholesterol and, 488, 490
isoprenoid metabolism, cholesterol synthesis and, 490, 493–494F
isoprenoids, 256, 269F
isotonic cells, 35F
IUMBM–Nicholson metabolic chart, 504F

J
Jacob, François, 635
Johnson, W. A., 386

K
Karrer, Paul, 223
Kelvin scale (K), units of, 26–27
Kendrew, John C., 2–3, 88–90, 122
keto group naming convention, 399
ketohexoses, 231F
ketone, general formula of, 5F
ketone bodies, 508–510
lipid metabolism, 508–510
liver functions and, 509–510F
mitochondria oxidation and, 510
ketopentoses, 231F
ketoses, 228–234F
cyclization of, 230–234F
Fischer projections of, 230–231F
structure of, 228–230F
Khorana, H. Gobind, 666
kinases, 158, 301, 314
ATP catalyzation, 310
enzyme regulation by covalent modification using, 158
metabolic pathway regulation and, 301
phosphorol group transfer, 314
kinetic constant, k_m, 144–147, 149T
kinetics, 23, 138–149
catalytic constant, k_{cat}, 143–145
catalytic proficiency, 144–147T
chemical reactions, 138–139F
enzyme properties and, 138–140
enzyme reactions, 139–140F
enzyme–substrate complex (ES), 139–140, 142–143
hyperbolic curve and, 146
kinetic constant, k_m, 144–147, 149T
kinetic mechanisms, 147
Lineweaver–Burk (double–reciprocal) plot, 146–147F
Michaelis–Menton equation, 140–144
multisubstrate reactions, 147–148F
ping-pong reactions, 148–149F
rate (velocity) equations, 138–139, 144–145
reversible inhibitors and, 148–149T
sequential reactions, 148–149F
substrate reactions, 138–147
Klenow fragment, 609–610F
KNF (sequential) model for enzyme regulation, 157–158F
knob-and-stalk mitochondria structure, 433F
Knowles, Jeremy, 174
Kornberg, Arthur, 183, 601, 603, 609
Koshland, Daniel, 157
Krebs, Edwin G., 375–376
Krebs, Hans, 385–386, 397
Krebs cycle, *see* citric acid cycle
Kuhn, Richard, 223

L
L-amino acids, 57–58F
lac operon, 651–655
binding repressor to the operon, 652F
repressor blocking RNA transcription, 651–652F
repressor structure, 652–653F
cAMP regulatory protein and, 653–655F
RNA transcription activation, 653–655
lactate, 360F, 361–362
buildup, 341
Cori cycle, 360F
gluconeogenesis precursor, 360F, 361–362
oxireductases catalyzation, 136
pyruvate reduction to, 340
lactate dehydrogenase, 102F
Lactobacillus, 340
lactose, 238, 239F
lactose intolerance, 350
lagging DNA strand synthesis, 608–609F, 613–614F
Landsteiner, Karl, 250
lateral diffusion, 275F
Leloir, Luis F., 223
Lesch, Michael, 569
Lesch–Nyhan syndrome, 569
leucine (L, Leu), 59F
nomenclature, 64T
structure of, 59F
synthesis of, 521–523F
leucine zipper, 96–97A
leukotrienes, 483, 485–486F
ligases enzymes, 138
light-gathering pigments, 444–448
accessory pigments, 447–448F
chlorophylls, 444–447F
photons (energy), 445–446
resonance energy transfer, 446
special pair, 446–447F
light reactions, 443
lignin synthesis from phenylalanine, 531–532F
limit dextrins, 242
Lind, James, 209–210
Lineweaver–Burk (double–reciprocal) plot, 146–147F
linkages, 4–5F, 8–9F
micromolecular structures of, 4–5F, 8–9F
peptide bonds, 67–68F
phosphate esters, 4–5F, 8
phosphoanhydride, 4–5F, 8F
phosphodiester, 8–9F
linoleate, 481F
lipid anchored proteins, 272–273F
lipid metabolism, 475–513
absorption and, 505–508
dietary lipids, 505
bile salts, 505F
pancreatic lipase action, 505F
lipoproteins, 505–508F
serum albumin, 508
cholesterol, synthesis of, 488, 490–494
isoprenoid metabolism and, 490, 493–494F
level regulation, 493
steps for, 488, 490
diabetes and, 511
eicosanoids synthesis of, 483–486F
ether, synthesis of, 487F
fatty acids, synthesis of, 475–481, 497F
activation reactions, 479F
β-oxidation and, 497F
desaturation, 479–481
elongation reactions, 477–479F
extension reactions, 479–481
initiation reaction, 477F
eukaryotic cell compartmentalization, 501–502
glycerophospholipids, synthesis of, 481–483F
hormone regulation, 502–504
IUMBM–Nicholson metabolic chart, 504F
ketone bodies, 508–510
liver functions and, 509–510F
mitochondria oxidation and, 510
oxidation of fatty acids, 494–501
acyl CoA synthase activation, 494
ATP generation from, 498–499
β-oxidation, 494–501F
mitochondria transport, 479–498
odd-chains, 499–500
unsaturated, 500–501
regulation of, 502–504
sphingolipids, synthesis of, 488–489F
triacylglycerols, synthesis of, 481–483F
lipid vitamins, 217–219F
α-tocopherol (vitamin E), 218F
cholecalciferol (vitamin D), 218–219F
phylloquinone (vitamin K), 218–219F
retinol (vitamin A), 217–218F
lipids, 9F, 256–293. *See also* fatty acids; lipid metabolism; membranes
absorption of, 505–508F
anchored membrane proteins, 272–273F
bilayers, 9, 10F, 269–270, 277–278F
biological membranes, 9–10F, 269–270
cholesterol and, 277–278F
membrane fluidity and, 276–277
phase transition of, 277F
defined, 9
dietary absorption, 505
diffusion of, 275–276F
eicosanoids, 268–269F
fatty acids, 9, 257–261
glycerophospholipids, 262–263T
isoprenoids, 256, 269F
linkages, 4–5F
macromolecular structure of, 9F
prostaglandins, 268–269
rafts, 277
sphingolipids, 263–266F
steroids, 9, 266–268F
structural and functional diversity, 256–257F
transverse diffusion, 275–276F
triacylglycerols, 261–262F
unusual membrane compositions, 274
vesicles (liposomes), 270F, 272F
waxes, 9, 268
Lipmann, Fritz Albert, 223, 311
lipoamide, 216–217F
lipoprotein lipase, coronary heart disease and, 507
lipoproteins, 505–508F
liver metabolic functions, 344–345F, 379–380F
lock-and-key theory of specificity, 180
loop structures, α helix and β strand and sheet connections, 98–99F
low-density lipoproteins (LDL), 507–508
lumen, 457–459F
Luria, Salvatore, 18
lyases enzymes, 137
lypoic acid, 216
lysine (K, Lys), 61F
catabolism, of, 542F
nomenclature, 64T
structure of, 61F
synthesis of, 520–522F

lysosomal storage diseases, 492F
lysosomes, eukaryotic cell structure and, 20F, 22
lysozyme, 6–7, 189–191F
  catalyzation by, 189–161F
  cleavage of, 189F
  conformation of, 186–190
  molecular structure, 6–7F
  reaction mechanism, 190–191F
lyxose, 229F

M
MacKinnon, Roderick, 280
MacLeod, Colin, 3, 573
macromolecules, 4–10
  condensation of, 40–41F
  hydrolysis of, 40F
  linkages, 4–5F, 8–9F
  lipids, 9
  membranes, 9–10
  noncovalent interaction in, 37–40F
  nucleic acids, 7–9F
  polysaccharides, 6–7F
  proteins, 6
  structure of, 4–10
magnesium (Mg), 3
major and minor grooves in double-stranded DNA, 582–583F
malate–aspartate shuttle, 348F
malate dedrogenase, citrus cycle reactions, 401–402
malate dehydrogenase, 102F
MALDI-TOF technique, 72F
maltose, 237, 239F
mammals, metabolic pathway in, 343T
mannose, 229
  conversion to fructose 6-phosphate, 351
maple syrup urine disease, 544
mass action ratio, *Q*, 306
mass spectrometry, 72F, 77–78F
matrix-assisted laser deabsorption ionization (MALDI), 72
Matthaei, J. Heinrich, 337, 666
McCarty, Maclyn, 3, 573
mechanistic chemistry, 162–164. *See also* enzymes
melanin synthesis from tyrosine, 531, 533F
melting curve, denaturation and, 584–585F
melting point, T_m, 584
membranes, 9–10F, 269–293
  biological, 9, 269–275
  chloroplasts, 458–460F
  cholesterol in, 277–278F
  diffusion of lipids, 275–276F
  double, 273F
  dynamic properties of, 275–277
  fluid mosaic model of, 274–275
  fluidity changes, 276–277
  freeze-fracture electron microscopy, 276–277F
  functions of, 269
  glycerol-3 phosphate, 9–10F
  glycerophospholipids, 9–10F
  lipid bilayers, 9, 10F, 269–270, 277–278F
    ampithatic lipids, 270F
    biological membranes, 9–10F, 269–270
    cholesterol and, 277–278F
    leaflets (monolayers) of, 270
    membrane fluidity and, 276–277
    phase transition of, 277F
  lipid rafts, 277
  lipid vesicles (liposomes), 270F, 272F
  macromolecular structure of, 9–10F
  osmotic pressure and, 34–35
  photosynthesis photosystems, 457–460
  plasma, 457F
  protein synthesis post-translational processing and, 691–694
    oligosaccharide chains, 694F
    secretory pathways, 691–692F
    signal peptide, 691–692F
  proteins, classes of, 10F, 270–273F
    α helix, 270–271F
    β barrel, 271–272F
    integral (transmembrane), 270–272F
    lipid anchored, 272–273F
    number and variety of proteins and lipids in, 273–274F
    peripheral, 272
  secretions, oligosaccharides and, 252F
  signal transduction across, 283–291
    adenylyl cyclase signaling pathway, 287–288F
    G proteins, 285–286F, 290
    inositol-phospholipid signaling pathway, 287–289F
    receptor tyrosine kinases, 290–291F
    receptors, 283–285
    signal transducers, 285–286
  solubility and, 34–35
    structure of, 10F
  thylakoid, 457–460F
  transport, 277–283
    active, 280–283F
    adenosine triphosphate (ATP), 282–283F
    channels for (animal), 279–280F
    characteristics of, 279T
    constant, K_{tr}, 281–282F
    endocytosis and exocytosis, 283–284F
    Gibbs free energy change, ΔG, 278–279
    molecular traffic and, 277–278
    passive, 280–282F
    permeability coefficients, 278–279F
    pores for (human), 279–280F
    potential, $\Delta\psi$, 279–280F
    proteins, 279–282
    thermodynamics and, 278–279
menaquinone, 220F
Mendel, Gregor, 270, 447, 469
Mendelian Inheritance in Man (MIM) numbers, 381–382
Menten, Maud L., 143
Meselson, Matthew, 601
messenger RNA, *see* mRNA
metabolic charts, 297F
metabolic pathways, 297–302
  defined, 297
  evolution of, 301–302
  feedback inhibition, 300
  feed-forward activation, 300
  flux in, 300F
  forms of sequences, 297–298F
  glycolysis, 325–354
  glucogenesis, 354–384
  regulation of, 299–301
  single and multiple steps of, 298–299F
  steady state in, 300F
metabolic precursors, 360–363, 529–532
  amino acids as, 529–532
  gluconeogenesis, 360–363
metabolism, 11, 198–200T. *See also* glycolysis; gluconeogenesis; metabolic pathways
  adenosine triphosphate (ATP), 198–199F, 304, 308–315
  allosteric enzyme phenomena, 153–154
  amino acids, 514–549
  amphibolic reactions, 295
  anabolic (biosynthetic) reactions, 294–295F, 302–303F
  autotrophs, 302–303
  bacteria adaptation and, 295–296
  biosynthetic (anabolic) pathways, 302303
  catabolic reactions, 295F, 303–304F
  cellular pathways, 302–304
  citric acid cycle, 303–304
  cobalamin and, 215–216F
  coenzymes, 198–200T, 316–320
  compartmentation, 304–305
  enzyme regulation and, 153–154
  experimental methods for study of, 321–322
  folate (tetrahyfolate) and, 213–214
  fuel, 295
  gene sequences and, 295–296
  Gibbs free energy change, ΔG, 306–312, 317–319
  glucose, 303
  heterotrophs, 302–303
  hydrolysis, 308–312, 316
  intermediary, 294
  interorgan, 304–305
  irreversible changes, 308–312
  lipids, 475–513
  nucleotide coenzymes and, 198–200
  nucleotides, 550–572
  nucleotidyl group transfer, 315F
  oxidation and, 303–304, 316–321
  phosphoryol group transfer, 312–315
  reaction network of, 294–297
  thioesters, 316
metabolite channeling, 158–159
metal-activated enzymes, 197
metalloenzymes, 197
methanol, 238F
methionine (M, Met), 60F, 216F
  catabolism by conversion of, 539–540F
  nomenclature, 64T
  residue, 76
  structure of, 60F, 216F
  synthesis of, 520–522F
methotrexate, structure of, 550
methylation, 560–564F
  cycle of reactions, 563F
  deoxyuridine monophosphate (dUMP) formation by, 560–564F
  nucleotide metabolism and, 560–564F
  restriction endonucleases catalysis by, 593, 595F
methylmalonyl CoA, 125–126F
Meyerhof, Otto, 331
micelles, 36F
Michaelis, Leonor, 142
Michaelis–Menton equation, 140–144
microheterogeneity, 248
microtubules, 23
Miescher, Friedrich, 573
mirror-image pairs of amino acids, 57F
Mitchell, Peter, 420
mitochondria, 21–22F, 418–421F
  active transport across membrane of, 435–436
  acyl CoA transport into, 497–498
  adenosine triphosphate (ATP) synthesis and, 421F, 435–436
  β-oxidation and, 497–498
  chemiosmotic theory, 420–423
  electron transport and, 435–436
  eukaryotic cell structure and, 20F, 21–22F
  knob-and-stalk structure, 433F
  number of, 418–419
  oxidation from, 21
  oxygen uptake in, 421F
  photosynthesis and, 22
  protonmotive force, 421–420F
  pyruvate entry into, 402–405F
  structure of, 419–420
mitochondrial genomes, 432F
mitosis, 20F
modified ends, mNRA, 658
molecular chaperones, 117–119F
  aggregation prevention by, 119
  chaperonin (GroE), 118–119F
  heat shock proteins, 117–118F
  protein folding assisted by, 117–119F

molecular weight, 6
molecular weight, amino acids and, 74–75T
Monod, Jacques, 157, 635
monolayers, 36F
monosaccharides, 227–236
  abbreviations for, 236T
  aldoses, 228–234F
  amino sugars, 235–236, 237F
  ball-and-stick models of, 228F, 235F
  boat conformations, 235F
  chair conformations, 235F
  chiral compounds, 228–230F
  conformations of, 234–235F
  cyclization of, 230–234
  deoxy sugars, 235
  derivatives of, 235–236F
  endo-envelope conformations, 234F
  epimers, 230
  Fischer projections of, 228–232F
  Haworth projections of, 232–235F
  ketoses, 228–234F
  sugar acids, 236, 238F
  sugar alcohols, 236, 237F
  sugar phosphates, 235
  trioses, 226
  twist conformation, 234F
monosaccharines, sucrose cleaved to, 348
Morse code, 667F
motifs (supersecondary structures), 100–101F
mRNA (messenger RNA), 9, 587, 658–663
  cap formation, 658–659F
  eukaryotic processing, 656, 658–663
  exons, 660
  genetic code and, 666–667F
  intron/extron gene organization, 660–662F
  introns, 658
  modified ends, 658
  polycistronic molecules, 679
  polydenylation of, 658, 660F
  protein synthesis and, 666–667F, 669–671F
  reading frames, 666–667F
  spliced precursors, 658–663
  spliceosomes, 662–663F
  tRNA anticodons base-paired with codons of, 669–671F
  wobble position, 670–671F
mucin secretions, 252F
multicellular organisms, metabolic pathways in, 305F
multienzyme complexes, 158–159
multifunctional enzymes, 158–159
multistep pathways, 298–299F
multisubstrate enzyme reactions, 147–148F
mutagenesis, site-directed, 167, 186
Mycobacterium tuberculosis, 296
Mycoplasma pneumoniae (*M. pneumoniae*), 108F
myoglobin (Mb), 122–129F
  heme prosthetic group, 122–123F
  oxygen binding, 123–129
  protein structure, study of, 122–129F
  tertiary structure of, 122–123F

N
N-linked oligosaccharides, 249–252F
N-terminus (amino terminus), 68, 74–76F
NADH (reduced nicotinamide adenine dinucleotide), 304, 319–320
  electron transfer from, 319–320, 426–427F
  glycolysis reactions, 334
  metabolic reactions, 304, 319–320
  shuttle mechanisms in eukaryotes, 436–439
NADPH (reduced nicotinamide adenine dinucleotide phosphate) reduction, 466–467
Nagyrapolt, Albert von Szent-Györgyi, 223
near-equilibrium reaction, K_{eq}, 307–308
negatively charged R groups, 62
Neisseria gonorrhea pilin, 105F
Némethy, George, 157
Nephila clavipes, 121
Neurospora crassa, 212, 322
neurotransmitters, signal transduction and, 284
neutral solutions, 43
niacin (vitamin B_3), 200–203F
nicotinamide adenine dinucleotide (NAD), 196F, 200–203F
nicotinamide adenine dinucleotide phosphate (NADP), 200–202F
nicotinamide mononucleotide (NMN), 200–202F
Nirenberg, Marshall, 666
nitric oxide synthesis from arginine, 530–531F
nitrogen (N), 3
nitrogen cycle, 515–517F
nitrogen fixation, 515
nitrogenases, 516–517
Nøby, Jens G., 44
noncompetitive inhibition, 149–151F
noncovalent interactions, 37–40F
  charge–charge, 37
  hydrogen bonds, 37–38F
  hydrophobic, 39–40F
  ion pairing, 37
  salt bridges, 37F
  van der Waals forces, 38–39F
noncyclic electronic transfer, 452
nonessential amino acids, 514, 529T
nonketotic hyperglycinemia, 544
nonreducing sugars, 238–239
nonsteroid anti-inflammatory drugs (NSAIDS), 486F
norepinephrine, 199F
nuclear magnetic resonance (NMR) spectroscopy, 90, 321
nucleases, 591–598
  alkaline hydrolysis, 591–592F
  DNA, 595–596F
  *Eco*RI and, 595–596F
  endonucleases, 591
  nucleic acid hydrolysis, 591–598
  restriction endonucleases, 593, 595–598
  ribonuclease A, 592–594
  RNA, 591–593F
nucleic acids, 2, 3, 7–9F. *See also* DNA; nucleosides; RNA
  chromatin, 588–591F
  cleavage of, 592F, 594F
  defined, 7
  double-stranded DNA, 579–586F
  functions of, 573–574
  history of, 573
  hydrogen bond sites of, 575–576F
  hydrolysis of, 591–598
    alkaline, 591–592F
    DNA, 593–596F
    *Eco*RI and, 595–596F
    ribonuclease A, 592–594
    RNA, 591–594F
  identification of, 3
  macromolecular structures of, 8–9F
  nucleases of, 591–598
  nucleosides, 575–577F
  nucleosomes, 588–590F
  nucleotides as building blocks, 574–579
    ribose and deoxyribose, 574F
    purines and pyrimidines, 574–575F
    nucleosides, 575–577F
    tautomeric forms, 575–576F
  restriction endonucleases, 593, 595–598
  RNA in cells, 587
  supercoiled DNA, 586–587F
nucleolus, 20
nucleophiles, 39–40
nucleophilic reactions, 39–41
nucleophilic substitution, 163
nucleoside triphosphates, 308–309
nucleosides, 239, 241, 575–577F
  chemical structures of, 575–577F
  glycosides, 239, 241F
  nomenclature, 576–578T
nucleosomes, 588–590F
nucleotide-group-transfer reaction, 604–605
nucleotide metabolism, 550–572
  adenosine 5′-monophosphate (AMP), 550–551F
  adenosine triphosphate (ATP) reactions, 551F
  allosteric regulation of eukaryotic ribonucleotide reductase, 561T
  base nomenclature, 552
  cytidine triphosphate (CTP) synthesis, 559–560F
  deoxythymidylate (dTMP) production, 560–564F
  deoxyuridine monophosphate (dUMP) methylation, 560–564F,
  DNA and RNA modification, 564–565F
  functions of, 550
  guanosine 5′-monophosphate (GMP), 550–551F
  inosine 5′-monophosphate (IMP) synthesis, 551–554F
  5-phosphoribosyl 1-pyrophosphate (PRPP), 551–552F, 555–556
  purine catabolism, 565–568
  purine nucleotides, synthesis of, 550–554F
  purine salvage, 564–565F
  pyrimidine catabolism, 568–570
  pyrimidine salvage, 564–565
  pyrimidine synthesis, 555–559F
  ribonucleotide and deoxyribonucleotide reduction, 560–562F
  salvage pathways, 564–565
  uridylate (UMP) synthesis, 556–557F
nucleotides, 198–199, 574–579
  anti conformation of, 577–578F
  chemical structure of, 574
  coenzyme metabolic roles, 198–199
  double-stranded DNA, 580–581F
  nomenclature, 577–578T
  nucleic acid building blocks, 574–579
    nucleosides, 575–577F
    purines and pyrimidines, 574–575F
    ribose and deoxyribose, 574F
    tautomeric forms, 575–576F
  phosphodiester linkages (3–5′) joining, 580–581F
  sin conformation of, 577–578F
nucleotidyl group transfer, 315F
nucleus, eukaryotic cells, 20
Nyhan, William, 569

O
O-linked oligosaccharides, 249–251F
odd-chain fatty acids, β-oxidation of, 499–500
Ogston, Alexander, 397
Okazaki, Reiji, 608
Okazki fragments, 608–611F
oligomeric protein, RNA polymerase, 363–637
oligomers (multisubunits), 103, 106, 108T
oligonucleotide-directed mutagenesis, 167
oligopeptide, 68
oligosaccharides, 227, 248–252F
  ABO blood group, 250–251F
  chain structure in post-translational processing, 694F
  diversity of chains, 248
  glycosidic subclasses, 249
  membrane secretions and, 252F
  N-linked, 249–252F
  O-linked, 249–251F
  synthesis of, 250–251

Online Mendelian Inheritance in Man (OMIM), 126
organelles, eukaryotic cells, 19–20F
orotidine 5′-monophosphate (OMP), 550–551F
osmotic pressure, solubility and, 34–35
oxidation, 21, 164, 385, 391–394
  acetyl CoA, 385, 391–394
  β-oxidation, 494–501F
  citric acid cycle reactions, 385, 391–394
  defined, 164
  fatty acids, 494–501
  glycerol, 361F
  mitochondria and, 21, 497–498
oxidation–reduction reactions, 164, 200–205, 221
  coenzymes, 200–205, 221, 316–320
  electron transfer from, 316–320
  electron transport and, 423–425T
  enzyme mechanism of, 164
  flavin mononucleotide (FMN), 204–205F
  NADH (reduced NAD), 316–320
  nicotinamide adenine dinucleotide (NAD), 200–203F
  reduction potentials of electron transfer components, 425T
  thioredoxin (human), 221F
oxidoreductases enzymes, 136
oxygen (O), 3, 29F
  sp^3 orbitals, 29F
  polarity of water and, 29F
oxygen binding, 123–129
  Bohr effect, 128F
  allosteric protein interactions, 127–129F
  carbamate adducts, 129F
  conformational changes from, 124–126F
  fractional saturation, 124–125F
  heme prosthetic group reversibility, 123–124
  hemoglobin (Hb), 123–129F
  hydrophic behavior and, 123–124F
  hyperbolic curve and, 124–126F
  myoglobin (Mb), 123–129F
  oxygenation and, 123
  positive cooperativity, 124
  sigmoidal (S-shaped) curves for, 124–126F
oxygen uptake in mitochondria, 421F
oxygenation, Calvin cycle of photosynthesis, 465–466F
oxyhemoglobin, 123
oxymyoglobin, 123

P

P/O (phosphorylated/oxygen) ratio, 436
packing ratio, 588
pancreatic lipase action, 505F
papain, pH and ionization of, 170–172F
parallel β sheets, 97–98F
parallel twisted sheet, domain fold, 106F
Parnas, Jacob, 331
passive membrane transport, 280–282F
Pasteur, Louis, 2, 331
Pasteur effect for glycolysis regulation, 347
Pauling, Linus, 94
pause sites, RNA transcription, 644
Pavlov, Ivan, 183
penicillin, 247–248F
pentose phosphate pathway, 364–369
  oxidative stage, 364–366F
  nonoxidative stage, 364–365F, 366–368F
  transketolase catalysis, 368F
  interconversions, 368–369F
  transaldolase catalysis, 368–369F
pepsin, 183
peptide bonds, 67–68. *See also* proteins
  acid-catalyzed hydrolysis of, 73F
  amino acids and, 67–68, 73F
  hydrolysis of, 40F
  peptide groups, 91–93F
  cis conformation, 91F, 93
  Ramachandran plots for, 92–93F
  rotation of, 91–92F
  trans conformation, 91F, 93
peptidyl transferase catalysis of, 681–682, 683F
polypeptide chains from, 91–93F
protein synthesis and, 681–682, 683F
residues, 67
resonance structure of, 91F
sequencing nomenclature, 68
structure of, 68F
peptidoglycans, 246–248F
peptidyl transferase catalysis of peptide bonds, 681–682, 683F
peptidylprolyl *cis/trans* isomerase (human), 104F
perchlorate (ClO_4), 36
periodic table of elements, 4F
perioxisomes, 20F, 22
peripheral proteins, 272
permeability coefficients, 278–279F
Perutz, Max, 2–3, 88–90, 94
pH, 43–52
  acid dissociation constant, K_a, 44–48T
  acid solutions, 43F
  base solutions, 42–43F
  buffered solutions, 50–52F
  calculation of, 49
  enzymatic rates and, 170–172F
  Henderson–Hasselbach equation for, 46–47
  indicators, 44F
  neutral solutions, 43F
  physiological uses, meter accuracy for, 44
  pK_a relation to 45–48T
  scale, 43–44
  titration of acid solutions, 47–48F
  water relations to, 43T
phase transition of lipid bilayers, 277F
phenylalanine (F, Phe), 59F
  lignin synthesis from, 531–532F
  nomenclature, 64T
  structure of, 59F
  synthesis of, 524–527F
phenylanyl-tRNA, 529F
phenylisothiocyanate (PITC) treatment, 73F
  amino acid treatment, 73F
  Edman reagent for sequencing residues, 74–75F
phenylthiocarbamoyl (PTC)-amino acid, 73F
phosphagens, phosphoryl group transfer, 314–315F
phosphate 4–5F, 8
  ester linkages, 4–5F, 8
  general formula of, 5F
  hydrolyses catalyzation, 137
phosphatidates, 262–264F
  formation of, 481F
  glycerophospholipid functions of, 262–264F
  structure of, 264F
phosphatidylinositol 3,4,5-*tris*phosphate (PIP_3), 290–291F
phosphatidylinositol 4,5-*bis*phosphate (PIP_2), 287–289F
5-phospho-β-D-ribosylamine (PRA), 553F
phosphoanhydride linkages, 4–5F
  general structure of, 4–5F
  nucleic acid structures and, 8F
phosphoarginine, 315F
phosphocreatine, 315F
phosphodiester linkages, 8–9F
  DNA synthesis of, 610, 612F
  nucleic acid structures and, 8–9F
  nucleotides joined by (3–5′) bonds, 580–581F
phosphoenolpyruvate (PEP), 154F, 315F, 338, 403F
phosphoenylpyruvate carboxykinase (PEPCK) reactions, 358F, 403
phosphofructokinase, 154–155F
phosphofruktokinase-1 (PFK-1), 330
  bacterial enzyme evolution, 364F
  catalysis, 330
  gluconeogenesis regulation, 363–364F
  glycolysis catalysis of, 330
  glycolysis regulation of, 345–346F
phosphoglycerate kinase catalysis, 335–336
phosphoglycerate mutase catalysis, 336–337F
2-phosphoglycolate, 180–181F
5-phosphoribosyl 1-pyrophosphate (PRPP), 551–553F, 555–556
phospholipids, 256
phosphopantetheine, 205–206F
phosphoric acid (H_3PO_4), titration of, 48
phosphorolysis, 371–376
  glycogen reaction, 371–372F
  glycogen regulation, 372–376
phosphorus (P), 3
phosphoryl, general formula of, 5F
phosphoryl group transfer, 312–315
phosphorylated state (GPa), glycogen phosphorylase, 375F
phosphorylation, protein synthesis regulation by, 687–688F
photoautotrophs, 303
photodimerization (direct repair), 622–623
photoheterotrophs, 303
photons (energy), 445–446
photosynthesis, 11F, 22, 439, 443–474
  atmospheric pollution and, 457
  bacterial photosystems, 448–458
    coupled, 453–455T
    cytochrome *bf* complex, 453–455F
    Gibbs free energy change, Δ*G*, 455–457
    internal membranes, 457
    photosystem I (PSI), 448, 450–453F
    photosystem II (PSII), 448–450F
    reaction equations, 450T, 452T, 455T
    reduction potentials, 455–457F
  biochemical process, 11F
  C_4 pathway, 469–471F
  Calvin cycle, 443, 461–467F
  carbon dioxide (CO_2) fixation, 461–467, 469–472
  carboxysomes, 469–470F
  cell structure, 22
  crassulacean acid metabolism (CAM), 471–472F
  dark reactions, 443
  electron transport compared to, 439
  energy flow, 11F
  eukaryotic (plant) photosystems, 458–461
    ATP synthase, 459–460F
    chloroplasts, 458–460F
    cyanobacteria evolution of, 459
    organization of components, 459–460F
  functions of, 443–444
  light-gathering pigments, 444–448
    accessory pigments, 447–448F
    chlorophylls, 444–447F
    photons (energy), 445–446
    resonance energy transfer, 446
    special pair, 446–447F
  light reactions, 443
  starch metabolism (plants), 467–469F
  sucrose metabolism (plants), 467–469F
photosystems, 448–461
  bacterial, 448–458
    coupled, 453–455T
    cytochrome *bf* complex, 453–455F
    Gibbs free energy change, Δ*G*, 455–457
    internal membranes, 457
    photosystem I (PSI), 448, 450–453F
    photosystem II (PSII), 448–450F
    reaction equations, 450T, 452T, 455T
    reduction potentials, 455–457F

photosystems (*Continued*)
eukaryotic (plant), 458–461
ATP synthase, 459–460F
chloroplasts, 458–460F
cyanobacteria evolution of, 459
organization of components, 459–460F
grana, 458
lumen, 458
stroma, 458
thylakoid membranes, 457–460F
Z-scheme, 455–456F
phycoerythrin, 447
phylloquinone (vitamin K), 218–219F
phylogenetic tree representation, 79–80F
Physeter catodon oxymyoglobin, 122F
Pin1 protein, 93
ping-pong enzyme reactions, 148–149F
pK_a, 45–48T, 63–67
acid dissociation parameter values, 45–48T
amino acids, ionization of and, 63–67F
buffer capacity and, 50–52F
free amino acid values, 66T
ionizable amino acid values, 168T
pH relation to, 45–48T
titration and, 47–48F, 64–65F
plasma, lipoproteins in, 508T. *See also* blood plasma
plasma membrane, 457F
plasmalogens, 263, 265F
plastoquinone, 220F
pleated β sheets, 97–98
polar substances, solubility of, 32–35
polarity of water, 29F
poly A tail, 658
polyacrylamide gel electrophoresis (PAGE), 70–71
polydenylation of mNRA, 658, 660F
polylinker, 597
polymerase chain reaction (PCR), 615–617F
polymerases, 603–615, 636–638
chain elongation, 604–606F, 637–638F
DNA replication and, 603–615
eukaryotic, 620T, 646T
interactions, 111F
nucleotide-group-transfer reaction, 604–605
proofreading for error correction, 607
protein types, 603–604T
RNA, 636–638
catalyzation by, 637–638F
chain elongation reactions, 637–638F
conformation changes, 642
eukaryotic factors, 646–648T
oligomeric protein, 363–637
transcription, 642, 646–648T
synthesis of, 607–615
binding DNA fragments, 609–611F
discontinuous, 608F
Klenow fragment, 609–610F
lagging DNA strands, 608–609F, 613–614F
Okazki fragments, 608–611F
phosphodiester linkage, 610, 612F
RNA primer for, 608–609
single-strand binding (SSB) protein, 613F
two DNA strands simultaneously, 607–615
polymers, 4–10
macromolecular structure of, 4–10
lipids, 9
membranes, 9–10
nucleic acids, 7–9F
proteins, 6
polysaccharides, 6–7F
polynucleotide, 7
polypeptides, 7, 68. *See also* proteins
polypeptide chains, 85–87, 91–93F
β strand and sheet structures, 97–99F
cotranslational modifications, 690–691
folding structures for protein stability, 99–101F
peptide bonds in, 91F
peptide groups in, 91–93F
post-translational modifications, 690–691
protein structure from, 85–87
protein synthesis modifications, 690–691F
polysaccharides, 6–7F. *See also* carbohydrates
cellulose, 243F
chitin, 244F
glycogen, 240–243F
heteroglycans, 240
homoglycans, 240
lysozyme catalyzation of, 189–190F
micromolecular structures of, 6–7F
starch, 240–242F
structure of, 240–241T
polyunsaturated fatty acids, 258, 260F
pores for (human) membrane transport, 279–280F
positive cooperativity, 124
positively charged R groups, 61–62
post-transcriptional RNA modification, 655–657F
post-translational processing, 689–694
glycosylation of proteins, 694F
oligosaccharide chains, 694F
polypeptide chain modifications, 689–694F
protein synthesis, 689–694
secretory pathways, 691–692F
signal hypothesis, 691–694
signal peptide, 691–692F
signal recognition particle (SRP), 691–693F
potassium (K), 3
prenylated protein membranes, 272
primary active membrane transport, 282
primary protein structure, 67, 79–81. *See also* amino acids
prochiral substrate binding, 397
prokaryotes, evolution and, 15–16F
prokaryotic cells, 17–18F
E. coli, 17F
ribosomes, eukaryotic cells compared to, 674–675F
structure of, 17–18F
proline (P, Pro), structure of, 59F
nomenclature, 64T
structure of, 59F
synthesis of, 523F
promoter recognition, RNA transcription, 641–642
promoter sequences, RNA transcription, 640–641F
proofreading for DNA replication error correction, 607, 674
propionate, gluconeogenesis presursor, 361–362
prostaglandins, 268–269
lipid structure and functions, 268–269
synthesis of, 483, 485–486F
prosthetic groups, 122, 197
biotin (vitamin B_7), 211–212F
coenzyme behavior of, 197
cytochromes, 221–222F
defined, 122
heme, 122–126F, 221–222F
oxygen binding in, 123–126F
oxygenation and, 122
phosphopantetheine, 205–206F
pyridoxal phosphate (vitamin B_6), 207–209F
proteasome from yeast, 534F
Protein Data Bank (PDB), 89–90, 116
protein disulfide isomerase (PDI), 113–114
protein machines, 108–109F
protein synthesis, 665–696
aminoacyl-tRNA synthetases, 670–673F
antibiotic inhibition of, 686F
anticodons, 668–671F
codons, 665–670T, 679–684F
energy expense of, 684–685
genetic code, 665–668T
mRNA (message RNA), 666–667F, 669–671F
post-translational processing, 689–694
glycosylation of proteins, 694F
oligosaccharide chains, 694F
polypeptide chain modifications, 689–694
secretory pathways, 691–692F
signal hypothesis, 691–694
signal peptide, 691–692F
signal recognition particle (SRP), 691–693F
regulation of, 685–690
attenuation, 688–689F
globin, 687–688F
heme availability and, 687–688F
ribosomal assembly in *E. coli*, 685–687F
trp operon in *E. coli*, 688–690F
ribosomes, 673–681F, 685–687
translation, 673–684
aminoacyl-tRNA docking sites for, 680–681F
chain elongation, 679–684F
elongation factors, 680–681F
eukaryotes, initiation in, 679
initiation of, 675–679F
microcycle steps for, 679–684
peptidyl transferase catalysis, 681–682, 683F
ribosomes, 673–674
Shine-Delgarno sequence, 677F, 679
termination of, 684
translocation of ribosome, 682–684F
tRNA (transfer RNA), 665–671F, 675–681F
protein turnover, 531–533
proteins, 6–7F, 55–133
α helix, 94–97F, 98–99
allosteric, 127–129F
amino acids and, 6F, 55–84
analytical techniques, 70–74
chromatography, 73–47F
mass spectrometry, 72–73F
polyacrylamide gel electrophoresis (PAGE), 70–71F
antibody binding to specific antigens, 129–130
β strands and sheets, 97–99F
biological functions of, 55–56, 119–129
classes of membrane proteins, 10F, 270–273F
coenzymes, 221
cytochrome *c* sequences, 79–81F
denaturation, 110–114F
diffusion of lipids, 275–276F
enzymes as, 6–7F
evolutionary relationships, 79–81
fibrous, 86, 119–121
folding and stability of, 99–103F, 114–119F
CASP, 116
characteristics of, 114–115F
charge–charge interactions and, 117
hydrogen bonding and, 115–116F
hydrophobic effect and, 114–115
molecular chaperones and, 117–119F
tertiary protein structure and, 99–103
van der Waals interactions and, 117
globular, 86, 122–129
glycosylation of, 694F
homologous, 79
hydrolysis of, 40F, 73–74F, 533F
linkages, 4–5F
loop and turn structures, 98–99F
macromolecular structures of, 6–7F
membranes, 10F, 270–273F
active transport, 280–283F
channels for transport (animal), 279–280F
integral (transmembrane), 270–272F
lipid anchored, 272–273F
number and variety of proteins and lipids in, 273–274F
passive transport, 280–282F
peripheral, 272
pores for transport (human), 279–280F

oxygen binding to myoglobin and hemoglobin, 123–129
peptide bonds, 40F, 67–68F, 91–93F
phylogenetic tree representation, 79–80F
polypeptide chains, 85–87, 91–93F, 99–101F
primary structure of, 67, 79–81
protein–protein interactions, 109–111
purification techniques, 68–70
quaternary structure of, 88, 103, 106–109F
renaturation, 112–113F
secondary structure of, 87
sequencing strategies, 74–79
cleavage by cyanogen bromide (CNBr), 76–77F
Edman degradation procedure, 74–75F
human serum albumin, 78–79F
mass spectrometry, 77–78F
structure of, 85–133
binding of antibodies to antigens, 129–130F
collagen, study of, 119–121F
conformations of, 91–98, 110–114
hemoglobin (Hb), study of, 122–129F
levels of, 87–88, 99–109
loops and turns, 98–99F
methods for determining, 88–90
myoglobin (Mb), study of, 122–129F
peptide group, 91–93F
subunits, 103, 106–109F
tertiary structure of, 87F, 99–106F
ubiquitination of, 533F
UV absorbance of, 60F
proteoglycans, 244–246F
proton leaks and heat production from ATP synthesis, 435
protonmotive force, 421–420F
proximity effect, 176–178F
psicose, 231F
pterin, 213–214F
purine, 8–9F, 574–575
catabolism of, 565–568
nucleotide structure, 574–575F
ring structure, 551–552F
salvage pathways, 564–565F
synthesis of nucleotides, 550–554F
nucleotides, 8–9F
puromycin, protein synthesis and, 686F
purple bacteria, photosynthesis in, 448–450F
pyranos, 231F, 234
pyridoxal (vitamin B_6), 207–209F
pyridoxal phosphate (PDP), 207–209F
pyrimidine, 8–9F, 574–575
catabolism of, 568–570
nucleotide structure, 574–575F
regulation of synthesis, 559
salvage pathways, 564–565
synthesis of, 555–559F
pyrophasphate, hydrolyses catalyzation, 137
pyrrolysine, structure of, 62–63F
pyruvate, 136–137, 315F, 338–340F, 387–391F
acetyl CoA, conversion to, 385, 387–391F
alanine, conversion to, 361F
citric acid cycle reactions, 385, 387–391F
gluconeogenesis conversion of, 356–360
gluconeogenesis precursor, 361
gluconeogenesis regulation, 363
glucose conversion from, 338–340F, 357–360F
glycolysis conversion of, 338–340F
lyases catalyzation, 137
metabolism to ethanol, 339–340F
mitochondria, entry into, 402–405F
oxireductases catalyzation, 136
oxidation of, 338–339F
polypeptide folding of, 101
transferases catalyzation, 136–137
pyruvate carcoxylase reaction, 357–358F
pyruvate dehydrogenase phosphorylase kinase (PDHK), 408F
pyruvate dehydrogenase structural core, 108F
pyruvate kinase, 101, 338, 346–347F
glycolysis catalysis of, 338
glycolysis regulation of, 346–347F
reduction to lactate, 340

Q

Q-cycle electron pathway, 430
quaternary protein structure, 88, 103, 106–109F
Escherichia coli (*E. coli*) oligomeric proteins, 108T
examples of, 107F
oligomers (multisubunits), 103, 106, 108T
protein machines, 108–109F
subunits, 103, 106–109F

R

R group amino acids, *see* side chains
R (relaxed) state, 126
racemization, 58
Racker, Efriam, 461
Ramachandran plots, 92–93F
Ramachandran, G. N., 92, 119
rate (velocity) equations, 138–139, 144–145
reaction coordinates, 165–166F
reactions, metabolic network of, 294–297F
reactive center, 196
reading frames, 666–667F
receptors, 283–285
recombinant DNA, 597–598F
recombination, *see* homologous recombination
reduced nicotinamide adenine dinucleotide, *see* NADH
reducing sugars, 238–239
reduction, 164. *See also* oxidation–reduction
Calvin cycle of photosynthesis, 466–467
defined, 164
deoxyribonucleotide, 560–562F
ribonucleotide, 560–562F
reduction potential, 317–319T, 425T, 455–457T
coenzymes, 317–319T
electron transport oxidation–reduction components, 425T
photosynthesis, 455–457F
reductive pentose phosphate cycle, *see* Calvin cycle
regeneration, Calvin cycle of photosynthesis, 466–467F
regulation, 153–158, 343–347, 363–364. *See also* inhibition
citric acid cycle, 406–407F
enzyme activity, 153–158
allosteric enzymes, 153–158F
concerted (symmetry) model for, 156–157F
cooperative binding and, 156F
covalent modification, 158F
phosphofructokinase, 154–155F
sequential (KNF) model for, 157–158F
sigmoidal (S shaped) curves for, 153F, 156F
gluconeogenesis, 363–364F
glycolysis, 343–347
hexokinase, 344–345
hexose transports, 343–344
metabolic pathway in mammals, 343F
Pasteur effect for, 347
phosphofruktokinase-1 (PFK-1), 345–346F
pyruvate kinases, 346–347F
hormones for, 502–504
IUMBM–Nicholson metabolic chart, 504F
lipid metabolism, 502–504
protein synthesis, 685–690
attenuation, 688–689F
globin, 687–688F
heme availability and, 687–688F
ribosomal assembly in *E. coli*, 685–687F
trp operon in *E. coli*, 688–690F
relative molecular mass, 6
renal glutamine metabolism, 547–548
renaturation, 112–113F
replisome, defined, 603
replisome model, 610, 612–615
residues, 5, 67–68F, 74–75F
amino acids, 67–68F, 74–75F, 166–168T
β strand and sheet turns, 99F
catalysis and, 166–168T
catalytic frequency distribution, 168T
collagen and formation of, 120–121F
Edman degradation procedure for, 74–75F
glycogen, cleavage of, 371–372F
ionizable amino acid functions, 166–168T
macromolecule structure of, 5
methionine, 76
peptide bond linkages, 67–68F
phenylisothiocyanate (PITC) treatment, 74–75F
pK_a values of ionizable amino acids, 168T
protein structure and, 120–121F
sequences of, 68, 74–75F
resonance energy transfer, 446
resonance stabilization, 310
respiration process, 340
restriction endonucleases, 593, 595–598
defined, 593
DNA and, 593,595–598
DNA fingerprints, 596–597F
hydrolysis and, 593,595
methylation, 593, 595F
nucleic acids and, 593, 595–598
recombinant DNA, 597–598F
restriction maps, 596
specificities of, 595T
types I and II, 593, 595
restriction maps, 596
retinol (vitamin A), 217–218F
retinol-binding protein (pig), 104F
reverse turns, protein structures, 99
reversible inhibition, 148–152F
rho-dependent RNA transcription termination, 644–645F
Rhodopseudomonas photosystem, 107F
Rhodospirillum rubrum, 484
riboflavin, 204–205F
ribofuranose, 233F
ribonuclease A (Rnase A), 90F, 111–113F
denaturation and renaturation of, 112–113F
disulfide bridges in, 112F
heat denaturation of, 111F
hydrolysis by, 592–594
ribonucleic acid, *see* RNA
ribopyranose, 233F
ribose, 7, 229F, 236F, 574F
cyclization of, 232–233F
monosaccharide structures of, 229F, 236F
nucleotide structure, 574F
sugar phosphate structure, 236F
ribosomal RNA, *see* rRNA
ribosomes, 108F, 673–681F
aminoacyl-tRNA binding sites in, 675, 677F
chain elongation and, 673–674, 682–684F
eukaryotic versus prokaryotic cells, interactions, 111F
protein synthesis, 673–681F, 685–687
regulation of protein synthesis, 685–687F
rRNA composition of, 674–675F
translocation by one codon, 682–684F
ribulose, 230–231F
ribulose 1,5-*bis*phosphate, 465–466F
ribulose 5-phosphate conversion, 367F
right turn structures, 98–99F

RNA (ribonucleic acid), 3, 9, 634–664
cell content, 587
classes of, 587
cleavage, 594F, 655–657F
discovery of, 3
eukaryotic mRNA processing, 656, 658–663
hydrolysis, 591–594
alkaline, 591–592F
nucleases and, 591–594
ribonuclease A, 592–594F
lac operon, 651–655
binding repressor to the operon, 652F
cAMP regulatory protein and, 653–655F
repressor blocking transcription, 651–652F
repressor structure, 652–653F
transcription activation, 653–655
messenger (mRNA), 9, 587, 656, 658–663
modified nucleotides, 564–565F
molecule types, 9
polymerase, 108F, 111F, 636–638
catalyzation by, 638F
chain elongation reactions, 637–638F
interactions, 111F
multisubunit, 108F
oligomeric protein, 363–637
post-transcriptional modification of, 655–657
ribosomal (rRNA) processing, 656–657F
transfer (tRNA) processing, 655–657F
ribosomal (rRNA), 9, 587, 656–657F
small nuclear (sRNA), 662–663F
stem–loop structures, 587–588F
synthesis of, *see* transcription
transfer (tRNA), 9, 587, 655–657F
types of, 635–636
RNA polymerase, 108F, 111F
RNA primer for DNA synthesis, 608–609
RNA transcription, 639–651
cAMP regulatory protein activation of, 653–655
eukaryotes, 646–649
chromatin and, 649
polymerase reactions, 646–648T
transcription factors, 648–649T
gene regulation, 649–651
initiation, 639–643
α subunits, 641–642T
gene orientation, 639–640F
polymerase changes in conformation, 642
process of, 643F
promoter recognition, 641–642
promoter sequences, 640–641F
lac repressor blockage of, 651–652F
termination, 644–645
hairpin formation, 644F
pause sites, 644
rho-dependent, 644–645F
rofecoxib (Vioxx), structure of, 486F
Rose, Irwin, 533
rRNA (ribosomal RNA), 9, 587, 656–657
cleavage, 656–657F
post-transcriptional modification, 656–657F
protein synthesis and, 674–675F
ribosome composition of, 674–675F
RS amino acid system configuration, 61F
rubisco (rubilose 1,5-*bis*phosphate carboxylase-oxygenase), 462, 464–466F

S

S-adenosylmethionine, 199F
saccharides, *see* carbohydrates; polysaccharides
Saccharomyces cerevisiae, 296F
salicylates, 486
Salmonella typhymurium, 514F, 528F
salt bridges, 37F
salvage pathways, 564–565
Sanger method for DNA sequencing, 616, 618
Sanger, Frederick, 616
saturated fatty acids, 258, 260F
Schiff bases, 121F, 208F, 332–333F
scurvy, ascorbic acid and, 209–210
seawater, properties of, 33F
second messengers, 285
secondary active membrane transport, 282, 283F
secondary protein structure, 87
secretory pathways, 691–692F
selenocysteine, structure of, 62–63F
semiconservative DNA replication, 602F
semiquinone anion, 220F
sequencing, 68, 74–81, 616–619
amino acid residues, 68, 74–75F
C-terminus (carboxyl terminus), 68, 76F
cytochrome *c*, 79–81F
DNA, 77F, 616–619F
dideoxynucleotides used for, 616, 618
parallel strands by synthesis, 618–619
Sanger method, 616, 618
Edman degradation procedure for, 74–77F
evolution relationships and, 79–81F
human serum albumin, 78–79F
N-terminus (amino terminus), 68, 74–76F
protein strategies, 76–79F
cleavage by cyanogen bromide (CNBr), 76–77F
human serum albumin, 78–79F
mass spectrometry, 77–78F
tryptic fingerprint, 77–79F
sequential enzyme reactions, 148–149F
sequential (KNF) model for enzyme regulation, 157–158F
serine (S, Ser), 56–57F, 60–61F
catabolism of, 536–537F
metabolic precursor use, 529–530F
nomenclature, 64T
RS amino acid system configuration, 61F
structure of, 56–57F, 60–61F
synthesis of, 523–524F
serine proteases, 183–189F
catalytic triad, 185F
catalysis modes for, 185–188
chymotrypsin, 183–188F
elastase, 183–185F
substrate binding, 186–188F
substrate specificity of, 184–185
trypsin, 183–185F
zymogens as inactive enzyme precursors, 183–184
serum albumin (human), 78–79F, 104F, 508
Shine-Delgarno sequence, 677F, 679
shuttle mechanisms, 436–439F
malate–aspartate shuttle, 348F
NADH in eukaryotes, 436–439F
shuttle mechanisms in eukaryote, 437F
side chains, 56, 59–62
alcohol groups with, 60–61
aliphatic R groups, 59
α helix proteins, 95
amino acid structure and, 56, 59
aromatic R groups, 59–60
hydrophic effect on, 114–115
hydrophobicity of amino acids with, 62
ionic states of, 64–65F
negatively charged R groups, 62
positively charged R groups, 61–62
protein folding and, 115–116
sulfur-containing R groups, 60
sigmoidal (S-shaped) curves, 124–126F, 153F, 156F
signal hypothesis, 691–694
signal peptide, 691–692F
signal recognition particle (SRP), 691–693F
signal transduction, 283–291
adenylyl cyclase signaling pathway, 287F
G proteins, 285–286F, 290
hormones receptors and binding, 284–287
hydrolysis and, 285–289F
inositol-phospholipid signaling pathway, 287–289F
insulin receptors, 290–291F
membrane cells, 283–291
pathways, 284–285, 287–289F
receptor tyrosine kinases, 285, 290–291F
receptors, 283–285
transducers, 285–286
sin conformation of nucleotides, 577–578F
single step pathways, 298–299F
single-strand binding (SSB) protein, 613F
single-strand DNA, 588
site-directed mutagenesis, 167, 186
small nuclear ribonucleic acid (snRNA), 662–663F
Smith, Michael, 167
sn-glycerol 3-phoshphate, 484
Söderbaum, H. G., 196
sodium (Na), 3
sodium chloride (NaCl), 33F, 37
sodium dodecyl sulfate (SDS), 36F
sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), 71F
sodium palmitate, 36
solubility, 32–36
amphipathic molecules, 36
cellular concentrations, 34F
chaotropes, 36
detergents, 36F
diffusion, 34F
electrolytes, 32–34
hydrated molecules, 34
hydrophilic substances, 32
hydrophobic substances, 35
ionic and polar substances, 32–35
nonpolar substances, 35–36
osmotic pressure, 34–35
solvated molecules, 34
surfactants, 36
water and, 32–36
solubilization, 36
solvated molecules, 34
solvation effects, 309–310
sorbitol conversion from glucose, 362G
sorbose, 231F
Sørensen, Søren Peter Lauritz, 44
sp^3 orbitals, 29F
space-filling models, 90F
DNA, 573F, 582–584F
proteins, 90F
special pair, 446–447F
specific heat of water, 31
sphingolipids, 263–266F
cerebrosides, 265, 266F
ceremide, 264, 265F
gangliosides, 265, 266F
genetic defects and, 265–266
pathways for formation and degradation of, 492F
sphingomyelins, 264, 265F
synthesis of, 488–489F
sphingomyelins, 264, 265F
spider silk strength, 121
spindle fibers, 56
spliced precursors, mNRA, 658–663
spliceosomes, 662–663F
squalene, cholesterol and, 488, 490
stacking interactions, double-stranded DNA, 582–583F, 585T
Stahl, Franklin, 601
Staphylococcus aureus (*S. Aureus*), 76, 247–248F
starch, 240–243
amylase, 242F
amylose, 241F
amylopectin, 241–242F
digestion of, 241–242

glucose storage (plants), 240–243F
metabolism (plants), 467–469F
structure of, 240–241T
synthesis of, 467–468F
starch, 240–243, 467–469
steady state, metabolic pathways, 300F
steady–state derivation, 141–142
stem length mutation, 270
stem–loop structures in RN, 587–588F
stereochemical numbering, 484
stereoisomers, 56, 59F
stereospecifity, 134–135
steroids, 9, 266–268F
cholesterol and, 266–268
isoprene structure of, 266F
lipid structures of, 266–267F
micromolecular structure of, 9
signal transduction and, 285
Strandberg, Bror, 89
Streptococcus pneumoniae, 3
Streptomyces, potassium channel protein, 107F
stroma, 458
substrates, 90F, 134–148, 175–182
binding properties, 139–140, 176, 178–181F, 185–188
binding sites, 90F, 674F
binding speed, 171–172T
diffusion-controlled reactions, 171–172T
enzymatic catalysis modes and, 175–182
induced fit, 179–180
proximity effect, 176–178F
transition–state stabilization, 176, 180–182F
weak binding and, 176, 179–181F
enzyme kinetics and, 138–148
enzyme reactions, 134–135, 138–147
enzyme–substrate complex (ES), 139–140, 142–143
Michaelis–Menton equation for, 140–144
multisubstrate reactions, 147–148F
prochiral binding, 397
rate (velocity) equations for, 138–139F, 144–145F
serine proteases and, 186–188F
specificity of, 184–185
stereospecifity of, 134–135
subunits, 103, 106–109F
succinate dehydrogenase complex, citrus cycle reactions, 399–401F
succinate:ubiquinone oxidoreductase (electron transfer complex II), 427–428F
Succinyl CoA, 216F
catalyzed structure of, 216F
thioester hydrolysis, 316
succinyl synthetase, citrus cycle reactions, 398–400F
sucralose, 240
sucrose, 238–239F
cleaved to monosaccharines, 348
metabolism (plants), 467–469F
structure of, 238–239F
synthesis of, 467–469F
sugar acids, 236, 238F
sugar alcohols, 236, 237F
sugar phosphates, 235
sugars, 235–236, 238–239
abbreviations for, 236T
disaccharides, 238–239
monosaccharides, 235–236F
nonreducing, 238–239
reducing, 238–239
sulfhydryl, general formula of, 5F
sulfur (S), 3
sulfur-containing R groups, 60
Sumner, James B., 135
supercoiled DNA, 586–587F
superoxide anions, 440–441
superoxide atoms, 440–441
superoxide dismutase, 175F
supersecondary structures (motifs), 100–101F
surfactants, solubility of, 36
sweetness receptors, 240
symport, membrane transport, 280–281F
Synechococcus elongatus, 470F
synonymous codons, 667
synthase, 395
ATP catalysis, 433–435F
defined, 395
glycogen reaction, 370–371F
synthesis, 13
adenosine triphosphate (ATP), 417–442
amino acids, 520–529
cancer drug inhibition of, 564
defined, 13
DNA, two strands simultaneously, 607–615
nucleotide metabolism and, 550–559
proteins, 665–696
purine nucleotides, 550–554F
pyrimidine, 555–559F
synthetase, defined, 395
Système International (SI) units, 26–27T

T

T (tense) state, 126
tagatose, 231F
tail growth, 373
talose, 229F
Tanaka, Koichi, 73
Tatum, Edward, 212, 634
tautomeric forms of nucleic acids, 575–576F
terminal electron acceptors and donors, 439–440
termination (stop) codons, 667F, 682, 684
terpenes, 256
tertiary protein structure, 87F, 99–106F
cytochrome *c* structure conservation, 101F
domains, 101–102, 106F
examples of, 104–105F
hemoglobin (Hb), 122–123F
intrinsically disordered (unstable) proteins, 102–103
motifs (supersecondary structures), 100–101F
myoglobin (Mb), 122–123F
polypeptide folding and stability of, 99–101F
protein stability and, 99–103
supersecondary structures (motifs), 100–101F
tetrahydrofolate, 213–214F
thermodynamics, 12–15, 278–280
activation energy, G‡, 14F
equilibrium constant, K_{eq}, 12, 14
Gibbs free energy change, ΔG, 12–15, 278–279
membrane potential, $\Delta\psi$, 279–280F
membrane transport and, 278–280
reaction rates and, 14–15
Thermus thermophilius, 675, 676F
thiamine (vitamin B_1), 206–207F
thiamine diphosphate (TDP), 206–207F
thiamine pyrophosphate (TPP), 206
Thiobacillus, 303F
thiocyanate (SCN), 36
thioesters, hydrolysis of, 316
thiol (sulfhydryl), general formula of, 5F
thiol-disulfide oxidoreductase, 105F
thioredoxin (human), 105F
coenzyme oxidation-reduction, 221F
oxidized, 221F
structure of, 105F
threonine (T, Thr), 58, 60–61F
catabolism of, 537–538
nomenclature, 64T
structure of, 58, 60–61F
synthesis of, 520–522F
threose, 229
thylakoid membranes, 457–460F
thymine (T), 8–9F
thyroxine, structure of, 63F
titration, 47–48F
acetic acid (CH_3COOH), 47F
acid solutions, 47–48F
amino acids, 64–65F
imazodole ($C_3H_4N_2$), 47F
ionization and, 64–65F
phosphoric acid (H_3PO_4), 48
pK_a values from, 45–48T, 64–65F
TψC arm, 668–669F
trans conformation, 91F, 93, 258, 259F
transaldolase catalysis, 368–369F
transanimation reactions, ammonia assimilation and, 518–519F
transducers, 285–286
bacterial, 285–286
eukaryotic, 285
G proteins, 285–286F
membrane signal transduction and, 285–286
transduction, *see* signal transduction
transfer RNA, *see* tRNA
transferases enzymes, 136–137, 395
transition–state stabilization, 180–182F
transition states, 163, 164–166
activation energy, 165F
catalyst stabilization for, 164–166
defined, 163
enzyme mechanisms and, 164–166
intermediates and, 165–166F
nucleophilic substitution, 163
reaction coordinates, 165–166F
transketolase catalysis, 368F
translation, 673–684. *See also* post-translational processing
chain elongation, 679–684F
aminoacyl-tRNA docking sites for, 680–681F
elongation factors, 680–681F
microcycle steps for, 679–684
peptidyl transferase catalysis, 681–682F
translocation of ribosome, 682, 684F
initiation of, 675–679F
eukaryotes, 679
initiation factors, 675, 677–679
ribosomes, 673–674
Shine-Delgarno sequence, 677F, 679
tRNA initiator, 675, 677F
protein synthesis and, 673–684
ribosomes and, 673–675F, 677F
aminoacyl-tRNA binding sites for, 675, 677F
eukaryotic versus prokaryotic, 674–675F
subunit composition of, 674–675F
Shine-Delgarno sequence, 675F, 679
termination of, 684
transmember (integral) proteins, 270–272F
transport, *see* electron transport; membranes
transport constant, K_{tr}, 281–282F
transverse (flip-flop) diffusion, 275–276F
triacylglycerols, 261–262F
digestion of, 262
structure of, 261F
synthesis of, 481–483F
Trichodesmium, 515F
triene, defined, 486
trifunctional enzymes, β-oxidation and, 498
triiodothryonine, structure of, 63F
triose phosphate isomerase (TPI), 107F, 172–174F
catalysis, 332–334F
diffusion-controlled reactions, 162F, 172–174F
trioses, 226
tripeptide, 68

tRNA (transfer RNA), 9, 587, 655–657, 665–671, 675–681
aminoacyl-tRNA synthetases, 670–673F
anticodons, 668–671F
cleavage, 655–656F
base-pairing, 669–670F
cloverleaf structure, 668–669F
genetic code and, 669–670F
isoacceptor molecules, 670–671
mRNA codons base-paired with anticodons of, 669–670F
post-transcriptional modification, 655–657F
protein synthesis and, 665–671F, 675–681F
three-dimensional (tertiary) structure of, 668–669F, 680
translation initiator, 675–681F
Watson-Crick base pairing, 670F
wobble position, 670–671F
trp operon, protein synthesis regulation by, 688–690F
trypsin, 76–77F, 183–185F
tryptic fingerprint, sequencing and, 77–79F
tryptophan (W, Trp), 58–60F
nomenclature, 64T
structure of, 58–60F
synthesis of, 524–527F
tryptophan biosynthesis enzyme, 105F
turn structures, α helix and β strand and sheet connections, 99F
twist conformations, 234F
type III triple helix, 119F
tyrosine (Y, Tyr), 58–60F
catabolism of, 541–542F
melanin synthesis from, 531, 533F
nomenclature, 64T
structure of, 58–60F
synthesis of, 524–527F

U
ubiquinol, 220
ubiquinol:cytochrome *c* oxidoreductase (electron transfer complex III), 428–430F
ubiquinone (coenzyme Q), 219–221F
ubiquitin, 533F
ubiquitination of proteins, 533F
UDP N-acetylglucosamine acyl transference, 104F
ultraviolet light absorption in double-stranded DNA, 584–585F
uncompetitive inhibition, 149–150F
uncouplers, 420–421F
uniport, membrane transport, 280, 281F
units for biochemistry, 26–27T
unphosphorylated state (GPb), glycogen phosphorylase, 347–375F
unsaturated fatty acids, 258, 260F, 500–501
uracil (U), 8
urea, structure of, 112
urea cycle, 542–547
amino acid metabolism and, 542–547
ancillary reactions to, 547
carbamoyl phosphate synthesis, 543F
conversion of ammonia to urea, 542–547
reactions of, 543–546F
uric acid, 566–569F
uridine diphosphate glucose (UDP-glucose), 200–201F
uridine triphosphate (UTP), 200–201F
uridylate (UMP) synthesis, 556–557F
UV absorbance of proteins, 60F

V
vacuoles, 20F, 22
valine (V, Val), 59F
nomenclature, 64T
structure of, 59F
synthesis of, 521–523F
van der Waals, Johannes Diderik, 38
van der Waals forces, 38–39F
van der Waals interactions, 117
van der Waals radii, 39T
vaporization of water, 32
variable arm, 668–669F
vesicles, 20F, 272F
eukaryotic cells, 22
liposomes, 270F, 272F
specialization, 20F
vitamins, 196, 198–199T
ascorbic acid (vitamin C), 209–211
biotin (vitamin B_7), 211–212F
cobalamin (vitamin B_{12}), 215–216F
deficiencies, 198T, 209–210, 214, 215
fat-soluble, 198
folate (vitamin B_9), 213–214F
functions of, 197–199T
history of, 198
lipid, 217–219F
α-tocopherol (vitamin E), 218F
cholecalciferol (vitamin D), 218–219F
phylloquinone (vitamin K), 218–219F
retinol (vitamin A), 217–218F
niacin (vitamin B_3), 200–203F
pyridoxal (vitamin B_6), 207–209F
sources, 199T
thiamine (vitamin B_1), 206–207F
water-soluble, 198
Voss-Andreae, Julian, 127

W
Walker, John E., 223
Warburg, Otto, 386
warfarin (rat poison), 220F
water, 28–54
acid disolution constants, 44–48
buffered solutions, 50–52
chemical properties of, 28, 39–52
concentration of, 41F
condensation of, 40–41F
hydrogen bonding in, 30–32, 37–38F
ice, formation of, 30–31F
insolubility of nonpolar substances, 35–36
ionization of, 41–43T
noncovalent interactions, 37–40F
charge–charge, 37
hydrogen bonds, 37–38F
hydrophobic, 39–40F
van der Waals forces, 38–39F
nucleophilic reactions, 39–41
pH scale and, 43–44, 49–52
physical properties of, 28–39
polarity of, 29F
solubility of ionic and polar substances, 32–35
specific heat of, 31
vaporization of, 32
water-soluble vitamins, 198
Watson, James D., 3, 573–574, 575, 601
Watson-Crick base pairing, 668–670F
Watson–Crick DNA model, 579, 601
waxes, lipid structure and functions, 9, 268
weak substrate binding, 179–179F
website accuracy, 401
Wilkins, Maurice, 579
Williams, Ronald, 420
Windaus, Adolf Otto Reinhold, 223
wobble position, 670–671F
Wöhler, Friedrich, 2
Wyman, Jeffries, 157

X
X-ray crystallography, 88–90F
X-ray diffraction pattern, 88F
xylose, 229F
xylulose, 231F

Y
yeast, 105F, 345–347F
FMN oxidoreductase, 105F
octamer enzyme, 345–346
proteasome from, 534F
pyruvate kinase regulation by, 347F
Young, William John, 331

Z
Z-DNA, 586F
Z-scheme, photosynthesis path, 455–456F
zwitterions (dipolar ions), 56
zymogens, 183–184

Common Abbreviations in Biochemistry

ACP	acyl carrier protein
ADP	adenosine 5′-diphosphate
AMP	adenosine 5′-monophosphate (adenylate)
cAMP	3′,5′-cyclic adenosine monophosphate
ATP	adenosine 5′-triphosphate
bp	base pair
1,3BPG	1,3-*bis*phosphoglycerate
2,3BPG	2,3-*bis*phosphoglycerate
CDP	cytidine 5′-diphosphate
CMP	cytidine 5′-monophosphate (cytidylate)
CoA	coenzyme A
CTP	cytidine 5′-triphosphate
DHAP	dihydroxyacetone phosphate
DNA	deoxyribonucleic acid
cDNA	complementary DNA
DNase	deoxyribonuclease
$E°$	reduction potential
$E°'$	standard reduction potential
EF	elongation factor
emf	electromotive force
ETF	electron-transferring flavoprotein
$\mathcal{F}$	Faraday's constant
FAD	flavin adenine dinucleotide
$FADH_2$	flavin adenine dinucleotide (reduced form)
F1,6BP	fructose 1,6-*bis*phosphate
FMN	flavin mononucleotide
$FMNH_2$	flavin mononucleotide (reduced form)
F6P	fructose 6-phosphate
ΔG	actual free-energy change
$\Delta G°'$	standard free-energy change
GDP	guanosine 5′-diphosphate
GMP	guanosine 5′-monophosphate (guanylate)
cGMP	3′,5′-cyclic guanosine monophosphate
G3P	glyceraldehyde 3-phosphate
G6P	glucose 6-phosphate
GTP	guanosine 5′-triphosphate
H	enthalpy
Hb	hemoglobin
HDL	high density lipoprotein
HETPP	hydroxyethylthiamine pyrophosphate
HPLC	high-pressure liquid chromatography
IDL	intermediate density lipoprotein
IF	initiation factor
eIF	eukaryotic initiation factor
IMP	inosine 5′-monophosphate
IP_3	inositol 1,4,5-*tris*phosphate
K_a	acid dissociation constant
k_{cat}	catalytic constant
K_{eq}	equilibrium constant
K_m	Michaelis constant
kb	kilobase pair
LDL	low density lipoprotein
LHC	light-harvesting complex
M_r	relative molecular mass
Mb	myoglobin
$NAD^{\oplus}$	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide (reduced form)
$NADP^{\oplus}$	nicotinamide adenine dinucleotide phosphate
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
$NMN^{\oplus}$	nicotinamide mononucleotide
NDP	nucleoside 5′-diphosphate
NMP	nucleoside 5′-monophosphate
NTP	nucleoside 5′-triphosphate
dNTP	deoxynucleoside triphosphate
P_i	inorganic phosphate (or orthophosphate)
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
2PG	2-phosphoglycerate
3PG	3-phosphoglycerate
PEP	phosphoenolpyruvate
PFK	phosphofructokinase
pI	isoelectric point
PIP_2	phosphatidylinositol 4,5-*bis*phosphate
PLP	pyridoxal phosphate
PP_i	inorganic pyrophosphate
PQ	plastoquinone
PQH_2	plastoquinol
PRPP	5-phosphoribosyl 1-pyrophosphate
PSI	photosystem I
PSII	photosystem II
Q	ubiquinone
QH_2	ubiquinol
RF	release factor
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
rRNA	ribosomal ribonucleic acid
snRNA	small nuclear ribonucleic acid
tRNA	transfer ribonucleic acid
RNase	ribonuclease
snRNP	small nuclear ribonucleoprotein
RPP	reductive pentose phosphate
Rubisco	ribulose 1,5-*bis*phosphate carboxylase-oxygenase
S	entropy
dTDP	deoxythymidine 5′-diphosphate
TF	transcription factor
dTMP	deoxythymidine 5′-monophosphate (thymidylate)
TPP	thiamine pyrophosphate
dTTP	deoxythymidine 5′-triphosphate
UDP	uridine 5′-diphosphate
UMP	uridine 5′-monophosphate (uridylate)
UTP	uridine 5′-triphosphate
v	velocity
V_{max}	maximum velocity
v_0	initial velocity
VLDL	very low density lipoprotein
XMP	xanthosine 5′-monophosphate

Abbreviations for amino acids are given on pages 57–62, and those for major pyrimidine and purine bases are given on page 575.

First position (5′ end)	Second position U	C	A	G	Third position (3′ end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

One- and three-letter abbreviations for amino acids

A	Ala	Alanine
B	Asx	Asparagine or aspartate
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
Z	Glx	Glutamate or glutamine